Time series models 2007

Norwegian University of Science and Technology
Department of Mathematical Sciences

Solutions to problem sheet 1, 2007

Exercise 1.1

a)

Let $S(c) = E[(Y-c)^2]$. Then
$$S(c) = E(Y^2) - 2cE(Y) + c^2$$
This gives
$$\frac{dS}{dc} = -2E(Y) + 2c = 0$$
for $c = E(Y)$, which leads to a global minimum since $\frac{d^2S}{dc^2} = 2 > 0$ for all c.

b)

$$\begin{aligned}
E[\big(Y-f(X)\big)^2|X] &= E[\big(Y-E(Y|X)+E(Y|X)-f(X)\big)^2|X] = \\
E[\big(Y-E(Y|X)\big)^2|X] &+ 2E[\big(Y-E(Y|X)\big)\big(E(Y|X)-f(X)\big)|X] + E[\big(E(Y|X)-f(X)\big)^2|X] = \\
E[\big(Y-E(Y|X)\big)^2|X] &+ 2\big(E(Y|X)-f(X)\big)E[\big(Y-E(Y|X)\big)|X] + E[\big(E(Y|X)-f(X)\big)^2|X] = \\
E[\big(Y-E(Y|X)\big)^2|X] &+ E[\big(E(Y|X)-f(X)\big)^2|X] \geq E[\big(Y-E(Y|X)\big)^2|X]
\end{aligned}$$

because $E(Y|X)$ is a function of X and $E(g(X)Y|X) = g(X)E(Y|X)$ for any function g such that $E(g(X)Y)$ exists.

It follows that
$$E[\big(Y-E(Y|X)\big)^2|X] \leq E[\big(Y-f(X)\big)^2|X]$$
for any function f. Hence $E[\big(Y-f(X)\big)^2|X]$ is minimized when $f(X) = E(Y|X)$.

c)

Since
$$E[\big(Y-E(Y|X)\big)^2] = E\Big(E[\big(Y-E(Y|X)\big)^2|X]\Big) \leq E\Big(E[\big(Y-f(X)\big)^2|X]\Big) = E[\big(Y-f(X)\big)^2]$$
it follows immediately that the random variable $f(X)$ that minimizes $E[\big(Y-f(X)\big)^2]$ is $f(X) = E(Y|X)$.

Exercise 1.2

a)

Let $X = (X_1, X_2, \ldots, X_n)$. Then

$$E[\big(X_{n+1} - f(X)\big)^2|X] = E[\big(X_{n+1} - E(X_{n+1}|X) + E(X_{n+1}|X) - f(X)\big)^2|X] =$$
$$E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X] + 2E[\big(X_{n+1} - E(X_{n+1}|X)\big)\big(E(X_{n+1}|X) - f(X)\big)|X]$$
$$+ E[\big(E(X_{n+1}|X) - f(X)\big)^2|X] =$$
$$E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X] + 2\big(E(X_{n+1}|X) - f(X)\big)E[\big(X_{n+1} - E(X_{n+1}|X)\big)|X]$$
$$+ E[\big(E(X_{n+1}|X) - f(X)\big)^2|X] =$$
$$E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X] + E[\big(E(X_{n+1}|X) - f(X)\big)^2|X] \geq E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X]$$

because $E(X_{n+1}|X)$ is a function of X and $E(g(X)X_{n+1}|X) = g(X)E(X_{n+1}|X)$ for any function g such that $E(g(X)X_{n+1})$ exists.

It follows that

$$E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X] \leq E[\big(X_{n+1} - f(X)\big)^2|X]$$

for any function f. Hence $E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X]$ is minimized when $f(X) = E(X_{n+1}|X)$.

b)

Since

$$E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2] = E\big(E[\big(X_{n+1} - E(X_{n+1}|X)\big)^2|X]\big)$$
$$\leq E\big(E[\big(X_{n+1} - f(X)\big)^2|X]\big) = E[\big(X_{n+1} - f(X)\big)^2]$$

it follows immediately that the random variable $f(X)$ that minimizes $E[\big(X_{n+1} - f(X)\big)^2]$ is again $f(X) = E(X_{n+1}|X)$.

c)

By b) the minimum mean-squared error predictor of X_{n+1} in terms of $X = (X_1, X_2, \ldots, X_n)$ when $X_t \sim IID(\mu, \sigma^2)$ is

$$E(X_{n+1}|X) = E(X_{n+1}) = \mu$$

d)

Suppose that $\sum_{i=1}^n \alpha_i X_i$ is an unbiased estimator for μ, that is, $\sum_{i=1}^n \alpha_i = 1$. Then

$$E[\big(\sum_{i=1}^n \alpha_i X_i - \mu\big)^2] = E[\big(\sum_{i=1}^n \alpha_i X_i - \overline{X}\big)^2] + 2E[\big(\sum_{i=1}^n \alpha_i X_i - \overline{X}\big)\big(\overline{X} - \mu\big)] + E[\big(\overline{X} - \mu\big)^2] \geq E[\big(\overline{X} - \mu\big)^2]$$

since the second term is zero: $E[\big(\sum_{i=1}^n \alpha_i X_i - \overline{X}\big)\big(\overline{X} - \mu\big)] = Cov(\sum_{i=1}^n \alpha_i X_i - \overline{X}, \overline{X}) = Cov(\sum_{i=1}^n \alpha_i X_i, \sum_{i=1}^n \frac{1}{n} X_i) - Cov(\sum_{i=1}^n \frac{1}{n} X_i, \sum_{i=1}^n \frac{1}{n} X_i) = \sum_{i=1}^n \frac{\alpha_i}{n}\sigma^2 - \sum_{i=1}^n \frac{1}{n^2}\sigma^2 = 0$.

e)

Again, suppose that $\sum_{i=1}^n \alpha_i X_i$ is an unbiased estimator for μ, that is, $\sum_{i=1}^n \alpha_i = 1$. Then

$$E[\big(X_{n+1} - \sum_{i=1}^n \alpha_i X_i\big)^2] = E[\big(X_{n+1} - \overline{X}\big)^2] + 2E[\big(X_{n+1} - \overline{X}\big)\big(\overline{X} - \sum_{i=1}^n \alpha_i X_i\big)] + E[\big(\overline{X} - \sum_{i=1}^n \alpha_i X_i\big)^2]$$
$$\geq E[\big(X_{n+1} - \overline{X}\big)^2]$$

since the second term is zero: $Cov\big(X_{n+1} - \overline{X}, \overline{X} - \sum_{i=1}^n \alpha_i X_i\big) = -Cov(\overline{X}, \overline{X}) + Cov(\overline{X}, \sum_{i=1}^n \alpha_i X_i) = 0$ as in d).

f)

$$E\big(S_{n+1}|S_1, \ldots, S_n\big) = E\big(S_n + X_{n+1}|S_1, \ldots, S_n\big) = S_n + E\big(X_{n+1}|S_1, \ldots, S_n\big) = S_n + \mu$$

since X_{n+1} is independent of $S_1, \ldots, S_n$.

Exercise 1.3

i)

$E(X_t)$ is independent of t since the distribution of X_t is independent of t and $E(X_t)$ exists.

ii)

Since $E[X_{t+h}X_t]^2 \leq E[X_{t+h}^2]E[X_t^2]$ for all integers t, h, and the joint distribution of X_{t+h} and X_t is independent of t, it follows that $E[X_{t+h}X_t]$ exists and is independent of t for every integer h.

Combining i) and ii) it follows that X_t is weakly stationary.

Exercise 1.4

a)

$E(X_t) = a$ is independent of t.

$$Cov\big(X_{t+h}, X_t\big) = \begin{cases} (b^2 + c^2)\sigma^2 & ; \quad h = 0 \\ 0 & ; \quad h = \pm 1 \\ bc\sigma^2 & ; \quad h = \pm 2 \\ 0 & ; \quad |h| > 2 \end{cases}$$

which is independent of t. That is, X_t is stationary.

b)

$E(X_t) = 0$ is independent of t.

$$Cov\big(X_{t+h}, X_t\big) = Cov\big(Z_1 \cos c(t+h) + Z_2 \sin c(t+h), Z_1 \cos ct + Z_2 \sin ct\big)$$
$$= \sigma^2\big(\cos c(t+h) \cos ct + \sin c(t+h) \sin ct\big) = \sigma^2 \cos ch$$

which is independent of t. That is, X_t is stationary.

c)

$E(X_t) = 0$ is independent of t.

$$Cov\big(X_{t+1}, X_t\big) = \sigma^2 \cos c(t+1) \sin ct$$

which is not independent of t. That is, X_t is not stationary (except in the special case when c is an integer multiple of 2π).

d)

$E(X_t) = a$ is independent of t.

$$Cov\big(X_{t+h}, X_t\big) = b^2\sigma^2$$

which is independent of t. That is, X_t is stationary.

e)

$E(X_t) = 0$ is independent of t.

$$Cov\big(X_{t+h}, X_t\big) = \sigma^2 \cos c(t+h) \cos ct$$

which is not independent of t. That is, X_t is not stationary (except in the special case when c is an integer multiple of 2π).

f)

$E(X_t) = 0$ is independent of t.

$$Cov\big(X_{t+h}, X_t\big) = E[X_{t+h}X_t] = E[Z_{t+h}Z_{t+h-1}Z_tZ_{t-1}] = \begin{cases} \sigma^4 & ; \quad h = 0 \\ 0 & ; \quad |h| > 0 \end{cases}$$

which is independent of t. That is, X_t is stationary, and it is seen that in fact $X_t \sim WN(0, \sigma^4)$.

Exercise 1.5

a)

The autocovariance function

$$\gamma_X(h) = \begin{cases} 1+\theta^2 & ; \quad h = 0 \\ \theta & ; \quad h = \pm 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

The autocorrelation function

$$\rho_X(h) = \begin{cases} 1 & ; \quad h = 0 \\ \frac{\theta}{1+\theta^2} & ; \quad h = \pm 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

For $\theta = 0.8$ it is obtained that

$$\gamma_X(h) = \begin{cases} 1.64 & ; \quad h = 0 \\ 0.8 & ; \quad h = \pm 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

$$\rho_X(h) = \begin{cases} 1 & ; \quad h = 0 \\ 0.488 & ; \quad h = \pm 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

b)
Let $\overline{X}_4 = \frac{1}{4}(X_1 + \ldots + X_4)$. Then

$$Var(\overline{X}_4) = Cov\big(\overline{X}_4, \overline{X}_4\big) = \frac{1}{16}\sum_{i=1}^{4}\sum_{i=1}^{4} Cov\big(X_i, X_j\big)$$
$$= \frac{1}{4}\big(\gamma_X(0) + \gamma_X(2)\big) = \frac{1}{4}\big(1.64 + 0.8\big) = 0.61$$

c)

$$Var(\overline{X}_4) = Cov\big(\overline{X}_4, \overline{X}_4\big) = \frac{1}{4}\big(\gamma_X(0) + \gamma_X(2)\big) = \frac{1}{4}\big(1.64 - 0.8\big) = 0.21$$

The negative lag 2 correlation in c) means that positive deviations of X_t from zero tend to be followed two time units later by a compensating negative deviation, resulting in smaller variability in the sample mean than in b) (and also smaller than if the time series X_t were IID(0, 1.64) in which case $Var(\overline{X}_4) = 0.41$).

Exercise 1.15

a)

Since s_t has period 12

$$\nabla_{12} X_t = \nabla_{12}(a + bt + s_t + Y_t) = 12b + Y_t - Y_{t-12}$$

so that

$$W_t := \nabla\nabla_{12} X_t = Y_t - Y_{t-1} - Y_{t-12} - Y_{t-13}.$$

Then $\mathrm{E}[W_t] = 0$ and

$$\begin{aligned}\mathrm{Cov}[W_{t+h}, W_t] &= \mathrm{Cov}[Y_{t+h} - Y_{t+h-1} - Y_{t+h-12} - Y_{t+h-13}, Y_t - Y_{t-1} - Y_{t-12} - Y_{t-13}] \\ &= 4\gamma(h) - 2\gamma(h-1) - 2\gamma(h+1) + \gamma(h-11) + \gamma(h+11) - 2\gamma(h-12) \\ &- 2\gamma(h+12) + \gamma(h+13) + \gamma(h-13)\end{aligned}$$

where $\gamma(\cdot)$ is the ACVF of Y_t. Since $\mathrm{E}[W_t]$ and $\mathrm{Cov}[W_{t+h}, W_t]$ are independent of t, W_t is stationary. Also note that $\nabla_{12} X_t$ is stationary.

b)

Using $X_t = (a + bt)s_t + Y_t$ it is obtained that

$$\nabla_{12} X_t = bts_t - b(t-12)s_{t-12} + Y_t - Y_{t-12} = 12bs_{t-12} + Y_t - Y_{t-12}.$$

Now let $U_t = \nabla_{12}^2 X_t = Y_t - 2Y_{t-12} + Y_{t-24}$. Then $\mathrm{E}[U_t] = 0$ and

$$\begin{aligned}\mathrm{Cov}[U_{t+h}, U_t] &= \mathrm{Cov}[Y_{t+h} - 2Y_{t+h-12} + Y_{t+h-24}, Y_t - 2Y_{t-12} + Y_{t-24}] \\ &= 6\gamma(h) - 4\gamma(h+12) - 4\gamma(h-12) + \gamma(h+2) + \gamma(h-24),\end{aligned}$$

which is independent of t. Hence U_t is stationary.

Exercise 2.1

$S(a, b) = E[(X_{n+h} - aX_n - b)^2]$ to be minimized wrt a and b. Now

$$S(a, b) = E[\big((X_{n+h} - \mu) - a(X_n - \mu) - b - a\mu + \mu\big)] = \gamma(0) + a^2\gamma(0) + (b + a\mu - \mu)^2 - 2a\gamma(h)$$

This gives

$$\frac{\partial S}{\partial a} = 2a\gamma(0) + 2\mu(b + a\mu - \mu) - 2\gamma(h)$$

$$\frac{\partial S}{\partial b} = 2(b + a\mu - \mu)$$

S is clearly minimized wrt b when for $b = \mu(1 - a)$. Substituting this value into $\frac{\partial S}{\partial a}$ and equating to zero leads to the result

$$a = \frac{\gamma(h)}{\gamma(0)} = \rho(h)$$

Hence, $S(a, b)$ is minimized when

$$a = \rho(h), \quad b = \mu(1 - \rho(h))$$

The BLP (best linear predictor) of X_{n+h} in terms of X_n is therefore $\mu + \rho(h)(X_n - \mu)$.

Exercise 2.3

a)

$$X_t = Z_t + 0.3Z_{t-1} - 0.4Z_{t-2}$$

$$\begin{aligned}
\gamma(0) &= 1 + 0.3^2 + 0.4^2 = 1.25 \\
\gamma(1) &= 0.3 - 0.4 \cdot 0.3 = 0.18 \\
\gamma(2) &= -0.4 \\
\gamma(h) &= 0, \quad h > 2 \\
\gamma(-h) &= \gamma(h)
\end{aligned}$$

b)

$$Y_t = \tilde{Z}_t - 1.2\tilde{Z}_{t-1} - 1.6\tilde{Z}_{t-2}$$

$$\begin{aligned}
\gamma(0) &= 0.25(1 + 1.2^2 + 1.6^2) = 1.25 \\
\gamma(1) &= 0.25(-1.2 + 1.6 \cdot 1.2) = 0.18 \\
\gamma(2) &= -1.6 \cdot 0.25 = -0.4 \\
\gamma(h) &= 0, \quad h > 2 \\
\gamma(-h) &= \gamma(h)
\end{aligned}$$

That is, we obtain the same ACVF as in a).

Exercise 2.5

$\sum_{j=1}^{\infty} \theta^j X_{n-j}$ converges absolutely (with probability 1) since

$$\begin{aligned}
E[\sum_{j=1}^{\infty} |\theta|^j |X_{n-j}|] &\leq \sum_{j=1}^{\infty} |\theta|^j E[|X_{n-j}|] \\
&\leq \sum_{j=1}^{\infty} |\theta|^j \sqrt{\gamma(0) + \mu^2} \quad \text{by Cauchy-Schwartz inequality} \\
&< \infty \quad \text{since} |\theta| < 1
\end{aligned}$$

That is, $\sum_{j=1}^{\infty} |\theta|^j |X_{n-j}| < \infty$ with probability 1.

Mean square convergence of $S_m = \sum_{j=1}^{m} \theta^j X_{n-j}$ as $m \to \infty$ can be verified by invoking Cauchy's criterion. For $m > k$

$$\begin{aligned}
E[|S_m - S_k|^2] &= E[(\sum_{j=k+1}^{m} \theta^j X_{n-j})^2] \\
&= \sum_{i=k+1}^{m} \sum_{j=k+1}^{m} \theta^{i+j} E[X_{n-i} X_{n-j}]
\end{aligned}$$

$$
\begin{aligned}
E[|S_m - S_k|^2] = E[(\sum_{j=k+1}^{m} \theta^j X_{n-j})^2] &= \sum_{i=k+1}^{m} \sum_{j=k+1}^{m} \theta^{i+j} E[X_{n-i} X_{n-j}] \\
&= \sum_{i=k+1}^{m} \sum_{j=k+1}^{m} \theta^{i+j} \big(\gamma(i-j) + \mu^2\big) \\
&\le \sum_{i=k+1}^{m} \sum_{j=k+1}^{m} |\theta|^{i+j} \big(\gamma(0) + \mu^2\big) = \big(\gamma(0) + \mu^2\big) \Big(\sum_{j=k+1}^{m} |\theta|^j \Big)^2 \\
&\to 0 \quad \text{as} \quad k, m \to \infty
\end{aligned}
$$

since $\sum_{j=1}^{\infty} |\theta|^j < \infty$. Hence, by Cauchy's mutual convergence criterion, mean square convergence is guaranteed.

Exercise 2.7

$$
\begin{aligned}
\frac{1}{1-\phi z} &= \frac{-\frac{1}{\phi z}}{1 - \frac{1}{\phi z}} \\
&= -\frac{1}{\phi z}\Big(1 + \frac{1}{\phi z} + \frac{1}{(\phi z)^2} + \ldots\Big) \\
&= -\sum_{j=1}^{\infty} (\phi z)^{-j}
\end{aligned}
$$

since $|\phi z| > 1$.

Exercise 2.8

$$
X_t = \phi X_{t-1} + Z_t
$$

$$
\begin{aligned}
X_t &= \phi X_{t-1} + Z_t \\
&= Z_t + \phi(Z_{t-1} + \phi X_{t-2}) \\
&= \ldots \\
&= Z_t + \phi Z_{t-1} + \ldots + \phi^n Z_{t-n} + \phi^{n+1} X_{t-n-1})
\end{aligned}
$$

That is

$$
X_t - \phi^{n+1} X_{t-n-1} = Z_t + \phi Z_{t-1} + \ldots + \phi^n Z_{t-n}
$$

First we calculate

$$
\begin{aligned}
\text{Var}(X_t - \phi^{n+1} X_{t-n-1}) &= \gamma(0)(1 + \phi^{2n+2}) - 2\phi^{n+1}\gamma(n+1) \\
&\le \gamma(0)(1 + |\phi|^{2n+2} + 2|\phi|^{n+1}) = 4\gamma(0)
\end{aligned}
$$

if X_t is stationary and $|\phi| = 1$

Next we calculate

$$\text{Var}(Z_t + \phi Z_{t-1} + \ldots + \phi^n Z_{t-n}) = n\sigma^2$$

if $|\phi| = 1$

Since clearly $n\sigma^2 > 4\gamma(0)$ for sufficiently large n, we have reached a contradiction. Hence X_t cannot be stationary if $|\phi| = 1$.

Exercise 2.10

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}$$

where $\phi = \theta = 0.5$

According to Section 2.3, equation (2.3.3), we obtain that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

where $\psi_0 = 1$, $\psi_j = (\phi + \theta)\phi^{j-1} = 0.5^{j-1}$ for $j = 1, 2, \ldots$.

$>$From Section 2.3, equation (2.3.5), we get

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$$

where $\pi_0 = 1$, $\pi_j = -(\phi + \theta)(-\theta)^{j-1} = -(-0.5)^{j-1}$ for $j = 1, 2, \ldots$.

Agrees with the results from ITSM.

Exercise 2.12

The given MA(1)-model is

$$X_t = Z_t - 0.6 Z_{t-1}$$

where $Z_t \sim \text{WN}(0, 1)$.

Observed that $\overline{x}_{100} = 0.157$

The variance of $\overline{x}_{100}$:

$$\begin{aligned}
\text{Var}[\overline{x}_{100}] &= \frac{1}{n} \sum_{h=-n}^{n} \Big(1 - \frac{|h|}{n}\Big)\gamma(h) \\
&= \frac{1}{100}\Big(\gamma(0) + 2 \cdot \frac{99}{100}\gamma(1)\Big) \\
&= \frac{1}{100}\Big(1.36 - 1.98 \cdot 0.6\Big) \\
&= 0.00172
\end{aligned}$$

That is, 95% confidence bounds for μ are approximately

$$\begin{aligned}&\overline{x}_{100} \pm 1.96\sqrt{0.00172}\\&= 0.157 \pm 1.96 \cdot 0.0415 = 0.157 \pm 0.0813 = 0.076,\ 0.238\end{aligned}$$

Reject H_0: $\mu = 0$ in favour of the alternative hypothesis H_1: $\mu \neq 0$ at significance level 0.05 since the 95% bounds for μ do not include the value 0.

Note: The conclusion would differ if the time series $X_t \sim IID(0, 1.36)$.

Exercise 2.13

a)

Assume an AR(1)-model

$$X_t = \phi X_{t-1} + Z_t$$

Since $\rho(h) = \phi^h$ $(h > 0)$ for an AR(1)-model, and it has been observed that $\hat{\rho(2)} = 0.145$, we shall assume that $\phi^2 << 1$. Using Bartlett's formula, the following approximate relations are obtained:

$$\text{Var}[\hat{\rho(1)}] \approx \frac{1}{n}\left(1 - \phi^2\right)$$

and

$$\text{Var}[\hat{\rho(2)}] \approx \frac{1}{n}\left(1 - \phi^2\right)\left(1 + 3\phi^2\right)$$

That is, 95% confidence bounds for $\rho(1)$ are approximately

$$\hat{\rho(1)} \pm \frac{1.96}{\sqrt{n}}\sqrt{1 - \phi^2}$$

Correspondingly, 95% confidence bounds for $\rho(2)$ are approximately

$$\hat{\rho(2)} \pm \frac{1.96}{\sqrt{n}}\sqrt{(1 - \phi^2)(1 + 3\phi^2)}$$

With $\phi = \hat{\phi} = \hat{\rho(1)}$, $n = 100$, $\hat{\rho(1)} = 0.438$, $\hat{\rho(2)} = 0.145$, these bounds become for $\rho(1)$: 0.262, 0.614, and for $\rho(2)$: -0.073, 0.369.

These values are not consistent with $\phi = 0.8$, since both $\rho(1) = 0.8$ and $\rho(2) = 0.64$ are outside these bounds.

b)

Assume an MA(1)-model

$$X_t = Z_t + \theta Z_{t-1}$$

Bartlett's formula gives the following approximate relations

$$\text{Var}[\hat{\rho(1)}] \approx \frac{1}{n}\left(1 - 3\rho(1)^2 + 4\rho(1)^4\right)$$

and

$$\text{Var}[\hat{\rho(2)}] \approx \frac{1}{n}\left(1 + 2\rho(1)^2\right)$$

That is, 95% confidence bounds for $\rho(1)$ are approximately

$$\hat{\rho(1)} \pm \frac{1.96}{\sqrt{n}}\sqrt{1 - 3\rho(1)^2 + 4\rho(1)^4}$$

Correspondingly, 95% confidence bounds for $\rho(2)$ are approximately

$$\hat{\rho(2)} \pm \frac{1.96}{\sqrt{n}}\sqrt{1 + 2\rho(1)^2}$$

With the numbers as in a), it is now obtained that these bounds become for $\rho(1)$: 0.290, 0.586, and for $\rho(2)$: -0.082, 0.378.

$\theta = 0.6$ leads to $\rho(1) = \frac{\theta}{1+\theta^2} = 0.4412$, $\rho(2) = 0$. It follows that the confidence bounds are consistent with these two values, and the data are therefore consistent with the MA(1)-model $X_t = Z_t + 0.6Z_{t-1}$

Exercise 2.14

$$X_t = A\cos(\omega t) + B\sin(\omega t), \quad t \in \mathbb{Z}$$

where A and B are uncorrelated random variables with zero mean and variance 1. This process is stationary with ACF $\rho(h) = \cos(\omega h)$.

a)

$$P_1 X_2 = \phi_{11} X_1$$

where $\gamma(0)\phi_{11} = \gamma(1)$, which gives $\phi_{11} = \rho(1) = \cos\omega$. Hence

$$P_1 X_2 = \cos(\omega) X_1$$

Also

$$E[(X_2 - P_1X_2)^2] = \gamma(0) - \phi_{11}\gamma(1) = \gamma(0)(1 - \cos^2\omega) = \sin^2\omega$$

Note: 2.14 is an example in which the matrix Γ_n in the equation $\Gamma_n\overline{\phi}_n = \overline{\gamma}_n$ is singular for $n \geq 3$. This is because $X_3 = (2\cos\omega)X_2 - X_1$.

b)

$$P_2 X_3 = \phi_{21} X_2 + \phi_{22} X_1$$

where

$$\gamma(0)\phi_{21} + \gamma(1)\phi_{22} = \gamma(1)$$
$$\gamma(1)\phi_{21} + \gamma(0)\phi_{22} = \gamma(2)$$

that is

$$\phi_{21} + (\cos\omega)\phi_{22} = \cos\omega$$
$$(\cos\omega)\phi_{21} + \phi_{22} = \cos 2\omega$$

Solving these equations give $\phi_{22}(\cos^2\omega - 1) = \cos^2\omega - 2\cos^2\omega + 1 = -\cos^2\omega + 1$, that is, $\phi_{22} = -1$, and then, $\phi_{21} = \cos\omega - \phi_{22}\cos\omega = 2\cos\omega$. Hence

$$P_2X_3 = (2\cos\omega)X_2 - X_1$$

and

$$\begin{aligned} E[(X_3 - P_2X_3)^2] &= \gamma(0) - \overline{\phi}_2\overline{\gamma}_2 \\ &= 1 - (2\cos\omega, -1)(\cos\omega, \cos 2\omega) \\ &= 1 - 2\cos^2\omega + \cos 2\omega = 0 \end{aligned}$$

c)

>From b) and stationarity, it follows that

$$P(X_{n+1}|X_n, X_{n-1}) = (2\cos\omega)X_n - X_{n-1}$$

with MSE = 0.

Since $(2\cos\omega)X_n - X_{n-1}$ is a linear combination of X_s, $-\infty < s \leq n$, and since it is impossible to find a predictor of this form with smaller MSE, we conclude that $\tilde{P}_nX_{n+1} = (2\cos\omega)X_n - X_{n-1}$ with MSE = 0.

Exercise 2.18

Given the MA(1) process

$$X_t = Z_t - \theta Z_{t-1}$$

where $|\theta| < 1$, and $Z_t \sim WN(0, \sigma^2)$. Represented as an AR(∞) process, it assumes the form

$$Z_t = X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \dots$$

Setting $t = n+1$ in the last equation and applying $\tilde{P}_n$ to each side, leads to the result

$$\tilde{P}_nX_{n+1} = -\sum_{j=1}^{\infty}\theta^j X_{n+1-j} = -\theta Z_n$$

Prediction error = $X_{n+1} - \tilde{P}_nX_{n+1} = Z_{n+1}$. Hence, MSE = $E[Z_{n+1}^2] = \sigma^2$.

Exercise 2.19

The given MA(1)-model is

$$X_t = Z_t - Z_{t-1}; \quad t \in \mathbb{Z}$$

where $Z_t \sim \mathrm{WN}(0, \sigma^2)$.

The vector $\mathbf{a} = (a_1, \ldots, a_n)'$ of the coefficients that provide the best linear predictor (BLP) of X_{n+1} in terms of $\mathbf{X} = (X_n, \ldots, X_1)'$ satisfies the equation

$$\Gamma_n \mathbf{a} = \gamma_{\mathbf{n}}$$

where the covariance matrix $\Gamma_n = Cov(\mathbf{X}, \mathbf{X})$ and $\gamma_{\mathbf{n}} = Cov(X_{n+1}, \mathbf{X}) = (\gamma(1), \ldots, \gamma(n))'$. Since $\gamma(0) = 2\sigma^2$, $\gamma(1) = -\sigma^2$, $\gamma(h) = 0$ for $|h| > 1$, it follows that

$$\Gamma_n = \sigma^2 \begin{pmatrix} 2 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 & 2 \end{pmatrix}$$

and $\gamma_{\mathbf{n}} = \sigma^2(-1, 0, \ldots, 0)'$. It can be shown, e.g. by induction, that the equations to be solved can be rewritten as follows

$$\begin{pmatrix} 2 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 3 & 2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 4 & 3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & n & (n-1) \\ 0 & 0 & 0 & 0 & \ldots & 1 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ \vdots \\ (-1)^{n-1} \\ 0 \end{pmatrix}$$

The solution is found to be given as follows

$$a_j = (-1)^j \frac{n+1-j}{n+1}$$

Hence it is obtained that

$$P_n X_{n+1} = \sum_{j=1}^{n} (-1)^j \frac{n+1-j}{n+1} X_{n+1-j}$$

The mean square error is

$$E[(X_{n+1} - P_n X_{n+1})^2] = \gamma(0) - \mathbf{a}'\gamma_{\mathbf{n}} = 2\sigma^2 + a_1\sigma^2 = \sigma^2\Big(1 + \frac{1}{n+1}\Big)$$

Exercise 2.20

We have to prove that

$$Cov(X_n - \hat{X}_n, X_j) = E[(X_n - \hat{X}_n)X_j] = 0$$

for $j = 1, \ldots, n-1$. This follows from equations (2.5.5) for suitable values of n and h with $a_0 = 0$ (since we may assume that $E[X_n] = 0$). This clearly implies that

$$E[(X_n - \hat{X}_n)(X_k - \hat{X}_k)] = 0$$

for $k = 1, \ldots, n-1$, since $\hat{X}_k$ is a linear combination of $X_1, \ldots, X_{k-1}$.

Exercise 2.21

In this exercise we shall determine the best linear predictor (BLP) $P(X_3|\mathbf{W}_\alpha)$ wrt three different vector variables $\mathbf{W}_\alpha$, $\alpha = a, b, c$. Let $\Gamma_\alpha = Cov(\mathbf{W}_\alpha, \mathbf{W}_\alpha)$ and $\gamma_\alpha = Cov(X_3, \mathbf{W}_\alpha)$.

The given MA(1)-model is

$$X_t = Z_t + \theta Z_{t-1}; \quad t \in \mathbb{Z}$$

where $Z_t \sim \mathrm{WN}(0, \sigma^2)$.

a)

In this case we have $\mathbf{W}_a = (W_1, W_2)' = (X_2, X_1)'$. Hence

$$\Gamma_a = \sigma^2 \begin{pmatrix} 1+\theta^2 & \theta \\ \theta & 1+\theta^2 \end{pmatrix}$$

and $\gamma_a = Cov(X_3, \mathbf{W}_a) = (\gamma(1), \gamma(2))' = \sigma^2(\theta, 0)$. The equation $\Gamma_a(a_1, a_2)' = \gamma_a$, or

$$(1+\theta^2)a_1 + \theta a_2 = \theta$$

$$\theta a_1 + (1+\theta^2)a_2 = 0$$

has the solution

$$a_1 = \frac{\theta(1+\theta^2)}{(1+\theta^2)^2 - \theta^2} \qquad a_2 = \frac{-\theta^2}{(1+\theta^2)^2 - \theta^2}$$

We obtain the BLP

$$P(X_3|X_2, X_1) = \frac{\theta}{(1+\theta^2)^2 - \theta^2}\left((1+\theta^2)X_2 - \theta X_1\right)$$

The mean square error

$$\begin{aligned} E[(X_3 - P(X_3|X_2, X_1))^2] &= \mathrm{Var}(X_3) - (a_1, a_2)\gamma_a = \sigma^2(1+\theta^2) - a_1\sigma^2\theta \\ &= \sigma^2(1+\theta^2)\left(1 - \frac{\theta^2}{(1+\theta^2)^2 - \theta^2}\right) \end{aligned}$$

b)

Here $\mathbf{W}_b = (W_1, W_2)' = (X_4, X_5)'$. With this choice, it follows that $\Gamma_b = \Gamma_a$, and $\gamma_b = \gamma_a$. It follows immediately that the BLP is given by

$$P(X_3|X_4, X_5) = \frac{\theta}{(1+\theta^2)^2 - \theta^2}\left((1+\theta^2)X_4 - \theta X_5\right)$$

And the mean square error is the same as in a)

$$E[(X_3 - P(X_3|X_4, X_5))^2] = \sigma^2(1+\theta^2)\left(1 - \frac{\theta^2}{(1+\theta^2)^2 - \theta^2}\right)$$

c)

Now, $\mathbf{W}_b = (W_1, W_2, W_3, W_4)' = (X_2, X_1, X_4, X_5)'$. It then follows that

$$\Gamma_c = \sigma^2 \begin{pmatrix} \Gamma_a & \bar{0} \\ \bar{0} & \Gamma_a \end{pmatrix}$$

where $\bar{0}$ denotes a 2×2 zero-matrix. Also, $\gamma_c = (\gamma_a', \gamma_a')'$. Hence, it follows that the solution to the equation $\Gamma_c(a_1, \ldots, a_4)' = \gamma_c$ is given by $a_3 = a_1$ and $a_4 = a_2$, where a_1 and a_2 are as given in a) or b). The BLP is therefore

$$P(X_3|X_2, X_1, X_4, X_5) = \frac{\theta}{(1+\theta^2)^2 - \theta^2}\left((1+\theta^2)[X_2 + X_4] - \theta[X_1 + X_5]\right)$$

with mean square error

$$\begin{aligned} E[(X_3 - P(X_3|X_2, X_1, X_4, X_5))^2] &= Var(X_3) - (a_1, a_2, a_3, a_4)\gamma_c = \sigma^2(1+\theta^2) - 2a_1\sigma^2\theta \\ &= \sigma^2(1+\theta^2)\left(1 - \frac{2\theta^2}{(1+\theta^2)^2 - \theta^2}\right) \end{aligned}$$

d)

See above.

Exercise 2.22

We shall determine the best linear predictor (BLP) $P(X_3|\mathbf{W}_\alpha)$ wrt three different vector variables $\mathbf{W}_\alpha$, $\alpha = a, b, c$. Let $\Gamma_\alpha = Cov(\mathbf{W}_\alpha, \mathbf{W}_\alpha)$ and $\gamma_\alpha = Cov(X_3, \mathbf{W}_\alpha)$.

The given causal (stationary) AR(1)-model is

$$X_t = \phi X_{t-1} + Z_t; \quad t \in \mathbb{Z}$$

where $Z_t \sim \mathrm{WN}(0, \sigma^2)$. Causality implies that $|\phi| < 1$. Hence, the ACVF $\gamma(h) = \sigma^2(1-\phi^2)^{-1}\phi^{|h|}$.

a)

In this case we have $\mathbf{W}_a = (W_1, W_2)' = (X_2, X_1)'$. Hence

$$\Gamma_a = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix}$$

and $\gamma_a = Cov(X_3, \mathbf{W}_a) = (\gamma(1), \gamma(2))' = \frac{\sigma^2}{1-\phi^2}(\phi, \phi^2)'$. The equation $\Gamma_a(a_1, a_2)' = \gamma_a$, or

$$a_1 + \phi a_2 = \phi$$

$$\phi a_1 + a_2 = \phi^2$$

has the solution

$$a_1 = \phi \qquad a_2 = 0$$

We obtain the BLP

$$P(X_3|X_2, X_1) = \phi X_2$$

The mean square error

$$E[(X_3 - P(X_3|X_2, X_1))^2] = \text{Var}(X_3) - (a_1, a_2)\gamma_a = \frac{\sigma^2}{1-\phi^2} - \frac{\sigma^2\phi^2}{1-\phi^2} = \sigma^2$$

b)

Here $\mathbf{W}_b = (W_1, W_2)' = (X_4, X_5)'$. With this choice, it follows that $\Gamma_b = \Gamma_a$, and $\gamma_b = \gamma_a$. It follows immediately that the BLP is given by

$$P(X_3|X_4, X_5) = \phi X_4$$

And the mean square error is

$$E[(X_3 - P(X_3|X_4, X_5))^2] = Var(X_3) - (a_1, a_2)\gamma_b = \frac{\sigma^2}{1-\phi^2} - \frac{\sigma^2\phi^2}{1-\phi^2} = \sigma^2$$

c)

Now, $\mathbf{W}_c = (W_1, W_2, W_3, W_4)' = (X_2, X_1, X_4, X_5)'$. It then follows that

$$\Gamma_c = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi^3 & \phi^4 \\ \phi^2 & \phi^3 & 1 & \phi \\ \phi^3 & \phi^4 & \phi & 1 \end{pmatrix}$$

where $\gamma_c = (\gamma_a', \gamma_a')'$. Hence, the following set of equations is obtained

$$a_1 + \phi a_2 + \phi^2 a_3 + \phi^3 a_4 = \phi$$

$$\phi a_1 + a_2 + \phi^3 a_3 + \phi^4 a_4 = \phi^2$$

$$\phi^2 a_1 + \phi^3 a_2 + a_3 + \phi a_4 = \phi$$

$$\phi^3 a_1 + \phi^4 a_2 + \phi a_3 + a_4 = \phi^2$$

It is seen that the first two equations give $a_2 = 0$, while the last two equations give $a_4 = 0$. Then it is found that

$$a_1 = a_3 = \frac{\phi}{1+\phi^2}$$

The BLP is therefore

$$P(X_3|X_2, X_1, X_4, X_5) = \frac{\phi}{1+\phi^2}[X_2 + X_4]$$

with mean square error

$$E[(X_3 - P(X_3|X_2, X_1, X_4, X_5))^2] = Var(X_3) - (a_1, a_2, a_3, a_4)\gamma_c = \frac{\sigma^2}{1-\phi^2} - \frac{\sigma^2}{1-\phi^2}\frac{2\phi^2}{1+\phi^2}$$
$$= \frac{\sigma^2}{1+\phi^2}$$

d)

See above.