TMA4285 Time series models
Solution to exercise 1, autumn 2018

August 28, 2018

Problem A.4

We can solve this problem using the moment generating function, $M_X(t) = E(e^{tX}), t \in \mathbb{R}$.

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{t}) &= M_{\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}+\boldsymbol{\mu}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}^T(\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}+\boldsymbol{\mu})}) = e^{\boldsymbol{t}^T\boldsymbol{\mu}}E(e^{\boldsymbol{t}^T\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}}) \\ &= e^{\boldsymbol{t}^T\boldsymbol{\mu}}M_{\boldsymbol{Z}}(\boldsymbol{t}^T\boldsymbol{\Sigma}^{1/2}) = e^{\boldsymbol{t}^T\boldsymbol{\mu}}e^{\frac{1}{2}\boldsymbol{t}^T(\boldsymbol{\Sigma}^{1/2})^2\boldsymbol{t}} = e^{\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}, \end{aligned}$$

which is the moment generating function for a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

Note that the standard normal multivariate distribution has $M_{\boldsymbol{Z}}(\boldsymbol{t}) = e^{\frac{1}{2}\boldsymbol{t}^T\boldsymbol{I}\boldsymbol{t}}$.

Problem A.5

Since $\boldsymbol{Y}$ is a linear combination of a Gaussian vector $\boldsymbol{X}$ (a special case of a Gaussian process), $\boldsymbol{Y}$ is also Gaussian. We specify the mean and covariance matrix of $\boldsymbol{Y} = \boldsymbol{a} + B\boldsymbol{X}$.

$$\begin{aligned} E(\boldsymbol{Y}) &= E(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{X}) = E(\boldsymbol{a}) + \boldsymbol{B}E(\boldsymbol{X}) = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu} \\ Cov(\boldsymbol{Y}) &= \boldsymbol{B}Cov(\boldsymbol{X})\boldsymbol{B}^T = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T \end{aligned}$$

Problem A.6

Let $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}$,$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$, and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$. Let $\boldsymbol{X} \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Proof proposition A.3.1:

i: $\boldsymbol{\Sigma}_{12} = 0 \rightarrow$ independence: If $\boldsymbol{\Sigma}_{12} = 0$ then $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^T = 0$, and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$. Then $M_{\boldsymbol{X}}(\boldsymbol{t}) = e^{\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}} = e^{\boldsymbol{t}_1^T\boldsymbol{\mu}_1+\frac{1}{2}\boldsymbol{t}_1^T\boldsymbol{\Sigma}_{11}\boldsymbol{t}_1}e^{\boldsymbol{t}_2^T\boldsymbol{\mu}_2+\frac{1}{2}\boldsymbol{t}_2^T\boldsymbol{\Sigma}_{22}\boldsymbol{t}_2} = M_{\boldsymbol{X}_1}(\boldsymbol{t}_1)M_{\boldsymbol{X}_2}(\boldsymbol{t}_2)$, which means that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent.

Independence $\rightarrow \boldsymbol{\Sigma}_{12} = 0$: We can use the same argument going backwards. If we have independence, we must have $M_{\boldsymbol{X}}(\boldsymbol{t}) = M_{\boldsymbol{X}_1}(\boldsymbol{t}_1)M_{\boldsymbol{X}_2}(\boldsymbol{t}_2)$, which is only achivelable if $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$.

ii: The conditional distribution of $\boldsymbol{X}_2|\boldsymbol{X}_1 = \boldsymbol{x}_1$ is $\mathcal{N}(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})$.

Proof:

Let $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix}$ such that $\boldsymbol{AX} = \boldsymbol{X}_1$, and let $\boldsymbol{B} = \begin{bmatrix} -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{I} \end{bmatrix}$ such that $\boldsymbol{BX} = -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{X}_1 + \boldsymbol{X}_2$. Then $\boldsymbol{X}_2 = \boldsymbol{BX} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{X}_1$. $\boldsymbol{AX}$ and $\boldsymbol{BX}$ are independent since $\boldsymbol{A\Sigma B}^T = 0$.

Next, find $E(\boldsymbol{BX})$ and $Cov(\boldsymbol{BX})$,

$$E(\boldsymbol{BX}) = \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1$$
$$Cov(\boldsymbol{BX}) = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$$

Since $\boldsymbol{AX}$ and $\boldsymbol{BX}$ are independent,

$$\boldsymbol{BX}|(\boldsymbol{AX} = \boldsymbol{X}_1) = \boldsymbol{x}_1 \sim \boldsymbol{BX} \sim \mathcal{N}(\boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}),$$

Therefore

$$\boldsymbol{X}_2|\boldsymbol{X}_1 = \boldsymbol{x}_1 \sim \boldsymbol{BX} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{X}_{\boldsymbol{1}}|\boldsymbol{X}_1 = \boldsymbol{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}).$$

Exercise 1.2

a)

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

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

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

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

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 $X = (X_1, X_2, \dots, X_n)$ 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 1.8

a.In order to show that X_t is WN(0,1) two **thing** needs to be verified:

b. $Cov(X_t, X_{t+h}) = \begin{cases} 1 & \text{if } h = 0 \\ 0 & \text{if } h \neq 0 \end{cases}$

a. $\mathbf{E[X_t] = 0}$

$$E[X_t] = \begin{cases} E[Z_t] & \text{if t is even} \\ E\left[\frac{(Z_{t-1}^2 - 1)}{\sqrt{2}}\right] & \text{if t is odd} \end{cases}$$

- $E[Z_t] = 0$
- $E\left[\frac{(Z_{t-1}^2 - 1)}{\sqrt{2}}\right] = 0$ since $E[Z_{t-1}^2] = 1$

b. $Cov(X_t, X_{t+h}) = \begin{cases} 1 & \text{if } h = 0 \\ 0 & \text{if } h \neq 0 \end{cases}$

If $h = 0$,

$$Cov(X_t, X_t) = Var(X_t) = \begin{cases} Var(Z_t) & \text{if t is even} \\ \frac{1}{2}Var(Z_{t-1}^2 - 1) & \text{if t is odd} \end{cases}$$

- $Var(Z_t) = 1$
- $\frac{1}{2}Var(Z_{t-1}^2 - 1) = 1$ since $Var(Z_{t-1}^2) = 2$:

$$\begin{aligned} Var(Z_{t-1}^2) &= E(Z_{t-1}^4) - [E(Z_{t-1}^2)]^2 \\ &= 3 - 1 \quad \text{(Remind kurtosis of N(0,1) r.v's)} \\ &= 2 \end{aligned}$$

If $h \neq 0$, $Cov(X_t, X_{t+h}) = E(X_t X_{t+h})$,

$$E(X_t X_{t+h}) = \begin{cases} E(Z_t Z_{t+h}) & \text{if t is even and t+h is even} \\ \\ E\left(\frac{1}{\sqrt{2}} Z_t (Z_{t+h-1}^2 - 1)\right) & \text{if t is even and t+h is odd} \\ \\ E\left(\frac{1}{2}(Z_t^2 - 1)(Z_{t+h-1}^2 - 1)\right) & \text{if t is odd and t+h is odd} \end{cases}$$

- $E(Z_t Z_{t+h}) = E(Z_t)E(Z_{t+h}) = 0$
- $E(\frac{1}{\sqrt{2}}(Z_t(Z_{t+h-1}^2 - 1))) = \frac{1}{\sqrt{2}} E(Z_t) E(Z_{t+h-1}^2 - 1) = 0$
- $E(\frac{1}{2}(Z_{t-1}^2 - 1)(Z_{t+h-1}^2 - 1)) = \frac{1}{2} E(Z_{t-1}^2 - 1) E(Z_{t+h-1}^2 - 1) = 0$

In order to determine if X_t is $IID(0,1)$ we need to see if X_t and X_{t+h} are independent when $h \neq 0$.

Let's assume t is odd, then:

$$E(X_t) = E(\frac{Z_{t-1}^2 - 1}{\sqrt{2}}) = E(\frac{X_{t-1}^2 - 1}{\sqrt{2}})$$

Which means X_t depends on X_{t-1}. Thus, X_t is not $IID(0,1)$ noise.

b.

If n is odd:

$$E(X_{n+1}|X_1, \ldots, X_n) = E(Z_{n+1}|Z_0, Z_2, Z_4, \ldots, Z_{n+1}) = 0$$

If n is even:

$$E(X_{n+1}|X_1, \ldots, X_n) = E(\frac{Z_n^2 - 1}{\sqrt{2}}|Z_0, Z_2, \ldots, Z_n) = \frac{Z_n^2 - 1}{\sqrt{2}}$$

Chapter 2

Problem 2.1. We find the best linear predictor $\hat{X}_{n+h} = aX_n + b$ of X_{n+h} by finding a and b such that $\mathbb{E}[X_{n+h} - \hat{X}_{n+h}] = 0$ and $\mathbb{E}[(X_{n+h} - \hat{X}_{n+h})X_n] = 0$. We have

$$\mathbb{E}[X_{n+h} - \hat{X}_{n+h}] = \mathbb{E}[X_{n+h} - aX_n - b] = \mathbb{E}[X_{n+h}] - a\mathbb{E}[X_n] - b = \mu(1-a) - b$$

and

$$\begin{aligned}
&\mathbb{E}[(X_{n+h} - \hat{X}_{n+h})X_n] = \mathbb{E}[(X_{n+h} - aX_n - b)X_n] \\
&\quad = \mathbb{E}[X_{n+h}X_n] - a\mathbb{E}[X_n^2] - b\mathbb{E}[X_n] \\
&\quad = \mathbb{E}[X_{n+h}X_n] - \mathbb{E}[X_{n+h}]\mathbb{E}[X_n] + \mathbb{E}[X_{n+h}]\mathbb{E}[X_n] \\
&\qquad - a\left(\mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2 + \mathbb{E}[X_n]^2\right) - b\mathbb{E}[X_n] \\
&\quad = \operatorname{Cov}(X_{n+h}, X_n) + \mu^2 - a\left(\operatorname{Cov}(X_n, X_n) + \mu^2\right) - b\mu \\
&\quad = \gamma(h) + \mu^2 - a\left(\gamma(0) + \mu^2\right) - b\mu,
\end{aligned}$$

which implies that

$$b = \mu(1-a), \quad a = \frac{\gamma(h) + \mu^2 - b\mu}{\gamma(0) + \mu^2}.$$

Solving this system of equations we get $a = \gamma(h)/\gamma(0) = \rho(h)$ and $b = \mu(1 - \rho(h))$ i.e. $\hat{X}_{n+h} = \rho(h)X_n + \mu(1 - \rho(h))$.

Problem 2.4. a) Put $X_t = (-1)^t Z$ where Z is random variable with $\mathbb{E}[Z] = 0$ and $\operatorname{Var}(Z) = 1$. Then

$$\gamma_X(t+h, t) = \operatorname{Cov}((-1)^{t+h}Z, (-1)^t Z) = (-1)^{2t+h}\operatorname{Cov}(Z, Z) = (-1)^h = \cos(\pi h).$$

b) Recall problem 1.4 b) where $X_t = Z_1\cos(ct) + Z_2\sin(ct)$ implies that $\gamma_X(h) = \cos(ch)$. If we let Z_1, Z_2, Z_3, Z_4, W be independent random variables with zero mean and unit variance and put

$$X_t = Z_1\cos\left(\frac{\pi}{2}t\right) + Z_2\sin\left(\frac{\pi}{2}t\right) + Z_3\cos\left(\frac{\pi}{4}t\right) + Z_4\sin\left(\frac{\pi}{4}t\right) + W.$$

Then we see that $\gamma_X(h) = \kappa(h)$.
c) Let $\{Z_t : t \in \mathbb{Z}\}$ be WN $\left(0, \sigma^2\right)$ and put $X_t = Z_t + \theta Z_{t-1}$. Then $\mathbb{E}[X_t] = 0$ and

$$\begin{aligned}
&\gamma_X(t+h, t) = \operatorname{Cov}(Z_{t+h} + \theta Z_{t+h-1}, Z_t + \theta Z_{t-1}) \\
&\quad = \operatorname{Cov}(Z_{t+h}, Z_t) + \theta\operatorname{Cov}(Z_{t+h}, Z_{t-1}) + \theta\operatorname{Cov}(Z_{t+h-1}, Z_t) \\
&\qquad + \theta^2\operatorname{Cov}(Z_{t+h-1}, Z_{t-1}) \\
&\quad = \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } |h| = 1, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

If we let $\sigma^2 = 1/(1+\theta^2)$ and choose θ such that $\sigma^2\theta = 0.4$, then we get $\gamma_X(h) = \kappa(h)$. Hence, we choose θ so that $\theta/(1+\theta^2) = 0.4$, which implies that $\theta = 1/2$ or $\theta = 2$.

Problem 2.8. Assume that there exists a stationary solution $\{X_t : t \in \mathbb{Z}\}$ to

$$X_t = \phi X_{t-1} + Z_t, \qquad t = 0, \pm 1, \ldots$$

GT exercises

Problem 1

Part a

- The best linear approximation $\tilde{X} = \hat{E}(X|Y)$, which can be expressed as
$$\tilde{X} = \alpha_0 + \sum_{t \in T} \alpha_t Y_t$$
is characterized by the projection onto the space
$$\{X* \in L^2 : X^* = \alpha_0 + \sum_{t \in T} \alpha_t Y_t\}$$
The existence of α_0 and α_t, $t \in T$ is guaranteed by the projection theorem. These values can be found making use of the orthogonality of the residuals as stated in the projection. That is,
$$E(Y_k(X - \alpha_0 - \sum_{t \in T} \alpha_t Y_t)) = 0 \quad \forall k \in T$$
$$E((X - \alpha_0 - \sum_{t \in T} \alpha_t Y_t)) = 0$$

- As stated in definition 2.7.3 (Time Series: Theory and Methods), $\hat{X} = E(X|Y)$ can be understood as the projection on the closed subspace of L^2, $M(Y)$, of all the random variables in L^2 that can be written in the form $\phi(Y)$ with $\phi : \mathbb{R}^n \to \mathbb{R}$, a Borel function. That is:
$$E(X|Y) = E_{M(Y)}X = P_{M(Y)}X$$
It can be proved that it accomplishes all the properties of a projection. Based on these properties we can say:
$$X = P_{M(Y)}X + (X - P_{M(Y)}X)$$
Which implies, $E(\hat{X}(X - \hat{X})) = 0$ according to the projection theorem.

Part b

Given that X and Y are simple random variables defined on the probability space $(\Omega, \mathcal{F}, P)$, they can be expressed in the form:

$$X = \sum_{i=1}^{n} a_i 1_{A_i}$$

with 1 the indicator function and $A_i \in \mathcal{A}$ with $A_i = \{X = a_i\}$ (Similarly Y is defined). It can be proved that $P(X|Y = y) = Q_y(X)$ is a probability measure. Then,

$$E(X|Y = y) = \sum_{i=1}^{n} a_i Q_y(X = a_i)$$

Now, let $b_i = \phi(a_i)$. Then,

$$\begin{aligned} E(\phi(X)|Y = y) &= \sum_{i=1}^{n} b_i Q_y(X = a_i) \\ &= \sum_{i=1}^{n} b_i P(X = a_i|Y = y) \\ &= \sum_{i=1}^{n} b_i \frac{P(X = a_i, Y = y)}{P(Y = y)} \\ &= \sum_{i=1}^{n} b_i \frac{f_{X,Y}(a_i, y)}{f_Y(y)} \\ &= \sum_{i=1}^{n} b_i f_{X|Y}(a_i|y) \\ &= \sum_{i=1}^{n} \phi(a_i) f_{X|Y}(a_i|y) \end{aligned}$$

Thus,

$$E(\phi(X)|Y = y) = \textstyle\sum_x \phi(x) f_{X|Y}(x|y)$$