Time Series
Examinations Solutions

Kacper Omieliańczyk (459458)

Task 1

We write the model as

$$X_t = -0.25X_{t-1} + Z_t$$

Since $\{Z_t\}$ is white noise with $\mathrm{Var}(Z_t) = 1$, it is well-known

$$\gamma(k) = \mathrm{Cov}(X_t, X_{t-k}) = \frac{\sigma_Z^2}{1-\phi^2}\phi^{|k|}$$

where $\phi = 0.25, \sigma_Z^2 = 1$. Thus

$$1 - \phi^2 = \frac{15}{16}, \quad \gamma(k) = \frac{16}{15}(-0.25)^{|k|}$$

In particular

$$\gamma(0) = \frac{16}{15}, \quad \gamma(1) = -\frac{4}{15}, \quad \gamma(2) = \frac{1}{15}$$

We seek (a_1, a_2) minimizing

$$\mathrm{MSE} = \mathbb{E}\left[(X_2 - a_1X_1 - a_2X_3)^2\right]$$

The normal equations are

$$\mathrm{Cov}(X_2, X_1) = a_1\mathrm{Cov}(X_1, X_1) + a_2\mathrm{Cov}(X_3, X_1)$$

$$\mathrm{Cov}(X_2, X_3) = a_1\mathrm{Cov}(X_1, X_3) + a_2\mathrm{Cov}(X_3, X_3)$$

In matrix form

$$\begin{pmatrix} \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \gamma(1) \end{pmatrix}$$

Numerically

$$R = \begin{pmatrix} 16/15 & 1/15 \\ 1/15 & 16/15 \end{pmatrix}, \quad v = \begin{pmatrix} -4/15 \\ -4/15 \end{pmatrix}$$

We compute $\det R$ and R^{-1}

$$\det R = \left(\frac{16}{15}\right)^2 - \left(\frac{1}{15}\right)^2 = \frac{17}{15}$$

$$R^{-1} = \frac{1}{\det R}\begin{pmatrix} 16/15 & -1/15 \\ -1/15 & 16/15 \end{pmatrix} = \begin{pmatrix} 16/17 & -1/17 \\ -1/17 & 16/17 \end{pmatrix}$$

Hence

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = R^{-1}v = \begin{pmatrix} -4/17 \\ -4/17 \end{pmatrix}$$

In other words

$$a_1 = a_2 = -\frac{4}{17}, \quad \hat{X}_2 = -\frac{4}{17}(X_1 + X_3)$$

A standard formula for the MSE is

$$\min \mathbb{E}[(X_2 - \hat{X}_2^2)] = \gamma(0) - v^T R^{-1} v = \gamma(0) - v^T a$$

We already have $\gamma(0) = 16/15$ and

$$v^T a = \begin{pmatrix} -4/15 & -4/15 \end{pmatrix} \begin{pmatrix} -4/17 \\ -4/17 \end{pmatrix} = \frac{32}{255}$$

Thus

$$\min \mathbb{E}[(X_2 - \hat{X}_2^2)] = \frac{16}{17}$$

□

Task 2

We compute autocovariances

$$\gamma(0) = \mathrm{Var}(X_t) = \mathrm{Var}(Z_t) + \theta^2 \mathrm{Var}(Z_{t-1}) = \sigma^2(1+\theta^2)$$
$$\gamma(1) = \mathrm{Cov}(X_t, X_{t-1}) = \mathrm{Cov}(Z_t + \theta Z_{t-1}, Z_{t-1} + \theta Z_{t-2}) = \sigma^2\theta$$

and $\gamma(h) = 0$ for $|h| \geqslant 2$. The lag-1 correlation

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta}{1+\theta^2}$$

Maximize $\rho(1)$ over θ

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{\theta}{1+\theta^2}\right] = \frac{(1+\theta^2) - 2\theta^2}{(1+\theta^2)^2} = \frac{1-\theta^2}{(1+\theta^2)^2} = 0 \implies \theta^2 = 1 \implies \theta = \pm 1$$

Of these two, $\rho(1) = \theta/(1+\theta^2)$ is largest when $\theta = +1$. Hence

$$\theta^* = 1, \quad \rho(1)\Big|_{\theta=1} = \frac{1}{2}$$

For any MA(1), the spectral density is

$$f_X(\omega) = \frac{\sigma^2}{2\pi}\left|1 + \theta e^{-i\omega}\right|^2 = \frac{\sigma^2}{2\pi}(1 + \theta^2 - 2\theta\cos\omega)$$

Plug in $\theta = 1$

$$f_X(\omega) = \frac{\sigma^2}{2\pi}(1 + 1 + 2\cos\omega) = \frac{\sigma^2}{\pi}(1+\cos\omega) = \frac{2\sigma^2}{\pi}\cos^2\left(\frac{\omega}{2}\right)$$

$\cos^2\left(\frac{\omega}{2}\right)$ is largest at $\omega = 0$ (value 1) and decreases to 0 at $\omega = \pi$. Thus at $\theta = 1$, low frequencies dominate the spectrum. As θ moves from 1:

- If θ decreases toward 0, the spectrum flattens out (approaches white noise),
- If θ becomes negative, the $+2\theta\cos\omega$ term shifts power toward high frequencies (since $\cos\omega < 0$ near $\omega = \pi$).

□

Task 3

In all cases we write

$$\phi(B)Y_t = \theta(B)Z_t$$

with

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots, \quad \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots$$

An ARMA(1, 1) requires

- after cancelling any common factors, both $\phi(B)$ and $\theta(B)$ are of order at most 1,
- the single AR-root lies outside the unit circle (causality),

- the single MA-root lies outside the unit circle (invertibility).

(i)

$$Y_t - \frac{5}{6}Y_{t-1} = Z_t - \frac{9}{20}Z_{t-1}$$

Here

$$\phi(B) = 1 - \frac{5}{6}B, \quad \theta(B) = 1 - \frac{9}{20}B$$

Both are already degree 1, with no common factor to cancel. The AR-root solves $1 - \frac{5}{6}z = 0 \implies z = 6/5$, and $|6/5| > 1$, so the process is stationary/causal. The MA-root is $1 - \frac{9}{20}z = 0 \implies z = 20/9, |20/9| > 1$, so it is invertible. Thus, (i) is causal, invertible ARMA(1, 1).

(ii)

$$Y_t - Y_{t-1} + \frac{1}{4}Y_{t-2} = Z_t - \frac{5}{4}Z_{t-1} + \frac{3}{8}Z_{t-2}$$

Here

$$\phi(B) = 1 - B + \frac{1}{4}B^2 = \left(1 - \frac{1}{2}B\right)^2$$

$$\theta(B) = 1 - \frac{5}{4}B + \frac{3}{8}B^2 = \left(1 - \frac{3}{4}B\right)\left(1 - \frac{1}{2}B\right)$$

There is a common factor $\left(1 - \frac{1}{2}B\right)$. Cancel it to get

$$\phi^*(B) = 1 - \frac{1}{2}B, \quad \theta^*(B) = 1 - \frac{3}{4}B$$

Now both are degree 1. The AR-root is $z = 2, |z| > 1$, so causal, and the MA-root is $z = \frac{4}{3}, |z| > 1$, so invertible. Thus, (ii) reduces to causal, invertible ARMA(1, 1).

(iii)

$$Y_t + \frac{1}{2}Y_{t-1} - \frac{1}{2}Y_{t-2} = Z_t - \frac{5}{4}Z_{t-1} + \frac{3}{8}Z_{t-2}$$

Here

$$\phi(B) = 1 + \frac{1}{2} - \frac{1}{2}B^2 \implies \phi(z) = 1 + \frac{1}{2}z - \frac{1}{2}z^2 = 0 \implies z = 2 \text{ or } z = -1$$

Since one root has $|z| = 1$, the AR part fails causality (on the boundary), and there is no cancellation with the MA. One checks the MA factors as in (ii), but they share no common factor with $\phi(B)$. This it is a true ARMA(2, 2), non-causal. Thus, (iii) is not causal ARMA(1, 1).

We take

$$Y_t - \frac{5}{6}Y_{t-1} = Z_t - \frac{9}{20}Z_{t-1}$$

and choose the state vector

$$x_t = \begin{pmatrix} Y_t \\ Z_t \end{pmatrix}, \quad \text{input } u_t = Z_t, \quad \text{output } y_t = Y_t$$

Then

$$\underbrace{\begin{pmatrix} Y_t \\ Z_t \end{pmatrix}}_{x_t} = \underbrace{\begin{pmatrix} 5/6 & -9/20 \\ 0 & 0 \end{pmatrix}}_{A} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{B} Z_t$$

and the observation equation is

$$y_t = \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{C} x_t$$

Written in standard state-space form:

$$\boxed{\begin{aligned} x_t &= Ax_{t-1} + Bu_t \\ y_t &= Cx_t \end{aligned}} \quad \text{where } A = \begin{pmatrix} 5/6 & -9/20 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

One checks immediately that the first row of the state equation reproduces $Y_t = \frac{5}{6}Y_{t-1} - \frac{9}{20}Z_{t-1} + Z_t$ □

Task 4

We are given

$$\begin{cases} X_t = b_t - (\theta_1 + \theta_2)Z_t - \theta_2 Z_{t-1} \\ b_t = b_{t-1} + (1 + \theta_1 + \theta_2)Z_t \\ \{Z_t\} \sim \mathrm{WN}(0, 1) \end{cases}$$

Here b_t, is itself (unobserved) random walk driven by the same white noise.

By definition,

$$\Delta X_t = X_t - X_{t-1}$$

We substitute the expressions for X_t and X_{t-1}:

$$\begin{aligned} \Delta X_t &= [b_t - (\theta_1 + \theta_2)Z_t - \theta_2 Z_{t-1}] - [b_{t-1} - (\theta_1 + \theta_2)Z_{t-1} - \theta_2 Z_{t-2}] \\ &= (b_t - b_{t-1}) - (\theta_1 + \theta_2)(Z_t - Z_{t-1}) - \theta_2(Z_{t-1} - Z_{t-2}) \end{aligned}$$

From the second equation,

$$b_t - b_{t-1} = (1 + \theta_1 + \theta_2)Z_t$$

Hence

$$\Delta X_t = (1 + \theta_1 + \theta_2)Z_t - (\theta_1 + \theta_2)(Z_t - Z_{t-1}) - \theta_2(Z_{t-1} - Z_{t-2})$$

Distribute the parantheses:

$$\begin{aligned} \Delta X_t &= (1 + \theta_1 + \theta_2)Z_t - (\theta_1 + \theta_2)Z_t + (\theta_1 + \theta_2)Z_{t-1} - \theta_2 Z_{t-1} + \theta_2 Z_{t-2} \\ &= \underbrace{[(1 + \theta_1 + \theta_2) - (\theta_1 + \theta_2)]}_{=1} Z_t + \underbrace{[(\theta_1 + \theta_2) - \theta_2]}_{=\theta_1} Z_{t-1} + \theta_2 Z_{t-2} \end{aligned}$$

Thus

$$\Delta X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

which is exactly an MA(2) model. Since $\Delta X_t = X_t - X_{t-1}$ is MA(2), it follows that $\{X_t\}$ is integrated of order 1 with an MA(2) error – that is,

$$X_t \sim \mathrm{ARIMA}(0, 1, 2)$$

$\square$

Task 5

Recall the notation:

- $X_n = (X_1, \ldots, X_n)^T$,
- $\hat{X}_n = (\hat{X}_1, \ldots, \hat{X}_n)^T$,
- the one-step prediction errors $\epsilon_t = X_t - \hat{X}_t$, so $\epsilon_n = (\epsilon_1, \ldots, \epsilon_n)^T$,
- the prediction-error variances $v_{t-1} = \mathbb{E}[\epsilon_t^2]$, collected in the diagonal matrix

$$D_N = \mathrm{diag}(v_0, v_1, \ldots, v_{n-1}), \quad v_0 = \mathbb{E}[\epsilon_1^2] = \mathrm{Var}(X_1)$$

- the lower-triangular „Yule-Walker" matrix:

$$\Phi_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi_{11} & 1 & 0 & \cdots & 0 \\ -\phi_{21} & -\phi_{22} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\phi_{n-1,1} & -\phi_{n-1,2} & \cdots & -\phi_{n-1,n-1} & 1 \end{pmatrix}$$

which encodes the t-th-step prediction

$$\hat{X}_t = \sum_{j=1}^{t-1} \phi_{t-1,j} X_{t-j}$$

(a) Let $C_n = \Phi_n^{-1}$. Because the Φ_n is unit-lower-triangular, so is its inverse C_n, and in particular $(C_n)_{ii} = 1$. Next we observe that the defining relation

$$\Phi_n X_n = \epsilon_n \qquad \Longrightarrow \qquad X_n = C_n \epsilon_n$$

Since $\mathbb{E}[\epsilon_n \epsilon_n^T] = D_n$, it follows immediately that

$$\Gamma_n = \mathbb{E}[X_n X_n^T] = \mathbb{E}[C_n \epsilon_n \epsilon_n^T C_n^T] = C_n D_n C_n^T$$

as claimed.
(b) By definition $\epsilon_n = X_n - \hat{X}_n$. Hence,

$$\hat{X}_n = X_n - \epsilon_n = C_n \epsilon_n - \epsilon_n = (C_n - I_n)\epsilon_n$$

Equivalently,

$$\hat{X}_n = (C_n - I)\epsilon_n$$

(c) We have

$$\gamma(0) = \mathrm{Var}(X_t) = 1 + 1.2^2 = 2.44, \quad \gamma(1) = \mathrm{Cov}(X_t, X_{t-1}) = -1.2, \quad \gamma(h) = 0 \ (h \geqslant 2)$$

Hence the t-th-step prediction coefficients $\{\phi_{t-1,j}\}$ come from solving

$$\underbrace{\begin{pmatrix} \gamma(0) & \gamma(1) & \dots \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \vdots & \ddots & \ddots \end{pmatrix}}_{R} \begin{pmatrix} \phi_{t-1,1} \\ \phi_{t-1,2} \\ \vdots \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \end{pmatrix}$$

By *(b)*,

$$\hat{X}_4 = (C_4 - I)\epsilon_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \phi_{11} & 0 & 0 & 0 \\ \phi_{21} & \phi_{22} & 0 & 0 \\ \phi_{31} & \phi_{32} & \phi_{33} & 0 \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix}$$

In particular

$$\hat{X}_4 = \phi_{31}\epsilon_1 + \phi_{32}\epsilon_2 + \phi_{33}\epsilon_3$$

Numerically (e.g. in R, see R-snippet 1) one finds

$$\phi_{31} \approx -0.723, \quad \phi_{32} \approx -0.469, \quad \phi_{33} \approx -0.230$$

so

$$\hat{X}_4 \approx -0.723\epsilon_1 - 0.469\epsilon_2 - 0.230\epsilon_3$$

- For $t = 2$, $\epsilon_2 = X_2 - \phi_{11} X_1$ with $\phi_{11} = \gamma(1)/\gamma(0) = -1.2/2.44 \approx -0.4918$. Hence

$$v_1 = \mathrm{Var}(\epsilon_2) = \gamma(0)\left(1 - \phi_{11}^2\right) \approx 1.8498$$

- For $t = 3$, one solves the 2x2 normal equations to get $\phi_{22} \approx -0.6487, \phi_{22} \approx -0.3190$. Then

$$v_2 = \mathrm{Var}(X_3 - \phi_{21}X_2 - \phi_{22}X_1) = \gamma(0) - (\gamma(1), \gamma(2)) \begin{pmatrix} \phi_{21} \\ \phi_{22} \end{pmatrix} \approx 1.6616$$

```r
# MA(1) parameters
theta <- -1.2

# autocovariances
gamma0 <- 1 + theta^2      # 2.44
gamma1 <- theta            # -1.2
gamma2 <- 0                # beyond lag 1

```

```r
# (ii) 2x2 for phi21,phi22
R2 <- matrix(c(gamma0, gamma1,
               gamma1, gamma0), 2, 2)
r2 <- c(gamma1, gamma2)
phi2 <- solve(R2, r2)
phi21 <- phi2[1];  phi22 <- phi2[2]

# (i) 3x3 for phi31,phi32,phi33
R3 <- matrix(c(gamma0, gamma1, gamma2,
               gamma1, gamma0, gamma1,
               gamma2, gamma1, gamma0), 3, 3)
r3 <- c(gamma1, gamma2, 0)
phi3 <- solve(R3, r3)
phi31 <- phi3[1];  phi32 <- phi3[2];  phi33 <- phi3[3]

cat("phi21=",phi21,"\nphi22=",phi22,
    "\nphi31=",phi31,"\nphi32=",phi32,"\nphi33=",phi33,"\n")
```

Listing 1: Solve for ϕ_{21}, ϕ_{22} and $\phi_{31}, \phi_{32}, \phi_{33}$ in R.

□

Task 6

First observe that „invertible MA(q)" means exactly that the MA-polynomial

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

has all it's zeros *outside* the unit circle. Equivalently there is a unique absolutely-summable sequence $\{\theta_j\}_{j=1}^{\infty}$ (with $\theta_j = 0$ form $j > q$) such that the infinite-past linear predictor of X_{t+1} based on $\{X_t, X_{t-1}, \ldots\}$ is

$$\hat{X}_{t+1|\infty} = \sum_{j=1}^{\infty} \theta_j (X_{t+1-j} - \hat{X}_{t+1-j})$$

A standard result in linear prediction theory then says

$$X_{t+1} - \hat{X}_{t+1|\infty} = Z_{t+1}$$

exactly, and that $\sum_{j=1}^{\infty} |\theta_j| < \infty$.

By definition our one-step estimator based on the first t observations is

$$\hat{X}_{t+1} = \sum_{j=1}^{t} \hat{\theta}_{t,j} (X_{t+1-j} - \hat{X}_{t+1-j})$$

where $\hat{\theta}_{t,j}$ solve the $t \times t$ Yule-Walker system. A key theorem is

Theorem 1 (Wold-Levinson)**.** *For an invertible process, as $t \to \infty$,*

$$\hat{\theta}_{t,j} \to \theta_j \quad \textit{for each fixed } j$$

and in fact $\sum_{j=1}^{\infty} |\theta_j| < \infty$.

Therefore we can write the „prediction error minus the true innovation" as

$$\begin{aligned}
X_{t+1} - \hat{X}_{t+1} - Z_{t+1} &= \left(X_{t+1} - \hat{X}_{t+1|\infty}\right) - \left(\hat{X}_{t+1} - \hat{X}_{t+1|\infty}\right) - Z_{t+1} \\
&= Z_{t+1} - Z_{t+1} - \sum_{j=1}^{\infty} \theta_j (X_{t+1-j} - \hat{X}_{t+1-j}) + \sum_{j=1}^{t} \hat{\theta}_{t,j} (X_{t+1-j} - \hat{X}_{t+1-j}) \\
&= \sum_{j=1}^{t} (\hat{\theta}_{t,j} - \theta_j) \Delta_{t+1-j} - \sum_{j=t+1}^{\infty} \theta_j \Delta_{t+1-j}
\end{aligned}$$

where we have written $\Delta_s = X_s - \hat{X}_s$. Because

1. Δ_s is uniformly L^2-bounded (it's just a linear combination of finitely many Z's),
2. $\hat{\theta}_{t,j} \to \theta_j$ for each fixed j, and
3. $\sum_{j=1}^{\infty} |\theta_j| < \infty$

one shows by dominated convergence (or the Cauchy criterion) that the right-hand side converges to zero in mean square. Hence

$$\mathbb{E}[(X_{t_1} - \hat{X}_{t+1} - Z_{t+1})^2] \to 0$$

which is exactly part (a).

By definition

$$v_t = \mathbb{E}[(X_{t+1} - \hat{X}_{t+1})^2]$$

Now

$$X_{t+1} - \hat{X}_{t+1} = (Z_{t+1}) + \underbrace{[X_{t+1} - \hat{X}_{t+1} - Z_{t+1}]}_{\to 0 \text{ in } L^2}$$

Call the bracketed term R_{t+1}. Then

$$v_t = \mathbb{E}[(Z_{t+1} + R_{t+1})^2] = \mathbb{E}[Z_{t+1}^2] + \mathbb{E}[R_{t+1}^2] + 2\mathbb{E}[Z_{t+1}R_{t+1}]$$

But R_{t+1} is a (mean-zero) function of $\{Z_s \colon s \leqslant t\}$, so it is uncorrelated with Z_{t+1}. Hence the cross-term vanishes, and

$$v_t = \sigma^2 + \mathbb{E}[R_{t+1}^2] \to \sigma^2 + 0 = \sigma^2$$

That completes part (b).

As we observe more and more of the past, our finite-sample predictor $\{\hat{X}_{t+1}\}$ converges for the „ideal" infinite-past best linear predictor (the Wold predictor), whose one-step-ahead error is exactly the current innovation Z_{t+1}. Thus eventually our prediction error variance converges to $\mathrm{Var}(Z_{t+1}) = \sigma^2$. $\square$