TSIA202A – Booklet of Exercises

Marco Cagnazzo, Mathieu Fontaine, François Roueff

Contents

1 General reminders and notation

1.1 Gaussian r.v.'s, vectors, processes

Except for the zero-variance case, a real valued **Gaussian random variable** X has the following probability density function (pdf):

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

A random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is a **Gaussian random vector** if and only if, $\forall u \in \mathbb{R}^n, Y = u^T\mathbf{X} = \sum_{i=1}^n u_i X_i$ is a Gaussian r.v. The pdf of a Gaussian vector is completely defined by the mean vector $\mu = \mathbb{E}[\mathbf{X}]$ and the covariance matrix $\Gamma = \mathbb{E}\left[\mathbf{X}^c \mathbf{X}^{cT}\right]$

A random process $\{X_t, t \in \mathbb{Z}\}$ is a **Gaussian random process** if and only if for all finite set of indexes $I \subset \mathbb{Z}, I = \{t_1, t_2, \ldots, t_n\}$, the random vector $[X_{t_1}, X_{t_2}, \ldots, X_{t_n}]^T$ is a Gaussian vector.

1.2 Functions of r.v.'s and of random processes

Let X be a real-valued r.v. and let g a real function. Let us suppose that g is derivable over $\mathbb{R}$, except for a set whose measure is zero, *e.g.*, a numerable set of points. If we define a new r.v. $Y = g(X)$, the pdf of Y is

Shortcut	Meaning
$\overline{X}$	Conjugate of X
A^T	Transpose pf A
A^H	Hermitian of A, *i.e.* $\overline{A^T}$
$\mathbb{N}_0$	Natural numbers including zero
$\mathbb{R}^+$	Positive real numbers: $\{x \in \mathbb{R} \| x > 0\}$
$\mathbb{R}_0^+$	Non-negative real numbers: $\{x \in \mathbb{R} \| x \geq 0\}$
$\mathbf{1}_A(x)$	Indicator function of set A: $\mathbf{1}_A(x) = 1$ if and only if $x \in A$; otherwise, $\mathbf{1}_A(x) = 0$
r.v.	random variable
pdf	probability density function
$X \sim P$	X is a r.v. distributed with law P
$\mathcal{N}\left(\mu, \sigma^2\right)$	Gaussian r.v. with mean μ and variance σ^2
$\mathbb{E}\left[X\right]$	Expectation of the r.v. X
X^c	Centered version of X: $X^c = X - \mathbb{E}\left[X\right]$
$\mathsf{Var}\left(X\right)$	Variance of the r.v. X: $\mathsf{Var}\left(X\right) = \mathbb{E}\left[\|X^c\|^2\right]$
$\mathsf{Cov}\left(X, Y\right)$	$\mathbb{E}\left[X^c\overline{Y^c}\right]$
$\{X_t, t \in \mathbb{Z}\}$	Discrete random process
s.o.1	A process $\{X_t, t \in \mathbb{Z}\}$ is *stationary at order 1* if and only if $\mathbb{E}\left[X_t\right]$ does not depend on t
s.o.2	A process $\{X_t, t \in \mathbb{Z}\}$ is *stationary at order 2*, if and only if $\forall t \in \mathbb{Z}$, $\mathbb{E}\left[\|X_t\|^2\right] < +\infty$ and $\forall t, h \in \mathbb{Z}, \mathsf{Cov}\left(X_t, X_{t+h}\right)$ does not depend on t
w.s.	weakly stationary, *i.e.*, s.o.1 and s.o.2
$\gamma_X(h)$	For $\{X_t, t \in \mathbb{Z}\}$ s.o.2, $\gamma_X(h) = \mathsf{Cov}\left(X_{t+h}, X_t\right) = \mathsf{Cov}\left(X_h, X_0\right)$
δ_h	The Kronecker's delta: $\delta : h \in \mathbb{Z} \to \delta_h$; if $h = 0$, $\delta_h = 1$; otherwise, $\delta_h = 0$

Table 1: Shortcuts and notation used throughout this document.

related to that of X as follows:

$$p_Y(y) = \begin{cases} 0 & \text{if the equation in the variable } x,\ g(x) = y, \text{ has no solution} \\ \sum_{i=1}^{N_y} \frac{p_X(x_i(y))}{|g'(x_i(y))|} & \text{if } g(x) = y \text{ has } N_y \geq 1 \text{ solutions, referred to as } \{x_i(y)\}_{i=1,\ldots,N_y} \end{cases}$$

We can also consider function of multiple r.v.'s. A particularly interesting case is when a random process is obtained by applying a function to another random process:

$$X_t = g_t(\{Z_s, s \in \mathbb{Z}\})$$

A special case is when the transformation is the same at each time (*i.e.* g does not depend on t) and it has a finite number of inputs. Apart from a time shift, this can be written as:

$$X_t = g(Z_t, Z_{t-1}, \ldots, Z_{t-k+1})$$

This is called a *moving transformation*. It can be shown that, for a moving transformation, if g is measurable and $\{Z_t, t \in \mathbb{Z}\}$ i.i.d., then $\{X_t, t \in \mathbb{Z}\}$ is strictly stationary.

A particularly interesting case of moving transformation is a linear filter:

$$Y_t = \sum_{n \in \mathbb{Z}} \alpha_n X_{t-n}$$

If the support of α is finite, this filter is called Finite Impulse Response (FIR); otherwise it is an Infinite Impulse Response (IIR).

1.2.1 Example: inversion of a FIR

Let us remember a particularly simple case of invertible filter. Let $\theta \in \mathbb{C}$ and $|\theta| < 1$. We introduce the following L^1 sequences:

$$
\begin{aligned}
a : n \in \mathbb{Z} &\to \delta_n - \theta\delta_{n-1} \\
b : n \in \mathbb{Z} &\to \begin{cases} \theta^n & \text{if } n \geq 0 \\ 0 & \text{otherwise} \end{cases} \\
c = (a * b) : n \in \mathbb{Z} &\to \sum_{k\in\mathbb{Z}} a_k b_{n-k}
\end{aligned}
$$

It is easy to find that $(a * b) = \delta$. In that case, we say that a FIR having a as impulse response, can be inverted by an IIR having b as impulse response, since the cascade of a and b will not change an input signal. Let us show that $c = \delta$.

$$
c_n = \sum_{k\in\mathbb{Z}} a_k b_{n-k} = b_n - \theta \cdot b_{n-1} = \begin{cases} 0 - \theta \cdot 0 = 0 & \text{if } n < 0 \\ 1 - \theta \cdot 0 = 1 & \text{if } n = 0 \\ \theta^n - \theta \cdot \theta^{n-1} = 0 & \text{if } n > 0 \end{cases} = \delta_n
$$

1.3 Autocovariance

$$
\begin{aligned}
\mathsf{Cov}\,(X,Y) &= \mathbb{E}\left[X^c\overline{Y^c}\right] \\
\mathsf{Cov}\,(X,Y) &= \mathbb{E}\left[X\overline{Y}\right] - \mathbb{E}\,[X]\,\mathbb{E}\left[\overline{Y}\right] \\
\overline{\mathsf{Cov}\,(X,Y)} &= \mathsf{Cov}\,(Y,X) \\
\mathsf{Cov}\,(X+a,Y) &= \mathsf{Cov}\,(X,Y) \\
\mathsf{Cov}\,(X,Y+a) &= \mathsf{Cov}\,(X,Y) \\
\mathsf{Cov}\,(aX,Y) &= a\mathsf{Cov}\,(X,Y) \\
\mathsf{Cov}\,(X,aY) &= \overline{a}\mathsf{Cov}\,(X,Y) \\
\mathsf{Cov}\,(X_1+X_2,Y) &= \mathsf{Cov}\,(X_1,Y) + \mathsf{Cov}\,(X_2,Y) \\
\mathsf{Cov}\,(X,Y_1+Y_2) &= \mathsf{Cov}\,(X,Y_1) + \mathsf{Cov}\,(X,Y_2)
\end{aligned}
$$

Table 2: Covariance properties. X, X_1, X_2, Y are complex or real r.v.'s; $a \in \mathbb{C}$.

The covariance of two r.v.'s has several interesting properties resumed in Tab. 2. Two real r.v. with null covariance are said to be *uncorrelated*. Two complex r.v. with null covariance are said to be *orthogonal*, while if also the *pseudo-covariance* $\mathbb{E}\,[XY]$ is null, they are said *uncorrelated*. Independent r.v.'s are uncorrelated while the converse is not true in general. A notable exception is when (X,Y) is a Gaussian vector (but not when X and Y are marginally Gaussian and not jointly Gaussian): in that case, uncorrelatedness implies independence.

The covariance allows to define a scalar product between two r.v.'s: $\langle X_t, X_s\rangle = \mathsf{Cov}\,(X_t, X_s)$. The (squared) norm of a r.v.'s is then its variance. Note that this scalar product is not affected by the mean of the r.v.'s, since neither the covariance is. For example, a zero-norm r.v. has a null variance, but can have any mean.

We can also introduce the concept of linear independent r.v.'s. $(X_1, \ldots, X_k)$ is a set of linearly independent r.v.'s if and only if $\forall a \in \mathbb{R}^k - \{\mathbf{0}\}, \|\sum_{i=1}^k a_i X_i\|^2 = \mathsf{Var}\left(\sum_{i=1}^k a_i X_i\right) > 0$.

Note also that, if $(X_1, \ldots, X_k)$ are not linearly independent, this means that one of the X_i can be expressed as a linear combination of the other r.v.'s, up to an additive constant, which does not affect the covariance. This constant is null in the case $\mathbb{E}\,[(X_1, \ldots, X_k)] = \mathbf{0}$.

For the sake of simplicity, let us prove that for some i, X_i is a linear combination of the other r.v.'s *only in the case of a centered vector*. In this case it must exist $a \in \mathbb{R}^k - \{\mathbf{0}\}$ such that $\mathsf{Var}\left(\sum_{i=1}^k a_i X_i\right) = 0$. The vector a must have at least one non-zero component, let it be a_j. Let also $Y = \sum_{i=1}^k a_i X_i$; since its variance is zero, $Y = \mathbb{E}[Y] = 0$. This implies:

$$\begin{aligned}
0 = \sum_{i=1}^{k} a_i X_i &= a_j X_j + \sum_{i \neq j} a_i X_i \\
a_j X_j &= -\sum_{i \neq j} a_i X_i \\
X_j &= -\sum_{i \neq j} \frac{a_i}{a_j} X_i
\end{aligned}$$

Then X_j is a linear combination of other r.v.'s. It can be shown that, if the X_i are not centered, the same result holds up to a constant: $X_j = -\sum_{i \neq j} \frac{a_i}{a_j} X_i + \sum_{i=1}^{k} \frac{a_i}{a_j} \mathbb{E}[X_i]$.

The **covariance matrix** of a complex-valued random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is $\Gamma = \mathbb{E}\left[\mathbf{X}^c \mathbf{X}^{cH}\right]$. In other words, $\Gamma_{i,j} = \mathsf{Cov}(X_i, X_j)$. It is an Hermitian, non-negative matrix, since for all $u \in \mathbb{C}^n$ the random variable $Y = u^H X$ shall have a non negative variance:

$$\begin{aligned}
0 \leq \mathsf{Var}(Y) &= \mathbb{E}\left[\|u^H \mathbf{X} - \mathbb{E}\left[u^H \mathbf{X}\right]\|^2\right] = \mathbb{E}\left[\|u^H (\mathbf{X} - \mathbb{E}[\mathbf{X}]\|^2\right] \\
&= \mathbb{E}\left[\|u^H \mathbf{X}^c\|^2\right] = \mathbb{E}\left[u^H \mathbf{X}^c \mathbf{X}^{cH} u\right] \\
&= u^H \mathbb{E}\left[\mathbf{X}^c \mathbf{X}^{cH}\right] u = u^H \Gamma u
\end{aligned}$$

The autocovariance function (acf) of a random process $\{X_t, t \in \mathbb{Z}\}$ is a function of two discrete variables t and s:

$$\gamma(t, s) = \mathsf{Cov}(X_t, X_s)$$

A **weakly stationary process** is a process s.o.1 and s.o.2, therefore, all X_t have finite quadratic mean, the mean of X_t is the same for all t and the autocovariance function only depend on the delay $t - s$:

$$\gamma(t, s) = \gamma(t - s) = \mathsf{Cov}(X_{t-s}, X_0)$$

In that case, we use a single-parameter notation for γ:

$$\gamma(h) = \mathsf{Cov}(X_h, X_0)$$

The acf of weakly stationary processes is an Hermitian and non-negative function. The maximum of $|\gamma|$ is in 0. The normalized acf, $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$ is referred to as autocorrelation function.

1.4 Noise

A **weak white noise** is a real-valued, weakly stationary process $\{X_t, t \in \mathbb{Z}\}$, with zero-mean and impulsive acf: $\gamma_X(h) = \sigma^2 \delta(h)$. In other words, for all $t \neq s$, X_t and X_s are uncorrelated variables.

A **strong white noise** is a real-valued, zero-mean, i.i.d. process. Note that a strong white noise is also a weak white noise, since i.i.d. implies weak stationarity and impulsive acf. On the contrary, a weak white noise is not necessarily a strong one, since uncorrelated r.v.'s may be dependent.

In both cases, we usually consider finite, positive variance $\sigma^2 = \mathsf{Var}(X_t)$.

2 Gaussian vectors

Exercise 2.1 (Functions of Gaussian random variables). Let $X \sim \mathcal{N}(0, 1)$, $a \in \mathbb{R}^+$ and $Y^a = X\mathbf{1}_{\{|X|<a\}} - X\mathbf{1}_{\{|X|\geq a\}}$.

1. Give the law of Y^a
2. Compute $\mathsf{Cov}(X, Y^a)$. For which value a_0 of a the covariance is null? Are X and Y^{a_0} independent?
3. Is (X, Y^{a_0}) a Gaussian vector?
4. For $a \neq a_0$, is (X, Y^a) a Gaussian vector?

3 Stationarity

Exercise 3.1 (Uncorrelated processes). Let $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ be two weakly stationary (w.s.), uncorrelated random processes. Show that $\{Z_t = X_t + Y_t, t \in \mathbb{Z}\}$ is weakly stationary. Find the covariance function of Z_t from those of X_t and Y_t and the spectral measure of Z_t from those of X_t and Y_t

Exercise 3.2 (Functions of strong white noise). Let $\{\epsilon_t, t \in \mathbb{Z}\}$ be a strong white noise with $\mathbb{E}\left[\epsilon_0^2\right] < \infty$. For each of the following processes (functions of the white noise), find out if they are weakly stationary or strictly stationary.

1. $W_t = a + b\epsilon_t + c\epsilon_{t-1}$, with a, b, c real numbers
2. $X_t = \epsilon_t \epsilon_{t-1}$
3. $Y_t = (-1)^t \epsilon_t$
4. $Z_t = \epsilon_t + Y_t$

Exercise 3.3 (Structured covariance matrix). Let us consider a real number ρ; we define $\Sigma_2 = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. Moreover, let $\forall t \in \mathbb{Z}$, Σ_t be a $t \times t$ matrix with diagonal elements equal to 1, and out-of-diagonal elements equal to ρ:

$$\Sigma_t = \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \dots & \dots & \dots & \dots & \dots \\ \rho & \rho & \rho & \dots & 1 \end{bmatrix}$$

1. Which condition on ρ must be imposed such that Σ_t is a covariance matrix for all t? Suggestion: decompose $\Sigma_t = \alpha I + A$, where A is matrix with easy-to-find eigenvalues.
2. Build a stationary process having Σ_t as auto-covariance matrix for all t.

4 Covariance, spectral measure and spectral density

Exercise 4.1 (Functions of weak white noise). Let $\{Z_t, t \in \mathbb{Z}\}$ be a weak white noise, centered, with variance σ^2. Let $a, b, c \in \mathbb{R}$. Are the following processes s.o.2? If yes, compute the autocovariance function and the spectral measure.

1. $X_t = a + bZ_0$
2. $X_t = Z_0 \cos(ct)$
3. $X_t = a + bZ_t + cZ_{t-1}$
4. $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$
5. $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$

Reminders:

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} \nu(d\lambda)$$

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda \qquad \text{if the density } f(\cdot) \text{ exists}$$

$$f(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma(h) e^{-ih\lambda} \qquad \text{if } \gamma \in L^1(\mathbb{Z})$$

Exercise 4.2 (Autocovariance function characterization). Let us introduce the following sequence on the integers:

$$\gamma : h \in \mathbb{Z} \to \gamma(h) = \begin{cases} 1 & \text{if } h = 0 \\ \rho & \text{if } |h| = 1 \\ 0 & \text{if } |h| > 1 \end{cases}$$

We want to show that such a function is an autocovariance function if and only if $|\rho| \leq \frac{1}{2}$.

1. Let Γ_k be a $k \times k$ matrix such that $\forall i, j \in \{1, 2, \ldots, k\}, \Gamma_k(i, j) = \gamma(i - j)$.

$$\Gamma_k = \begin{bmatrix} 1 & \rho & 0 & 0 & \ldots & 0 & 0 \\ \rho & 1 & \rho & 0 & \ldots & 0 & 0 \\ 0 & \rho & 1 & \rho & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & \rho & 1 \end{bmatrix}$$

   Find the recurrence equation among the determinants of matrices Γ_k

2. Show that if $|\rho|$ is not greater than a given value, Γ_k is positive definite for all k. Use or the previous point or the Herglotz theorem.

3. Build a s.o.2 process having $\gamma(h)$ as autocovariance function. [**Hint:** use Question 3 of Exercise 4.2.]

Exercise 4.3 (Band-limited stationary process). Let $S(f) = \mathbf{1}_{(-f_0, f_0)}(f)$, with $f_0 \in (0, \pi)$ be the spectral density of a stationary process.

1. Compute the autocovariance function.

2. Is it ℓ^1?

Exercise 4.4 (Process generated by linear combination). Let γ be the autocovariance function of a stationary, zero-mean process. Let us suppose that it exist a finite subset of this process such that the corresponding autocovariance matrix is not invertible, *i.e.*, it is not full rank.

1. Show that either $\gamma(0) = 0$, or it exists $k \geq 1$ such that:

   - $X_{k+1} \in \mathsf{Vect}(X_1, \ldots, X_k)$; and
   - $(X_1, \ldots, X_k)$ is a set of linearly independent vectors: $\forall a \in \mathbb{R}^k - \{\mathbf{0}\}, \mathsf{Var}\left(\sum_{i=1}^k a_i X_i\right) > 0$.

2. Let Γ_k be the autocovariance matrix of $X_1, \ldots, X_k$. Find a property of its minimum eigenvalue.

3. Show that the process $\{X_t, t \in \mathbb{Z}\}$ is linearly predictable, *i.e.*, for all $p \geq 1$, there exists a set of k scalars $\phi_{p,1}, \phi_{p,2}, \ldots, \phi_{p,k}$ such that:

$$X_{k+p} = \sum_{\ell=1}^{k} \phi_{p,\ell} X_\ell. \tag{1}$$

4. Show that $\sup_{p \geq 1} \sum_{\ell=1}^k |\phi_{p,\ell}|^2 < \infty$.

5. Deduce that, if in addition $\lim_{|t| \to \infty} \gamma(t) = 0$, then $\gamma(0) = 0$.

5 Linear filtering, ARMA processes

Exercise 5.1 (Linear filtering and stationarity). Let $\beta \in \mathbb{R}$, $\{S_t, t \in \mathbb{Z}\}$ a w.s., periodical (period = 4) real process, and $\{X_t, t \in \mathbb{Z}\}$ a w.s. real process, uncorrelated with S_t.

Let us consider the process $\{Y_t = \beta t + S_t + X_t, t \in \mathbb{Z}\}$.

1. Is $\{Y_t, t \in \mathbb{Z}\}$ w.s.?

2. Let us refer to the back-shift operator as B, and let us consider the process $\{\bar{S}_t = \left(1 + B + B^2 + B^3\right) \circ S_t, t \in \mathbb{Z}\}$. Show that γ is periodic and that $\bar{S}_t = S_0 + S_1 + S_2 + S_3$
3. Let us consider the process $\{Z_t = (1 - B) \circ \left(1 + B + B^2 + B^3\right) \circ S_t, t \in \mathbb{Z}\}$. Show that $\{Z_t, t \in \mathbb{Z}\}$ is w.s. and compute γ_Z as a function of γ_X (autocovariance functions).
4. Find the shape of the spectral measure μ of $\{S_t, t \in \mathbb{Z}\}$.
5. Find the spectral measure of $(1 - B^4) \circ Y_t$ as a function of the spectral measure of $\{X_t, t \in \mathbb{Z}\}$.

Exercise 5.2 (Characterization of MA(q)). Let $q \in \mathbb{Z}$ and $q > 0$. Let $\{X_t, t \in \mathbb{Z}\}$ be a centered w.s. real process and let γ be its autocovariance function. Let us suppose that γ has a compact support, *i.e.* $\forall t > q, \gamma(t) = 0$.

We also introduce

$$\mathcal{H}_t = \mathsf{Vect}(X_s, s \leq t)$$
$$\widetilde{X}_t = \mathsf{Proj}(X_t | \mathcal{H}_{t-1})$$

1. Recall why $Z_t = X_t - \widetilde{X}_t$ is a white noise.
2. Show that $X_t \perp \mathcal{H}_{t-q-1}$.
3. Deduce that $X_t \in \mathsf{Vect}(Z_s, s \in \{t, t-1, \dots t-q\})$.
4. Show that $\{X_t, t \in \mathbb{Z}\}$ is a MA(q) process.

Exercise 5.3 (Sum of MA processes). Let $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ be two real uncorrelated MA processes of order q and p respectively:

$$X_t = \epsilon_t + \sum_{n=1}^{q} \theta_n \epsilon_{t-n} \qquad\qquad Y_t = \eta_t + \sum_{n=1}^{p} \rho_n \eta_{t-n}$$

where $\forall n \in \{1, \dots, q\}, \theta_n \in \mathbb{R}$, $\forall n \in \{1, \dots, p\}, \rho_n \in \mathbb{R}$, $\{\epsilon_t, t \in \mathbb{Z}\}$ and $\{\eta_t, t \in \mathbb{Z}\}$ are white noises whose variances are respectively noted as σ_ϵ^2 and σ_η^2. Let us also introduce $\{Z_t = X_t + Y_t, t \in \mathbb{Z}\}$.

1. Which kind of process is $\{Z_t, t \in \mathbb{Z}\}$?
2. Let us consider the case p=1 , q=1, 0<θ_1<1 and 0<ρ_1<1. Show that $\{\epsilon_t, t \in \mathbb{Z}\}$ and $\{\eta_t, t \in \mathbb{Z}\}$ are uncorrelated.
3. For p=1 , q=1, θ_1=ρ_1=θ and 0<θ<1, what is the innovation process for $\{Z_t, t \in \mathbb{Z}\}$?
4. For p=1 , q=1, 0<θ_1<1 and 0<ρ_1<1, compute the variance of the innovation of $\{Z_t, t \in \mathbb{Z}\}$.

Exercise 5.4 (Sum of AR processes). Let $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ be two real uncorrelated AR(1) processes such that

$$X_t = aX_{t-1} + \epsilon_t$$
$$Y_t = bY_{t-1} + \eta_t$$

where $a \in]0, 1[$, $b \in]0, 1[$. Moreover, $\{\epsilon_t, t \in \mathbb{Z}\}$ and $\{\eta_t, t \in \mathbb{Z}\}$ are white noises whose variances are respectively noted as σ_ϵ^2 and σ_η^2. Let us also introduce $\{Z_t = X_t + Y_t, t \in \mathbb{Z}\}$.

1. Show that there exists a white noise $\{\xi_t, t \in \mathbb{Z}\}$ and a real number $\theta \in]-1, 1[$ such that:

$$Z_t - (a + b)Z_{t-1} + abZ_{t-2} = \xi_t - \theta\xi_{t-1}.$$

2. Show that:

$$\xi_t = \epsilon_t + (\theta - b)\sum_{k=0}^{\infty}\theta^k \epsilon_{t-1-k} + \eta_t + (\theta - a)\sum_{h=0}^{\infty}\theta^h \eta_{t-1-h}.$$

3. Compute the prediction of Z_{t+1} when $(X_s, s \le t)$ and $(Y_s, s \le t)$ are all known.

4. Compute the prediction of Z_{t+1} when $(Z_s, s \le t)$ are all known.

5. Compare the variances of the prediction errors in the two previous cases.

Exercise 5.5 (Forward/backward prediction of a MA(1) process). Let $\{X_t = Z_t + \theta Z_{t-1}, t \in \mathbb{Z}\}$ be a real w.s. process, with $\{Z_t, t \in \mathbb{Z}\}$ centered white noise and $\theta \in]-1, 1[$.

1. Find the best (in terms of MSE) linear prediction of X_3 as a function of X_1 and X_2.

2. Find the best linear prediction of X_3 as a function of X_4 and X_5.

3. Find the best linear prediction of X_3 as a function of X_1, X_2, X_4 and X_5.

Exercise 5.6 (Canonical representation of an ARMA process). Let $\{X_t, t \in \mathbb{Z}\}$ be a centered, s.o.2 process satisfying the recurrence equation

$$X_t - 2X_{t-1} = \epsilon_t + 4\epsilon_{t-1}$$

where $\{\epsilon_t, t \in \mathbb{Z}\}$ is a white noise with variance σ^2.

1. Compute the spectral density of $\{X_t, t \in \mathbb{Z}\}$.

2. Compute the canonical representation of $\{X_t, t \in \mathbb{Z}\}$.

3. What is the variance of the innovation of $\{X_t, t \in \mathbb{Z}\}$?

4. Find a representation of X_t as a function of $(\epsilon_s, s \le t)$.

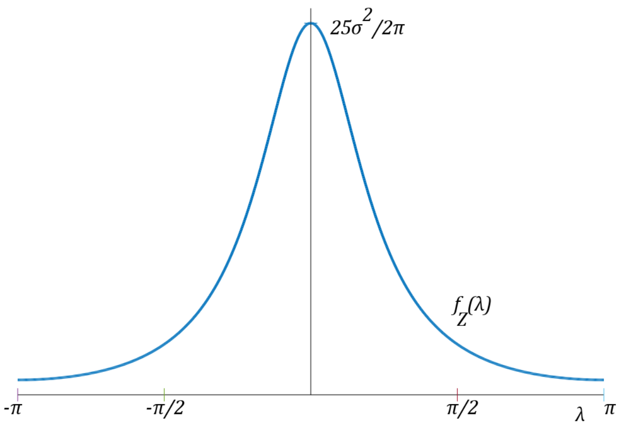

Figure 1: $f_Z(\lambda) = \frac{\sigma^2}{2\pi}\frac{8\cos\lambda + 17}{5 - 4\cos\lambda}$

Exercise 5.7 (ACF of an AR(1) process). Let $\{X_t, t \in \mathbb{Z}\}$ be a w.s. process defined by:

$$X_t - \phi X_{t-1} = \epsilon_t$$

where $\phi \in]-1, 1[$ and $\{\epsilon_t, t \in \mathbb{Z}\}$ is a centered WN with variance σ_ϵ^2.

1. Compute the weights ψ_i of the representation

$$X_t = \sum_{k \in \mathbb{Z}} \psi_k \epsilon_{t-k}$$

2. Deduce the autocovariance function of $\{X_t, t \in \mathbb{Z}\}$.

6 Solutions

Solution of Exercise 2.1 1. The r.v. Y satisfies the following equation: $Y = \begin{cases} X & \text{if } |X| < a \\ -X & \text{if } |X| > a \end{cases}$

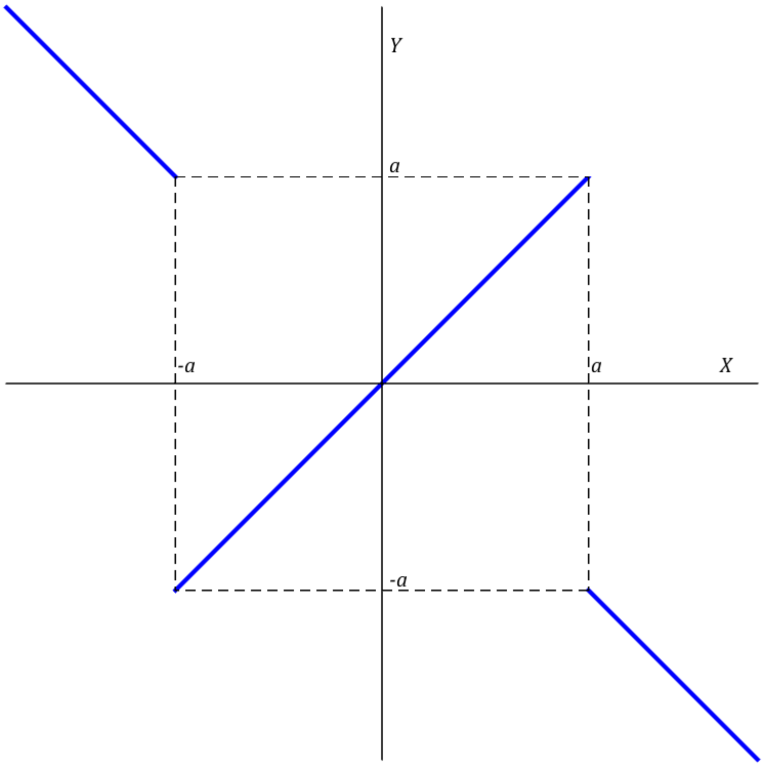

Figure 2: $Y = g(X)$

$$\text{If } |y| < a \qquad p_Y(y) = p_X(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$$

$$\text{If } |y| > a \qquad p_Y(y) = p_X(-y) = \frac{e^{-\frac{(-y)^2}{2}}}{\sqrt{2\pi}}$$

Thus, $Y \sim \mathcal{N}(0,1)$

2. Let us compute the covariance of X and Y^a:

$$\begin{aligned}
\mathsf{Cov}(X, Y^a) &= \mathbb{E}[XY^a] = \mathbb{E}\left[X^2 \mathbf{1}_{\{|X|<a\}} - X^2 \mathbf{1}_{\{|X|\geq a\}}\right] \\
&= \mathbb{E}\left[X^2 \left(\mathbf{1}_{\{|X|<a\}} - \mathbf{1}_{\{|X|\geq a\}}\right)\right] = \mathbb{E}\left[X^2 \left(2\mathbf{1}_{\{|X|<a\}} - 1\right)\right] \\
&= 2\mathbb{E}\left[X^2 \mathbf{1}_{\{|X|<a\}}\right] - \mathbb{E}\left[X^2\right] = \sqrt{\frac{2}{\pi}} \int_{-a}^{a} x^2 e^{-\frac{x^2}{2}}\, dx - 1 = h(a)
\end{aligned}$$

The function $h : a \to h(a)$ is continuous and strictly increasing. Moreover $h(0) = -1$ and $\lim_{a\to+\infty} h(a) = \mathbb{E}\left[X^2\right] = 1$. Therefore, $\exists a_0 \in]0, +\infty[$: $h(a_0) = 0$. For such a value a_0, X and Y^{a_0} are uncorrelated but they are not independent, since $Y|X$ is deterministic. Another way to show that X and Y^{a_0} are not independent is the following. Since they are both Gaussian, if they were independent, the vector (X, Y^{a_0}) would be a Gaussian Vector, therefore $X + Y^{a_0}$ would be Gaussian. But this is impossible, since $X + Y^{a_0} = 2X\mathbf{1}_{|X|<a_0}$ cannot be larger than $2a_0$. This also answers to points 3. As for point 4, the since $X + Y^a$ is not a Gaussian r.v. for any real positive a, the vector (X, Y^a) cannot be a Gaussian vector.

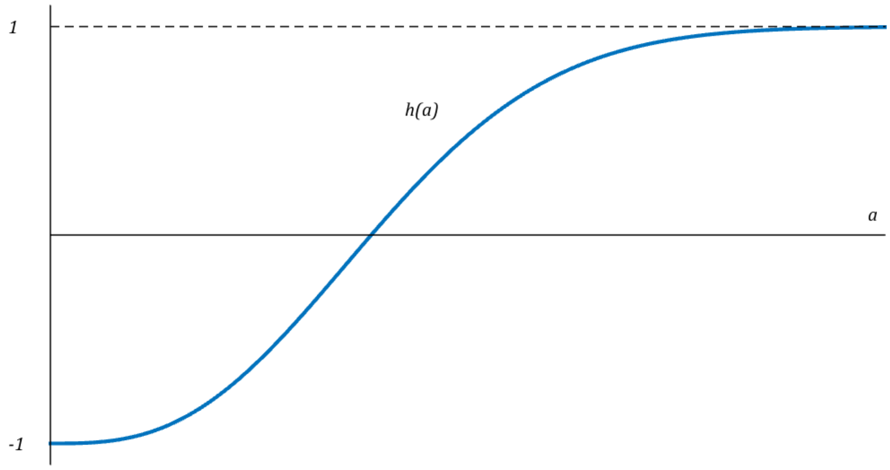

Figure 3: Function $h(a) = \mathsf{Cov}\left(X, Y^a\right)$

Solution of Exercise 3.1 First, since $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ are w.s.,

$$\mathbb{E}\left[X_t\right] = \mu_X \qquad \mathbb{E}\left[Y_t\right] = \mu_Y$$

$$\mathsf{Cov}\left(X_t, X_s\right) = \gamma_X(t-s) \qquad \mathsf{Cov}\left(Y_t, Y_s\right) = \gamma_Y(t-s)$$

Moreover, $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ are uncorrelated, meaning that $\forall t, s, \mathsf{Cov}\left(X_t, Y_s\right) = 0$, therefore we find:

$$\begin{aligned}
\mathbb{E}\left[Z_t\right] &= \mathbb{E}\left[X_t + Y_t\right] = \mu_X + \mu_Y \\
\mathsf{Cov}\left(Z_t, Z_s\right) &= \mathsf{Cov}\left(X_t + Y_t, X_s + Y_s\right) = \mathsf{Cov}\left(X_t, X_s\right) + \mathsf{Cov}\left(X_t, Y_s\right) + \mathsf{Cov}\left(Y_t, X_s\right) + \mathsf{Cov}\left(Y_t, Y_s\right) \\
&= \gamma_X(t-s) + \gamma_Y(t-s)
\end{aligned}$$

Therefore $\{Z_t, t \in \mathbb{Z}\}$ is w.s. with $\mathbb{E}\left[Z_t\right] = \mu_X + \mu_Y$ and $\gamma_Z(h) = \gamma_X(h) + \gamma_Y(h)$. From the previous point we deduce that the spectral measure of $\{Z_t, t \in \mathbb{Z}\}$ is the sum of those of $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$.

Solution of Exercise 3.2 We remind that if g is a measurable moving transformation, it preserves the strict stationarity, meaning that, since $\{\epsilon_t, t \in \mathbb{Z}\}$ is strictly stationary, so $g(\epsilon)$ is.

1. and 2. We are in the case of a moving transformation. In both cases g is measurable, so $\{W_t, t \in \mathbb{Z}\}$ and $\{X_t, t \in \mathbb{Z}\}$ are strictly stationary.

3. This is not a moving transformation. Actually, Y_t is alternatively equal to ϵ_t and $-\epsilon_t$. Since $\{\epsilon_t, t \in \mathbb{Z}\}$ is iid, the pdf of Y_t is

$$p_Y(y) = \begin{cases} p_\epsilon(y) & \text{if } t \text{ is even} \\ p_\epsilon(-y) & \text{if } t \text{ is odd} \end{cases}$$

Therefore, if the pdf of ϵ_t is symmetric, $\{Y_t, t \in \mathbb{Z}\}$ is iid; otherwise, it is not strictly stationary.

As for weak stationarity, it is achieved if $\mathbb{E}\left[\epsilon_t\right] = 0$. This actually implies that $\mathbb{E}\left[Y_t\right] = 0$. Moreover,

$$\mathsf{Cov}\left(Y_t, Y_s\right) = \begin{cases} \mathbb{E}\left[\epsilon_0^2\right] & \text{if } t = s \\ \mathsf{Cov}\left(\pm\epsilon_t, \pm\epsilon_s\right) = 0 & \text{otherwise} \end{cases}$$

Thus, Y_t is w.s. if $\mathbb{E}\left[\epsilon_t\right] = 0$.

4. In that case,$Z_t = 2\epsilon_t$ if t is even, and $Z_t = 0$ if t is odd ,implying that:

$$\mathbb{E}\left[Z_t\right] = \begin{cases} 0 & \text{if } t \text{ is even} \\ 0 & \text{if } t \text{ is odd} \end{cases} \qquad \mathsf{Var}\left(Z_t\right) = \begin{cases} 4\sigma_\epsilon^2 & \text{if } t \text{ is even} \\ 0 & \text{if } t \text{ is odd} \end{cases}$$

Therefore $\{Z_t, t \in \mathbb{Z}\}$ is s.o.1, but it is s.o.2 if and only if $\sigma_\epsilon^2 = 0$: in that case, $\epsilon_t = Z_t = 0$ for all t.

Solution of Exercise 3.3 1. A covariance matrix is an Hermitian, non-negative matrix. Since ρ is real, matrices Σ_t are Hermitian. As for non-negativity, it is equivalent to the fact that the eigenvalues of Σ_t, let them be $\{\lambda_1, \lambda_2, \ldots, \lambda_t\}$, are all non-negative.

Let us define A as a $t \times t$ matrix such that $A_{i,j} = \rho$ for all i and j. Then we have $\Sigma_t = (1-\rho)I_t + A$. Now, $\lambda_i = (1-\rho) + \omega_i$, where ω_i is the i-th eigenvalue of A. Since the rank of A is 1, $t-1$ of its eigenvalues are equal to 0. Let us say that ω_t is the remaining, non null eigenvalue. Moreover, $\mathsf{Tr}(A) = \sum_{i=1}^t \omega_i = \omega_t$, but also $\mathsf{Tr}(A) = t\rho$, thus $\omega_t = t\rho$. In conclusions we have

$$\begin{aligned} \forall i \in \{1, 2, \ldots, t-1\}, \lambda_i &= 1-\rho \\ \lambda_t &= 1 - \rho + t\rho = 1 + (t-1)\rho \end{aligned}$$

The non-negativity conditions are:

$$\begin{aligned} 1 - \rho &\geq 0 & 1 + (t-1)\rho &\geq 0 \\ \rho &\leq 1 & \rho &\geq -\frac{1}{t-1} \to_{t\to+\infty} 0^- \end{aligned}$$

In conclusion, $0 \leq \rho \leq 1$.

2. Let us consider a process $\{X_t = \alpha\epsilon_t + \beta Z, t \in \mathbb{Z}\}$, with $\{\epsilon_t, t \in \mathbb{Z}\}$ being a real-valued, zero-mean, unitary-variance strong white noise, Z a real-valued, zero-mean, unitary-variance r.v. independent from any ϵ_t, and $\alpha, \beta \in \mathbb{R}$. We would have:

$$\mathsf{Cov}\left(X_t, X_{t+h}\right) = \mathbb{E}\left[(\alpha\epsilon_t + \beta Z)(\alpha\epsilon_{t+h} + \beta Z)\right] = \alpha^2\mathbb{E}\left[\epsilon_t\epsilon_{t+h}\right] + \beta^2\mathbb{E}\left[Z^2\right] = \alpha^2\delta_h + \beta_2$$

$$\Sigma_t = \begin{bmatrix} \alpha^2+\beta^2 & \beta^2 & \beta^2 & \ldots & \beta^2 \\ \beta^2 & \alpha^2+\beta^2 & \beta^2 & \ldots & \beta^2 \\ \beta^2 & \beta^2 & \alpha^2+\beta^2 & \ldots & \beta^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \beta^2 & \beta^2 & \beta^2 & \ldots & \alpha^2+\beta^2 \end{bmatrix}$$

$$\begin{aligned} \alpha^2 + \beta^2 &= 1 & \beta^2 &= \rho \\ \alpha^2 &= 1-\rho & \beta^2 &= \rho \\ \alpha &= \sqrt{1-\rho} & \beta &= \sqrt{\rho} \end{aligned}$$

Since $\rho \in [0,1]$, then also $\alpha, \beta \in [0,1]$.

Solution of Exercise 4.1 1. $X_t = a + bZ_0$ is a constant with respect to t, thus strictly stationary.

$$\mathbb{E}\left[X_t\right] = a \qquad \mathsf{Cov}\left(X_t, X_{t+h}\right) = \mathsf{Cov}\left(a + bZ_0, a + bZ_0\right) = b^2\sigma^2 < +\infty$$

Since the acf is a constant, the spectral measure is $\nu(d\lambda) = b^2\sigma^2\delta(d\lambda)$.

2. $X_t = Z_0\cos(ct)$

$$\begin{aligned} \mathbb{E}\left[X_t\right] = 0 \qquad \mathsf{Cov}\left(X_t, X_{t+h}\right) &= \mathbb{E}\left[|Z_0|^2\cos(ct)\cos(ch+ct)\right] \\ &= \frac{\sigma^2}{2}\left[\cos(ch) + \cos(c(2t+h))\right] \end{aligned}$$

The covariance of X_t and X_{t+h} depends on t, thus the process is not s.o.2.

3. $X_t = a + bZ_t + cZ_{t-1}$

$$\mathbb{E}[X_t] = a \qquad \mathsf{Cov}(X_t, X_{t+h}) = \mathsf{Cov}(bZ_t + cZ_{t-1}, bZ_{t+h} + cZ_{t+h-1})$$
$$= (c^2 + b^2)\gamma_Z(h) + bc\gamma_Z(h-1) + bc\gamma_Z(h+1)$$
$$= (c^2 + b^2)\delta_h + bc\delta_{h-1} + bc\delta_{h+1}$$

Thus, $\mathsf{Cov}(X_t, X_{t+h})$ does not depend on t and $\mathsf{Var}(X_t) = \gamma_X(0) = c^2 + b^2 < +\infty$. Therefore, it is a w.s. process. Finally,

$$f(\lambda) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \gamma(h)e^{-ih\lambda} = \frac{1}{2\pi}\left(b^2 + c^2 + 2bc\cos\lambda\right)$$

4. $X_t = Z_1\cos(ct) + Z_2\sin(ct)$

$$\mathbb{E}[X_t] = 0 \qquad \mathsf{Cov}(X_t, X_{t+h}) = \mathsf{Cov}(Z_1\cos(ct) + Z_2\sin(ct), Z_1\cos(c(t+h)) + Z_2\sin(c(t+h)))$$
$$= \sigma^2[\cos(ct)\cos(c(t+h)) + \sin(ct)\sin(c(t+h))]$$
$$= \frac{1}{2}\sigma^2[\cos(2ct + 2ch) + \cos(ch) + \cos(ch) - \cos(2ct + 2ch)] = \sigma^2\cos(ch)$$

Thus, $\mathsf{Cov}(X_t, X_{t+h})$ does not depend on t and $\mathsf{Var}(X_t) = \sigma^2 < +\infty$. Therefore, it is a w.s. process. Finally,

$$\nu(d\lambda) = \frac{\sigma^2}{2}[\delta(d\lambda - c) + \delta(d\lambda + c)]$$

5. $X_t = Z_t\cos(ct) + Z_{t-1}\sin(ct) \Rightarrow \mathbb{E}[X_t] = 0$

$$\begin{aligned}\mathsf{Cov}(X_t, X_{t+h}) &= \mathsf{Cov}(Z_t\cos(ct) + Z_{t-1}\sin(ct), Z_{t+h}\cos(c(t+h)) + Z_{t+h-1}\sin(c(t+h)))\\ &= \sigma^2[\delta_h\cos(ct)\cos(c(t+h)) + \delta_{h-1}\cos(ct)\sin(c(t+h)) + \delta_{h+1}\sin(ct)\cos(c(t+h)) +\\ &+ \delta_h\sin(ct)\sin(c(t+h))]\\ &= \sigma^2\left[\delta_h\cos(ch) + \delta_{h-1}\frac{1}{2}(\sin(c(2t+h)) + \sin(ch)) + \delta_{h+1}\frac{1}{2}(\sin(c(2t+h)) - \sin(ch))\right]\end{aligned}$$

Thus, $\mathsf{Cov}(X_t, X_{t+h})$ depends on t, the process is not s.o.2.

Solution of Exercise 4.2 Let us define the sequence $d : k \in \mathbb{N}_0 \to \mathsf{det}(\Gamma_{k+1})$. We have the following:

$$k = 0 \qquad \Gamma_1 = [1] \qquad d_0 = \mathsf{det}(\Gamma_1) = 1$$

$$k = 1 \qquad \Gamma_2 = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad d_1 = \mathsf{det}(\Gamma_2) = 1 - \rho^2$$

For $k \geq 2$, we can write Γ_{k+1} as a block matrix:

$$\Gamma_{k+1} = \left[\begin{array}{c|cccccc} 1 & \rho & 0 & 0 & \dots & 0 & 0 \\ \hline \rho & & & & & & \\ 0 & & & & \Gamma_k & & \\ \dots & & & & & & \\ 0 & & & & & & \end{array}\right] = \left[\begin{array}{c|c|cccccc} 1 & \rho & 0 & 0 & \dots & 0 & 0 \\ \hline \rho & 1 & \rho & 0 & \dots & 0 & 0 \\ \cline{2-8} 0 & \rho & & & & & \\ \dots & \dots & & & \Gamma_{k-1} & & \\ 0 & 0 & & & & & \end{array}\right]$$

Therefore we have:

$$d_k = \mathsf{det}(\Gamma_{k+1}) = \mathsf{det}(\Gamma_k) - \rho\mathsf{det}\left[\begin{array}{c|cccccc} \rho & \rho & 0 & 0 & \dots & 0 & 0 \\ \hline 0 & & & & & & \\ 0 & & & & \Gamma_{k-1} & & \\ \dots & & & & & & \\ 0 & & & & & & \end{array}\right] = d_{k-1} - \rho^2 d_{k-2}$$

Thus we have that the sequence d is the solution of the following recurrent equation:

$$\begin{cases} d_k = d_{k-1} - \rho^2 d_{k-2} \\ d_0 = 0 \\ d_1 = 1 - \rho^2 \end{cases} \tag{2}$$

The characteristic equation is $x^2 - x + \rho^2 = 0$, with solutions

$$x_0 = \frac{1 - \sqrt{1 - 4\rho^2}}{2} \qquad\qquad x_1 = \frac{1 + \sqrt{1 - 4\rho^2}}{2}$$

Therefore, the sequence d_k has the following form:

$$d_k = \begin{cases} \alpha x_0^k + \beta x_1^k & \text{if } x_0 \neq x_1 \Leftrightarrow |\rho| \neq \frac{1}{2} \\ (\alpha + \beta k) x_0^k & \text{if } x_0 = x_1 \Leftrightarrow |\rho| = \frac{1}{2} \Rightarrow x_0 = x_1 = \frac{1}{2} \end{cases}$$

where α and β are defined by the initial conditions.

2. We have now to show that the matrices Γ_k are positive definite given some condition on ρ. Using the expression Eq. (2) for the sequence of determinants, we have to find under which conditions on ρ, the determinants are all positive: $d_k > 0 \forall k \in \mathbb{N}_0$.

We have to consider three cases, with respect to the discriminant of the characteristic equation $x^2 - x + \rho^2 = 0$: positive, null and negative discriminant. Since $\Delta = 1 - 4\rho^2$, these conditions correspond respectively to $|\rho| < \frac{1}{2}$, $|\rho| = \frac{1}{2}$, and $|\rho| > \frac{1}{2}$.

If $\rho = |1/2|$, by applying the initial condition, one can easily find that $\alpha = 1$ and $\beta = 1/2$. In that case $d_k = (1 + \frac{k}{2}) \left(\frac{1}{2}\right)^k > 0 \forall k$. Then the Γ_k matrices are all definite positive, thus they can be autocovariance matrices.

If $|\rho| \neq \frac{1}{2}$, one can find that $\alpha = \frac{\rho^2 - x_0}{\sqrt{\Delta}} = \frac{1}{2} - \sqrt{\Delta}\left(\frac{1}{2} + \frac{\rho^2}{\Delta}\right)$ and $\beta = \frac{x_1 - \rho^2}{\sqrt{\Delta}} = \frac{1}{2} + \sqrt{\Delta}\left(\frac{1}{2} + \frac{\rho^2}{\Delta}\right)$. Now, if $|\rho| < \frac{1}{2}$ then $\Delta > 0$ and both α and β are real. It can also be proven that $\beta > 1$, $\alpha < 0$ and $|\beta| - |\alpha > |1$. Since $0 < x_0 < x_1$, $|\beta||x_1|^n > |\alpha||x_0|^n$, proving that $\forall k \in \mathbb{N}_0, d_k > 0$, *q.d.e.*.

Finally, if $|\rho| > \frac{1}{2}$, it can be shown that d_k has sinusoidal terms, hence it can be negative, which prevents Γ_k from being an autocovariance matrix.

As alternative method, we can use the **Herglotz theorem**, stating that $\gamma(h)$ is positive if and only if it exists a positive measure ν such that $\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} \nu(d\lambda)$. Here we can use the density: $\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda$ where

$$f(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma(h) e^{-ih\lambda} \frac{1}{2\pi} (1 + 2\rho \cos \lambda)$$

The density is non-negative for all λ if and only if $|\rho| leq \frac{1}{2}$, *q.d.e.*.

3. Let us consider a weak white noise $\{\epsilon_t, t \in \mathbb{Z}\}$ and a process $\{X_t = a\epsilon_t + b\epsilon_{t-1}, t \in \mathbb{Z}\}$, with $a, b \in \mathbb{R}$. Then, the new process is real-valued and centered: $\mathbb{E}[X_t] = 0$. Moreover,

$$\begin{aligned} \mathsf{Cov}(X_t, X_{t+h}) = \mathbb{E}[X_t X_{t+h}] &= \mathbb{E}\left[a^2 \epsilon_t \epsilon_{t+h} + b^2 \epsilon_{t-1} \epsilon_{t-1+h} + ab \epsilon_{t+h} \epsilon_{t-1} + ab \epsilon_t \epsilon_{t-1+h}\right] \\ &= \left(a^2 + b^2\right) \delta_h + ab \left(\delta_{h-1} + \delta_{h+1}\right) \end{aligned}$$

Finally, we find a and b by setting:

$$\left(a^2 + b^2\right) = 1$$
$$ab = \rho$$

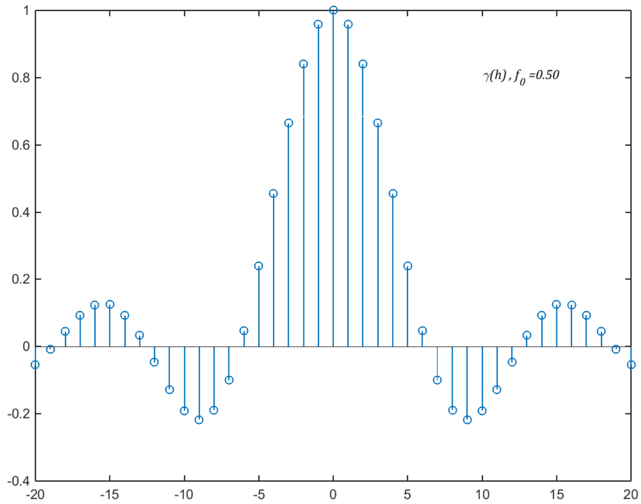

Figure 4: Example of autocovariance function for a band-limited stationary process, Exercise 4.3.

implying $\left(a^2+b^2\right)+2ab = 1+2\rho$ and thus $a+b=\sqrt{1+2\rho}$. Then we have:

$$\begin{aligned}
b &= \sqrt{1+2\rho}-a \\
b^2 &= a^2+1+2\rho-2a\sqrt{1+2\rho} \\
a^2+b^2 &= 2a^2+1+2\rho-2a\sqrt{1+2\rho} \\
1 &= 2a^2+1+2\rho-2a\sqrt{1+2\rho} \\
2a^2+2\rho-2a\sqrt{1+2\rho} &= 0 \\
a &= \frac{\sqrt{1+2\rho}\pm\sqrt{1-2\rho}}{2} \\
b &= \frac{\sqrt{1+2\rho}\mp\sqrt{1-2\rho}}{2}
\end{aligned}$$

Note that, since $|\rho| \le \frac{1}{2}$, $a, b \in \mathbb{R}$.

Solution of Exercise 4.3

$$\begin{aligned}
\gamma(h) &= \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda \\
&= \int_{-f_0}^{f_0} e^{ih\lambda} d\lambda \\
&= \frac{1}{ih}\left(e^{ihf_0}-e^{-ihf_0}\right) \\
&= 2\frac{\sin(hf_0)}{h} = 2f_0 \mathsf{Sinc}(f_0 h)
\end{aligned}$$

An example of this function is given in Fig. 4. It is not L^1 since in that case its density would have been continuous.

Solution of Exercise 4.4 1. Let $W = \{\ell \in \mathbb{Z}^+ | (X_1,\ldots,X_\ell)$ is a set of linearly independent vectors$\}$. If this set is empty, this means that even (X_1) is not a set of linearly independent vectors, thus $\exists a \in \mathbb{R}^+$ such that $\mathsf{Var}\,(aX_1)=0$. Since $a \neq 0$, $\gamma(0)=\mathsf{Var}\,(X_1)=0/a=0$.

If W is not empty, we define k as the maximum value in W. Since the elements of W are drawn from $\mathbb{Z}^+$, we have $k \geq 1$. Then, by our choice of k, $(X_1, \ldots, X_{k+1})$ is a not set of linearly independent vectors, while $(X_1, \ldots, X_k)$ is such. This imply $X_{k+1} \in \mathsf{Vect}(X_1, \ldots, X_k)$.

2. Since the autocovariance matrix is invertible, its smallest eigenvalue is positive

3. We have to show that, $\forall p \geq 1, X_{k+p} \in \mathsf{Vect}(X_1, \ldots, X_k)$. If $\gamma(0) = 0$ this is trivial. Otherwise, we will prove it by recurrence.

3.1. The basis of the recurrence is already proved: $X_{k+1} \in \mathsf{Vect}(X_1, \ldots, X_k)$

3.2. We have to prove that, if $\forall \ell < p, X_{k+\ell} \in \mathsf{Vect}(X_1, \ldots, X_k)$, then, also $X_{k+p} \in \mathsf{Vect}(X_1, \ldots, X_k)$.

By stationarity, $X_{k+1} \in \mathsf{Vect}(X_1, \ldots, X_k) \Rightarrow X_{k+p} \in \mathsf{Vect}(X_p, \ldots, X_{p+k-1})$.

By recurrence hypothesis, each of $(X_p, \ldots, X_{p+k-1})$ is in $\mathsf{Vect}(X_1, \ldots, X_k)$. Therefore, the same for X_{k+p}, *q.e.d.*

4. We rewrite Eq.(1) as $X_{k+p} = \varphi_p^T \mathbf{X} = \mathbf{X}^T \varphi_p$, where φ_p is the vector of the scalars $\phi_{p,1}, \ldots, \phi_{p,k}$ and $\mathbf{X}$ is the random vector $[X_1, \ldots, X_k]^T$. We have

$$\gamma(0) = \mathbb{E}\left[|X_{k+p}|^2\right] = \mathbb{E}\left[\varphi_p^H \mathbf{X}\mathbf{X}^H \varphi_p\right] = \varphi_p^H \Gamma_k \varphi_p \geq \lambda_{\mathsf{min}} \|\varphi_p\|^2 \Leftrightarrow \|\varphi_p\|^2 \leq \frac{\gamma(0)}{\lambda_{\mathsf{min}}} < +\infty$$

5.

$$\begin{aligned}\gamma(0) = \mathsf{Cov}\left(X_{k+p}, X_{k+p}\right) &= \mathsf{Cov}\left(X_{k+p}, \sum_{\ell=1}^{k} \phi_{p,\ell} X_\ell\right) = \sum_{\ell=1}^{k} \mathsf{Cov}\left(X_{k+p}, \phi_{p,\ell} X_\ell\right) \\ &= \sum_{\ell=1}^{k} \phi_{p,\ell}\gamma(p+k-\ell) \leq \sum_{\ell=1}^{k} \sqrt{\frac{\gamma(0)}{\lambda_{\mathsf{min}}}}\gamma(p+k-\ell)\end{aligned}$$

By passing to the limit for $p \to +\infty$, we obtain $\gamma(0)$ for the left-hand term and 0 for the right-hand term.

Solution of Exercise 5.1 We know that, $\forall t, k \in \mathbb{Z}, S_{t+4k} = S_t$

1. $\mathbb{E}[Y_t] = \mathbb{E}[\beta t + S_t + X_t] = \beta t + \mu_S + \mu_X$. Therefore $\{Y_t, t \in \mathbb{Z}\}$ is not w.s. unless $\beta = 0$.

2.1.

$$\forall k \in \mathbb{Z}, \qquad \gamma_S(h) = \mathsf{Cov}\left(S_t, S_{t+h}\right) = \mathsf{Cov}\left(S_t, S_{t+h+4k}\right) = \gamma_S(h+4k)$$

Therefore γ_S is periodic with period equal to 4.

2.2. By applying the operator $\left(1 + B + B^2 + B^3\right)$ on S, we obtain:

$$\begin{aligned}\forall t \in \mathbb{Z}, \qquad & \bar{S}_t = S_t + S_{t-1} + S_{t-2} + S_{t-3} \qquad \Rightarrow \\ \forall t \in \mathbb{Z}, \qquad & \bar{S}_t - \bar{S}_{t-1} = S_t - S_{t-4} = 0 \qquad \Rightarrow \\ \forall t \in \mathbb{Z}, \qquad & \bar{S}_t = \bar{S}_0 = S_0 + S_1 + S_2 + S_3\end{aligned}$$

3. First, we observe that, given a process $\{W_t, t \in \mathbb{Z}\}$, $(1-B) \circ \left(1 + B + B^2 + B^3\right) \circ W_t = (1 - B^4) \circ W_t$. Therefore,

$$Z_t = (1 - B^4) \circ (\beta t + S_t + X_t) = \beta t + S_t + X_t - \beta(t-4) - S_{t-4} - X_{t-4} = 4\beta + X_t - X_{t-4}$$

Then, $\mathbb{E}[Z_t] = 4\beta$ and:

$$\mathsf{Cov}\left(Z_t, Z_{t+h}\right) = \mathsf{Cov}\left(X_t - X_{t-4}, X_{t+h} - X_{t+h-4}\right) = 2\gamma_X(h) - \gamma_X(h-4) - \gamma_X(h+4)$$

Therefore $\{Z_t, t \in \mathbb{Z}\}$ is w.s. and $\gamma_Z(h) = 2\gamma_X(h) - \gamma_X(h-4) - \gamma_X(h+4)$.

4. As an autocovariance function, γ_S is Hermitian, but since $\{S_t, t \in \mathbb{Z}\}$ is real, it is symmetric: $\gamma_S(-h) = \gamma_S(h)$. Moreover, we have shown that γ_S is periodic, thus defined by the values of its period. We set:

$$\begin{aligned}&\gamma_S(0) = \gamma_0 \\ &\gamma_S(1) = \gamma_1 \\ &\gamma_S(2) = \gamma_2 \\ &\gamma_S(3) = \gamma_S(-1) = \gamma_S(1) = \gamma_1\end{aligned}$$

Thus γ_S has three degrees of freedom. Let us now show that a function

$$\eta(h) = a + b\cos\left(\frac{\pi}{2}h\right) + c\cos(\pi h)$$

satisfies all the constraint of γ_S. First we observe that η is real, periodical of period 4 and symmetric. Moreover,

$$\begin{aligned} \eta(0) &= a + b + c \\ \eta(1) &= a - c \\ \eta(2) &= a - b + c \end{aligned}$$

Finally, the parameters a, b, c are found by solving

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} \Rightarrow \begin{array}{l} a = \frac{\gamma_0}{4} + \frac{\gamma_1}{2} + \frac{\gamma_2}{4} \\ b = \frac{\gamma_0}{2} - \frac{\gamma_2}{2} \\ c = \frac{\gamma_0}{4} - \frac{\gamma_1}{2} + \frac{\gamma_2}{4} \end{array}$$

As for the spectral measure, from $\gamma_S(h) = a + b\cos\left(\frac{\pi}{2}h\right) + c\cos(\pi h)$, we have that $\nu_S(d\lambda) = a\delta_0(d\lambda) + \frac{b}{2}\delta_{\frac{\pi}{2}}(d\lambda) + \frac{b}{2}\delta_{-\frac{\pi}{2}}(d\lambda) + c\delta_\pi(d\lambda)$.

5.

$$\begin{aligned} \gamma_Z(h) &= 2\gamma_X(h) - \gamma_X(h-4) - \gamma_X(h+4) \Rightarrow \\ f_Z(\lambda) &= 2f_X(\lambda) - f_X(\lambda)e^{-i4\lambda} - f_X(\lambda)e^{i4\lambda} \\ &= 2f_X(\lambda)\left(1 - \frac{e^{i4\lambda} + e^{-i4\lambda}}{2}\right) = 2f_X(\lambda)\left[1 - \cos(4\lambda)\right] = 4f_X(\lambda)\sin^2(2\lambda) \end{aligned}$$

Solution of Exercise 5.2 1. For a centered w.s. process $\{X_t, t \in \mathbb{Z}\}$, the innovation process is defined at each t as the difference between X_t and its projection on the *linear past* of the process. Thus, $\{Z_t, t \in \mathbb{Z}\}$ is the innovation process of $\{X_t, t \in \mathbb{Z}\}$, and as such, it is a white noise (Corollary 2.4.1 in the text book). Let us prove that in this special case.

It is easy to see that $\mathbb{E}[Z_t] = 0$. We also have that $Z_t \in \mathcal{H}_t$, since both X_t and $\widetilde{X}_t$ are in $\mathcal{H}_t$.

$$\mathsf{Proj}(Z_t|\mathcal{H}_{t-1}) = \mathsf{Proj}(X_t - \widetilde{X}_t|\mathcal{H}_{t-1}) = \widetilde{X}_t - \widetilde{X}_t = 0 \Rightarrow Z_t \perp \mathcal{H}_{t-1} \Rightarrow Z_t \perp \widetilde{X}_t \Rightarrow$$

$$\mathbb{E}\left[|X_t|^2\right] = \mathbb{E}\left[|Z_t|^2 + |\widetilde{X}_t|^2\right] = \mathbb{E}\left[|Z_t|^2\right] + \mathbb{E}\left[|\widetilde{X}_t|^2\right] \Rightarrow \mathbb{E}\left[|Z_t|^2\right] = \mathbb{E}\left[|X_t|^2\right] - \mathbb{E}\left[|\widetilde{X}_t|^2\right] \quad (3)$$

$$\forall s < t, Z_s \in \mathcal{H}_s \subseteq \mathcal{H}_{t-1} \Rightarrow Z_t \perp Z_s \Leftrightarrow \mathsf{Cov}(Z_t, Z_s) = 0 \quad (4)$$

Eq. (3) shows that $\mathsf{Cov}(Z_t, Z_t)$ does not depend on t and Eq. (3) shows that $\mathsf{Cov}(Z_t, Z_{t+h})$ does not depend on t neither, and is null. Therefore, $\{Z_t, t \in \mathbb{Z}\}$ is a weak white noise.

2. $\forall s \leq t - q - 1, \mathsf{Cov}(X_t, X_s) = \gamma_X(t-s) = 0$ since $t - s > q$. This means that $\forall s \leq t - q - 1, X_t \perp X_s$, q.e.d.

3. We know that $X_t \perp \mathcal{H}_{t-q-1}$ and $X_t \in \mathcal{H}_t$, *i.e.*, X_t is in the orthogonal complement of $\mathcal{H}_{t-q-1}$ in $\mathcal{H}_t$, which is a space with $q+1$ dimensions. In this space, the set $(Z_s, s \in \{t, t-1, \ldots, t-q\})$ is made up of orthogonal vectors, so it is a basis, implying $X_t \in \mathsf{Vect}(Z_s, s \in \{t, t-1, \ldots, t-q\})$.

4. From the previous, we can write $X_t = \sum_{p=0}^q \theta_{t,p} Z_{t-p}$. The coefficients of the projection on the orthogonal basis are found as:

$$\begin{aligned} \theta_{t,p} &= \mathsf{Cov}(X_t, Z_{t-p}) = \mathsf{Cov}\left(X_t, X_{t-p} - \widetilde{X}_{t-p}\right) \\ &= \gamma_X(p) - \mathsf{Cov}\left(X_t, \widetilde{X}_{t-p}\right) \end{aligned}$$

By stationarity, $\mathsf{Cov}\left(X_t, \widetilde{X}_{t-p}\right)$ does not depend on t, thus $\theta_{t,p}$ also only depends on p, and can be referred to as θ_p. In conclusion, we can write:

$$\forall t \in \mathbb{Z}\; X_t = \sum_{p=0}^q \theta_p Z_{t-p},$$

with $\{Z_t, t \in \mathbb{Z}\}$ a white noise: this is the definition of MA(q) process.

Solution of Exercise 5.3 1. Let us compute the average and the covariance for the sum of the MA processes:

$$\begin{aligned}\mathbb{E}[Z_t] &= \mathbb{E}[X_t] + \mathbb{E}[Y_t] = 0 \\ \mathsf{Cov}(Z_{t+h}, Z_t) &= \mathsf{Cov}(X_{t+h} + Y_{t+h}, X_t + Y_t) = \gamma_X(h) + \gamma_(h)\end{aligned}$$

Thus, $\{Z_t, t \in \mathbb{Z}\}$ is a w.s. process. Moreover, since $\gamma_Z(h) = \gamma_X(h) + \gamma_(h)$, the support of $\gamma_Z(h)$ is $s = \max\{p, q\}$. As shown in Exercise 5.2, this implies that $\{Z_t, t \in \mathbb{Z}\}$ is an MA(s) process.

2. Let us use the shortcuts $\theta = \theta_1$ and $\rho = \rho_1$. The process X can be seen as the filtering of the WN ϵ with an FIR filter with impulse response $a : n \in \mathbb{Z} \to \delta_n + \theta\delta_{n-1}$. This means that ϵ can be recovered from X by applying the inverse filter with impulse response

$$b : n \in \mathbb{Z} \to \begin{cases} (-\theta)^n & \text{if } n \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Similarly, we can recover η from Y. We have

$$\epsilon_t = \sum_{k=0}^{+\infty} (-\theta)^k X_{t-k} \qquad \eta_t = \sum_{k=0}^{+\infty} (-\rho)^k Y_{t-k}$$

$$\mathbb{E}[\epsilon_t, \eta_s] = \mathbb{E}\left[\sum_{k=0}^{+\infty} (-\theta)^k X_{t-k} \sum_{\ell=0}^{+\infty} (-\rho)^\ell Y_{s-\ell}\right] = \sum_{k=0}^{+\infty}\sum_{\ell=0}^{+\infty} (-\theta)^k (-\rho)^\ell \mathbb{E}[X_{t-k} Y_{s-\ell}] = 0 \text{ q.e.d.}$$

3. In this case, introducing $\xi_t = \epsilon_t + \eta_t$, we have $Z_t = \epsilon_t + \eta_t + \theta(\epsilon_{t-1} + \eta_{t-1}) = \xi_t + \theta\xi_{t-1}$. Since $|\theta| < 1$, we know that this is a canonical MA representation and thus ξ is the innovation process.

4. In this case we have:

$$\begin{aligned} X_t &= \epsilon_t + \theta\epsilon_{t-1} & Y_t &= \eta_t + \rho\eta_{t-1} \quad \Rightarrow \\ \gamma_X(h) &= \sigma_\epsilon^2\left[(1+\theta^2)\delta_h + \theta\delta_{h-1} + \theta\delta_{h+1}\right] & \gamma_Y(h) &= \sigma_\eta^2\left[(1+\rho^2)\delta_h + \rho\delta_{h-1} + \rho\delta_{h+1}\right] \end{aligned}$$

In Question 1 we have shown that Z must be MA(1). This means that it must exist a WN ϕ and a real number α such that ϕ is the innovation of Z and

$$\begin{aligned} Z_t &= \phi_t + \alpha\phi_{t-1} \\ \gamma_Z(h) &= \sigma_\phi^2\left[(1+\alpha^2)\delta_h + \delta_{h-1} + \delta_{h+1}\right] \end{aligned}$$

The unknown α and σ_ϕ^2 can be found by the identity $\gamma_Z(h) = \gamma_X(h) + \gamma_Y(h)$:

$$\sigma_\phi^2\left[(1+\alpha^2)\delta_h + \delta_{h-1} + \delta_{h+1}\right] = \sigma_\epsilon^2\left[(1+\theta^2)\delta_h + \theta\delta_{h-1} + \theta\delta_{h+1}\right] + \sigma_\eta^2\left[(1+\rho^2)\delta_h + \rho\delta_{h-1} + \rho\delta_{h+1}\right]$$

$$\begin{cases} \sigma_\phi^2(1+\alpha^2) = \sigma_\epsilon^2(1+\theta^2) + \sigma_\eta^2(1+\rho^2) \\ \sigma_\phi^2\alpha = \sigma_\epsilon^2\theta + \sigma_\eta^2\rho \end{cases}$$

Let us first set $a = \sigma_\epsilon^2(1+\theta^2) + \sigma_\eta^2(1+\rho^2)$ and $b = \sigma_\epsilon^2\theta + \sigma_\eta^2\rho$. We find that $\alpha = \frac{b}{\sigma_\phi^2}$ and then:

$$\sigma_\phi^2\left(1 + \frac{b^2}{\sigma_\phi^4}\right) = a \qquad \sigma_\phi^2 + \frac{b^2}{\sigma_\phi^2} - a = 0$$

$$\sigma_\phi^4 - a\sigma_\phi^2 + b^2 = 0 \qquad \sigma_\phi^2 = \frac{1}{2}\left(a \pm \sqrt{a^2 - 4b^2}\right)$$

$$\sigma_\phi^2 = \frac{1}{2}\left[\sigma_\epsilon^2(1+\theta^2) + \sigma_\eta^2(1+\rho^2) \pm \sqrt{\sigma_\epsilon^4(1-\theta^2)^2 + \sigma_\eta^4(1-\rho^2)^2 + 2\sigma_\epsilon^2\sigma_\eta^2(1+\theta^2)(1+\rho^2)}\right]$$

Solution of Exercise 5.4 Let us observe that $\epsilon_t = X_t - aX_{t-1}$ and $\eta_t = Y_t - bY_{t-1}$. We can write the following:

$$\begin{aligned} Z_t - (a+b)Z_{t-1} + abZ_{t-2} &= X_t + Y_t - aX_{t-1} - aY_{t-1} - bX_{t-1} - bY_{t-1} + abX_{t-2} + abY_{t-2} \\ &= X_t - aX_{t-1} - b(X_{t-1} - bX_{t-2}) + Y_t - bY_{t-1} - a(Y_{t-1} - bY_{t-2}) \\ &= \epsilon_t - b\epsilon_{t-1} + \eta_t - a\eta_{t-1} = W_t + V_t \end{aligned}$$

Now, both $\{W_t = \epsilon_t - b\epsilon_{t-1}, t \in \mathbb{Z}\}$ and $\{V_t = \eta_t - a\eta_{t-1}, t \in \mathbb{Z}\}$ are MA(1) processes, and thus their sum is also a MA(1) process, meaning that it exists a WN ξ and a real number $\theta \in]-1, 1[$ such that $Z_t - (a+b)Z_{t-1} + abZ_{t-2} = \xi_t - \theta\xi_{t-1}$, *q.e.d.*

2. From the previous point, we can write

$$\xi_t - \theta\xi_{t-1} = \epsilon_t - b\epsilon_{t-1} + \eta_t - a\eta_{t-1} \tag{5}$$
$$(1 - \theta B) \circ \xi_t = (1 - bB) \circ \epsilon_t + (1 - aB) \circ \eta_t \tag{6}$$

where we use the back-shift operator B. The left-hand term of this equation can be read as the filtering of ξ with a FIR with impulse response $h_k = \delta_k - \theta\delta_{k-1}$. As shown in Exercise 5.3, this filter can be inversed by applying a filter with impulse response

$$g : n \in \mathbb{Z} \to \begin{cases} (\theta)^n & \text{if } n \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Apply the inverse filter to both members of Eq. (6), we get:

$$\begin{aligned} \xi_t &= (1 - bB) \sum_{n\geq 0} \theta^n \epsilon_{t-n} + (1 - aB) \sum_{n\geq 0} \theta^n \eta_{t-n} \\ &= (1 - bB) \left(\epsilon_t + \sum_{n\geq 1} \theta^n \epsilon_{t-n} \right) + (1 - aB) \left(\eta_t + \sum_{n\geq 1} \theta^n \eta_{t-n} \right) \\ &= \epsilon_t + \sum_{n\geq 1} \theta^n \epsilon_{t-n} - b \sum_{n\geq 0} \theta^n \epsilon_{t-1-n} + \eta_t + \sum_{n\geq 1} \theta^n \eta_{t-n} - a \sum_{n\geq 0} \theta^n \eta_{t-1-n} \\ &= \epsilon_t + \sum_{m\geq 0} \theta^{m+1} \epsilon_{t-1-m} - b \sum_{n\geq 0} \theta^n \epsilon_{t-1-n} + \eta_t + \sum_{m\geq 0} \theta^{m+1} \eta_{t-1-m} - a \sum_{n\geq 0} \theta^n \eta_{t-1-n} \\ &= \epsilon_t + (\theta - b) \sum_{k\geq 0} \theta^k \epsilon_{t-1-k} + \eta_t + (\theta - a) \sum_{h\geq 0} \theta^h \eta_{t-1-h} \quad q.e.d. \end{aligned}$$

3. We write the following:

$$\begin{aligned} Z_{t+1} &= (a+b)Z_t - abZ_{t-1} + \xi_{t+1} - \theta\xi_t \\ &= (a+b)Z_t - abZ_{t-1} + \epsilon_{t+1} + (\theta - b) \sum_{k\geq 0} \theta^k \epsilon_{t-k} + \eta_{t+1} + (\theta - a) \sum_{h\geq =} \theta^h \eta_{t-h} - \theta\xi_t \\ &= [\epsilon_{t+1} + \eta_{t+1}] + \left[(a+b)Z_t - abZ_{t-1} + (\theta - b) \sum_{k\geq 0} \theta^k \epsilon_{t-k} + (\theta - a) \sum_{h\geq 0} \theta^h \eta_{t-h} - \theta\xi_t \right] \end{aligned} \tag{7}$$

If we know $(X_s \forall s \leq t)$ and $(Y_s \forall s \leq t)$, we also know Z_t, Z_{t-1}. Moreover, by applying an inverse filtering, we know also $\epsilon_{t-k} \forall k \geq 0$ and $\eta_{t-h} \forall h \geq 0$. On the contrary, we do not know ϵ_{t+1} nor η_{t+1}, and both are uncorrelated with $(X_s \forall s \leq t)$ and $(Y_s \forall s \leq t)$ Therefor the first term in the right-hand part of Eq. (7) is the innovation, while the second term is the prediction.

4. In this case we do not know separately $(X_s \forall s \leq t)$ and $(Y_s \forall s \leq t)$, but only their sum. We write therefore:

$$\begin{aligned} Z_{t+1} &= (a+b)Z_t - abZ_{t-1} + \xi_{t+1} - \theta\xi_t \\ &= \xi_{t+1} + (a+b)Z_t - abZ_{t-1} - \theta\xi_t \\ &= \xi_{t+1} + \widetilde{Z}_t \end{aligned}$$

Thus ξ_{t+1} is the innovation and $\widetilde{Z}_t = (a+b)Z_t - abZ_{t-1} - \theta\xi_t$ is the prediction. Again, ξ_t is obtained by inverse filtering of $Z_t - (a+b)Z_{t-1} + abZ_{t-2}$.

5. In the first case,

$$\mathbb{E}\left[|\eta_{t+1} + \epsilon_{t+1}|^2\right] = \sigma_\eta^2 + \sigma_\epsilon^2.$$

In the second we have:

$$\begin{aligned}\xi_t &= \epsilon_t + (\theta - b)\sum_{k\geq 0}\theta^k\epsilon_{t-1-k} + \eta_t + (\theta - a)\sum_{h\geq 0}\theta^h\eta_{t-1-h}\\ &= \epsilon_t + (\theta - b)\alpha_t + \eta_t + (\theta - a)\beta_t\end{aligned}$$

with:

$$\alpha_t = \sum_{k\geq 0}\theta^k\epsilon_{t-1-k} \qquad\qquad \beta_t = \sum_{h\geq 0}\theta^h\eta_{t-1-h}$$

Therefore ξ is expressed as the sum of four uncorrelated processes. We can then compute its variance, referred to as σ^2, as the sum of the four variances:

$$\sigma^2 = \mathsf{Var}\left(\xi_t\right) = \sigma_\epsilon^2 + (\theta-b)^2\mathsf{Var}\left(\alpha_t\right) + \sigma_\eta^2 + (\theta-a)^2\mathsf{Var}\left(\beta_t\right)$$

We have:

$$\begin{aligned}\mathsf{Var}\left(\alpha_t\right) &= \mathbb{E}\left[\sum_{k\geq 0}\theta^k\epsilon_{t-1-k}\sum_{\ell\geq 0}\theta^\ell\epsilon_{t-1-\ell}\right] = \sum_{k\geq 0}\sum_{\ell\geq 0}\theta^k\theta^\ell\gamma_\epsilon(k-\ell)\\ &= \sum_{k\geq 0}\sum_{\ell\geq 0}\theta^k\theta^\ell\sigma_\epsilon^2\delta_{k-\ell} = \sigma_\epsilon^2\sum_{k\geq 0}\theta^2 k = \frac{\sigma_\epsilon^2}{1-\theta^2}\end{aligned}$$

and, likewise, $\mathsf{Var}\left(\beta_t\right) = \frac{\sigma_\eta^2}{1-\theta^2}$. In conclusion,

$$\begin{aligned}\sigma^2 = \mathsf{Var}\left(\xi_t\right) &= \sigma_\epsilon^2 + (\theta-b)^2\frac{\sigma_\epsilon^2}{1-\theta^2} + \sigma_\eta^2 + (\theta-a)^2\frac{\sigma_\eta^2}{1-\theta^2}\\ &= \sigma_\epsilon^2\left[1 + \frac{(\theta-b)^2}{1-\theta^2}\right] + \sigma_\eta^2\left[1 + \frac{(\theta-a)^2}{1-\theta^2}\right]\end{aligned}$$

Thus we see that the variance of the innovation in the second case is always larger than that of the first case, unless $\theta = a = b$.

Solution of Exercise 5.5 We observe that $\{X_t, t\in\mathbb{Z}\}$ is a MA(1) process, thus, if γ be the autocovariance function of $\{X_t, t\in\mathbb{Z}\}$, its support is $\{-1, 0, +1\}$. In facts, we have:

$$\gamma(h) = \mathbb{E}\left[(Z_t + \theta Z_{t-1})(Z_{t+h} + \theta Z_{t+h-1})\right] = \sigma^2\left[(1+\theta^2)\delta_h + \theta\delta_{h-1} + \theta\delta_{h+1}\right]$$

1. The linear prediction of X_3 is written as:

$$\widehat{X}_3 = \alpha X_1 + \beta X_2.$$

Our problem consists in minimizing the mean square error $\mathbb{E}\left[\left(X_3 - \widehat{X}_3\right)^2\right]$. The optimal solution is found the the error $(X_3 - \widehat{X}_3)$ is orthogonal to data (X_1, X_2). Thus we have:

$$\begin{aligned}\mathsf{Cov}\left(X_3 - \widehat{X}_3, X_1\right) &= 0 & \mathsf{Cov}\left(X_3 - \widehat{X}_3, X_2\right) &= 0\\ \mathsf{Cov}\left(X_3 - \alpha X_1 - \beta X_2, X_1\right) &= 0 & \mathsf{Cov}\left(X_3 - \alpha X_1 - \beta X_2, X_2\right) &= 0\\ \gamma(2) - \alpha\gamma(0) - \beta\gamma(1) &= 0 & \gamma(-1) - \alpha\gamma(1) - \beta\gamma(0) &= 0\\ -\alpha\sigma^2(1+\theta^2) - \beta\sigma^2\theta &= 0 & \sigma^2\theta - \alpha\sigma^2\theta - \beta\sigma^2(1+\theta^2) &= 0\end{aligned}$$

This is a linear system, and we can actually get rid of σ^2:

$$\begin{bmatrix} (1+\theta^2) & \theta \\ \theta & (1+\theta^2) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ \theta \end{bmatrix}$$

We find:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\theta^4+\theta^2+1} \begin{bmatrix} (1+\theta^2) & -\theta \\ -\theta & (1+\theta^2) \end{bmatrix} \begin{bmatrix} 0 \\ \theta \end{bmatrix} = \begin{bmatrix} \frac{-\theta^2}{\theta^4+\theta^2+1} \\ \frac{\theta+\theta^3}{\theta^4+\theta^2+1} \end{bmatrix}$$

2. If we now set

$$\widehat{X}_3 = \alpha X_5 + \beta X_4.$$

and we look for α, β minimizing the MSE, we end up exactly with the same equation as before, since for real processes, $\gamma(h) = \gamma(-h)$. Therefore, the same optimal values of the coefficients are found:

$$\alpha = \frac{-\theta^2}{\theta^4+\theta^2+1} \qquad\qquad \beta = \frac{\theta+\theta^3}{\theta^4+\theta^2+1}$$

3. Let us define the spaces $V_1 = \mathsf{Vect}(X_1, X_2)$ and $V_2 = \mathsf{Vect}(X_4, X_5)$. Any element of V_1 is uncorrelated to any element of V_2 (*i.e.*, they are orthogonal):

$$\begin{aligned} &\mathsf{Cov}\left(aX_1 + bX_2, cX_4 + dX_5\right) \\ =&ac\mathsf{Cov}\left(X_1, X_4\right) + ad\mathsf{Cov}\left(X_1, X_5\right) + bc\mathsf{Cov}\left(X_2, X_4\right) + bd\mathsf{Cov}\left(X_2, X_5\right) \\ =&ac\gamma(-3) + ad\gamma(-4) + bc\gamma(-2) + bd\gamma(-3) = 0 \end{aligned}$$

Thus, $\mathsf{Vect}(X_1, X_2, X_4, X_5) = V_1 \oplus V_2$, which implies that

$$\widehat{X}_3 = \mathsf{Proj}(X_3 | V_1 \oplus V_2) = \mathsf{Proj}(X_3|V_1) + \mathsf{Proj}(X_3|V_2) = \widehat{X}_{3,1} + \widehat{X}_{3,2}$$

Since $\widehat{X}_{3,1}$ and $\widehat{X}_{3,2}$ are orthogonal, when we impose $\mathsf{Cov}\left(X_3 - \widehat{X}_3, X_i\right) = 0$, with $i \in \{1,2,4,5\}$, only one between $\widehat{X}_{3,1}$ and $\widehat{X}_{3,2}$ gives a non-zero covariance (depending on i). Therefore, we end up with $\mathsf{Cov}\left(X_3 - \widehat{X}_{3,1}, X_i\right) = 0$ or $\mathsf{Cov}\left(X_3 - \widehat{X}_{3,2}, X_i\right) = 0$, *i.e.*, the same equations as in Questions 1 and 2. Therefore we find the same partial solutions. In conclusion:

$$\begin{aligned} \widehat{X}_3 &= \frac{-\theta^2}{\theta^4+\theta^2+1} X_1 + \frac{\theta+\theta^3}{\theta^4+\theta^2+1} X_2 + \frac{\theta+\theta^3}{\theta^4+\theta^2+1} X_4 + \frac{-\theta^2}{\theta^4+\theta^2+1} X_5 \\ &= \frac{\theta+\theta^3}{\theta^4+\theta^2+1}(X_2 + X_4) - \frac{\theta^2}{\theta^4+\theta^2+1}(X_1 + X_5) \end{aligned}$$

Solution of Exercise 1 1. Let us first rewrite the equation defining X as an ARMA(p, q) equation:

$$X_t - \sum_{k=1}^{p} \phi_k X_{t-k} = \epsilon_t + \sum_{k=1}^{p} \theta_k \epsilon_{t-k} \tag{8}$$

Let us introduce the polynomials $\Phi(z)$, $\Theta(z)$:

$$\Phi(z) = 1 - \sum_{k=1}^{p} \phi_k z^k \qquad\qquad \Theta(z) = 1 + \sum_{k=1}^{p} \theta_k z^k$$

Introducing the backshift operator B, the ARMA equation (Eq. (14)) can be written as:

$$\Phi(B)X = \Theta(B)\epsilon \tag{9}$$

Now we have just to check that a) $\Phi(z)$ and $\Theta(z)$ do not have common roots and that b) $\Phi(z)$ does not vanish on the unit circle of $\mathbb{C}$. This is straighforward since the only root of Φ is $1/2$ while the only root of Θ is $-1/4$. We can then apply theorem 3.3.2: X is the unique w.s. solution of Eq. (14), and it admits a spectral density function given by:

$$f_Z(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-i\lambda})\right|^2}{\left|\Phi(e^{-i\lambda})\right|^2}$$

In our case we have the following function, shown in Fig. 1:

$$f_Z(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|1 + 4e^{-i\lambda}\right|^2}{\left|1 - 2e^{-i\lambda}\right|^2} = \frac{\sigma^2}{2\pi} \frac{8\cos\lambda + 17}{5 - 4\cos\lambda}.$$

2. We remind that a canonical representation of an ARMA process is characterized by the fact that X is a causal and invertible filtering of weak noise. This is equivalent to say that neither Φ nor Θ vanish on the closed unit disk $\Delta_1 = \{z \in \mathbb{C} : |z| \leq 1\}$.

A given representation of an ARMA process is not necessarily canonical but it is possible to get a canonical representation by using an *all-pass filter*. We recall that, given $\psi \in \ell^1$, the filter F_ψ is an all-pass filter if and only if:

$$\forall z \in \Gamma_1, \left| \sum_{k \in \mathbb{Z}} \psi_k z^k \right| = c,$$

where $\Gamma_1 = \{\{z \in \mathbb{C} : |z| = 1\}$ is the complex unit circle and $c > 0$ is a constant.

A key property of all-pass filters is that they transform a WN process A_t into another WN process B_t. To prove this, let us first recall that, since $\psi \in \ell^1$, then theorem 3.1.2 and corollary 3.1.3 apply. Thus $B = F_\psi(A)$ is a w.s. centered process, with spectral density function

$$f_B(\lambda) = \frac{\sigma_A^2}{2\pi} \left| \sum_{k \in \mathbb{Z}} \psi_k e^{-ik\lambda} \right|^2 = \frac{\sigma_A^2}{2\pi} c^2,$$

where we applied the definition of all-pass filter for $z = e^{-i\lambda} \in \Gamma_1$. We also have that:

$$f_B(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \gamma_B(h) e^{-ik\lambda}$$

Comparing the two last equations and remembering that the Discrete-Time Fourier Transform is injective for ℓ^1 sequences, we get $\gamma_B(h) = c^2 \sigma_A^2 \delta_h$, $\square$.

A second, crucial property of all-pass filters is that they can be used to invert the moduli of the roots of a polynomial (example 3.2.2): let Q be a polynomial defined by $Q(z) = \prod_{k=1}^q (1 - \nu_k z)$, such that none of the ν_k have neither unitary nor zero modulus. We observe that $Q(0) = 1$ and that the q roots of Q are ν_k^{-1} for $k = 1, \ldots, q$.

Now we define the polinomial $\widetilde{Q}(z) = \prod_{k=1}^n \left(1 - \overline{\nu_k^{-1}} z\right)$ and the function $\Xi : z = \frac{Q}{\tilde{Q}}(z)$. Ξ is a rational function with poles $\overline{\nu_k} \neq \Gamma_1$. Then we know that it exists a unique ℓ^1 sequence ξ_k such that $\Xi(z) = \sum_{k \in \mathbb{Z}} \xi_k z^k$. Let us now prove that the filter F_ξ is then an all-pass. First, we have:

$$\left| \sum_{k \in \mathbb{Z}} \xi_k z^k \right| = \prod_{k=1}^n \frac{|1 - \nu_k z|}{\left|1 - \overline{\nu_k^{-1}} z\right|}. \tag{10}$$

Now, since $z \in \Gamma_1 \Rightarrow \bar{z} = z^{-1}$, for any $k = 1, \ldots, n$ and for $z \in \Gamma_1$ we have:

$$\begin{aligned}\left|1 - \overline{\nu_k^{-1}} z\right| &= \left|-\overline{\nu_k^{-1}} z\right| \left|-\overline{\nu_k} z^{-1} + 1\right| = \left|\overline{\nu_k^{-1}}\right| |z| \left|1 - \overline{\nu_k} z^{-1}\right| = \left|\overline{\nu_k^{-1}}\right| |z| \left|1 - \overline{\nu_k z}\right| \\ &= \left|\overline{\nu_k^{-1}}\right| \left|\overline{1 - \nu_k z}\right| = \left|\overline{\nu_k^{-1}}\right| \left|1 - \nu_k z\right|.\end{aligned}$$

Replacing in Eq. (10), we get:

$$\left|\sum_{k\in\mathbb{Z}}\xi_k z^k\right| = \prod_{k=1}^{n}\frac{|1-\nu_k z|}{\left|\overline{\nu_k^{-1}}\right||1-\nu_k z|} = \prod_{k=1}^{n}|\nu_k| = c > 0 \qquad \square$$

Equipped with the all-pass filter properties, we can rewrite an ARMA filter in canonical form. Let us consider an all-pass filter in the form $\Xi = a\frac{Q}{\widetilde{Q}}$. The roots ν_k^{-1} and the constant a will be defined later on. If Φ has no roots on Γ_1 we know that F_ϕ is invertible. Likewise, by construction Ξ is invertible. Using a fraction notation to refer to inverse filters, we can formally rewrite Eq. (9) as:

$$X = \frac{\Theta}{\Phi}\circ\epsilon = \frac{\Theta}{\Phi}\circ\frac{\Xi}{\Xi}\circ\epsilon = \left(\frac{\Theta}{\Phi\Xi}\right)\circ(\Xi\circ\epsilon) = \frac{\widetilde{\Theta}}{\widetilde{\Phi}}\circ\eta$$

with:

$$\frac{\widetilde{\Theta}}{\widetilde{\Phi}} = \frac{\Theta}{\Phi\Xi} \qquad\qquad \eta = \Xi\circ\epsilon.$$

We already know that η is a WN process, since it is an all-pass filtering of a WN. We have to show that we can build such a $\Xi(B) = a\frac{Q}{\widetilde{Q}}(B)$ that $\widetilde{\Theta}$ and $\widetilde{\Phi}$ do not have roots in the closed unit disk Δ_1. This is always possible since we can write:

$$\frac{\widetilde{\Theta}(z)}{\widetilde{\Phi}(z)} = \frac{\Theta(z)}{\Phi(z)}\frac{1}{\Xi(z)} = \frac{\prod_{k=1}^{q}(1-\nu_k^{(\theta)}z)}{\prod_{k=1}^{p}(1-\nu_k^{(\phi)}z)}\frac{1}{a}\prod_{k=1}^{n}\frac{1-\overline{\nu_k^{-1}}z}{1-\nu_k z}$$

where $\nu_k^{(\phi)}$ (resp. $\nu_k^{(\theta)}$) are the inverse of the roots of Φ (resp. of Θ). Now we build Ξ such that we cancel out the roots of Φ and of Θ in Δ_1. More precisely, to cancel out a given $\nu_k^{(\theta)}$ we introduce as a root of Q the number $\nu_k = \nu_k^{(\theta)}$ and to cancel out a given $\nu_k^{(\phi)}$ we introduce as a root of Q the number $\nu_k = \left(\overline{\nu_k^{(\theta)}}\right)^{-1}$.

In our case, we have: $\frac{\Theta(z)}{\Phi(z)} = \frac{1+4z}{1-2z}$ with roots $-\frac{1}{4}$ and $\frac{1}{2}$. To cancel out these roots, we set:

$$\frac{\widetilde{\Theta}(z)}{\widetilde{\Phi}(z)} = \frac{\Theta(z)}{\Phi(z)}\frac{1}{\Xi(z)} = \frac{1+4z}{1-2z}\cdot\frac{1}{a}\frac{1+\frac{1}{4}z}{1+4z}\frac{1-2z}{1-\frac{1}{2}z} = \frac{1}{a}\frac{1+\frac{1}{4}z}{1-\frac{1}{2}z}$$
$$\Xi(z) = a\frac{1+4z}{1+\frac{1}{4}z}\frac{1-\frac{1}{2}z}{1-2z}$$

Since $\forall z\in\Gamma_1, |\Xi(z)| = c$, given that $\Xi(1) = a\frac{5}{5/4}\frac{1/2}{-1} = -2a$, choosing $a = -1/2$ we get $\forall z\in\Gamma_1\ |\Xi(z)| = |\Xi(1)| = 1$. This also implies $f_\eta(\lambda) = f_\epsilon(\lambda)$ and thus $\mathsf{Var}\,(\eta) = \mathsf{Var}\,(\epsilon)$. In conclusion,

$$\boxed{\begin{aligned}\frac{\widetilde{\Theta}(z)}{\widetilde{\Phi}(z)} &= \frac{-2-\frac{1}{2}z}{1-\frac{1}{2}z}\\ \eta &= -\frac{1}{2}\frac{1+4z}{1+\frac{1}{4}z}\frac{1-\frac{1}{2}z}{1-2z}\epsilon\\ X_t-\frac{1}{2}X_{t-1} &= -2\eta_t-\frac{1}{2}\eta_{t-1}\end{aligned}}$$

3. Let us recall here the results of theorem 3.5.1. The canonical representation of an ARMA process is desirable since it express the former as an *causal* and *inversible* filtering of WN:

$$X_t = \widetilde{\phi}_1 X_{t-1} + \ldots + \widetilde{\phi}_p X_{t-p} + \widetilde{\theta}_0\eta_t + \widetilde{\theta}_1\eta_{t-1} + \ldots + \widetilde{\theta}_q\eta_{t-q}$$

This means that there exist two causal ℓ^1 sequences, ξ and $\widetilde{\xi}$, such that:

$$X = F_\xi(\eta) \tag{11}$$
$$\eta = F_{\widetilde{\xi}}(X) \tag{12}$$

From Eq. (11), since ξ is causal, we deduce that $\mathcal{H}_X^t \subseteq \mathcal{H}_Z^t$. From Eq. (12), since $\widetilde{\xi}$ is causal, we deduce that $\mathcal{H}_Z^t \subseteq \mathcal{H}_X^t$. In conclusion, $\mathcal{H}_X^t = \mathcal{H}_Z^t$. If we set:

$$\widehat{X}_t = \widetilde{\phi}_1 X_{t-1} + \ldots + \widetilde{\phi}_p X_{t-p} + +\widetilde{\theta}_1 \eta_{t-1} + \ldots + \widetilde{\theta}_q \eta_{t-q}$$

we see that $X_t - \widehat{X}_t = \widetilde{\theta}_0 \eta_t$. Since η is WN, $X_t - \widehat{X}_t \perp \mathcal{H}_\eta^{t-1}$ but then $X_t - \widehat{X}_t \perp \mathcal{H}_X^{t-1}$. This means that $\widehat{X}_t$ is the projection of X_t onto its linear past, and therefore $\widetilde{\theta}_0 \eta_t$ is the innovation process of X.

The canonical form gives therefore a direct access to the innovation of an ARMA process.

Now we can answer immediately to the question. The variance of the innovation is:

$$\mathsf{Var}\left(-2\eta_t\right) = 4\mathsf{Var}\left(\eta_t\right) = 4\mathsf{Var}\left(\epsilon_t\right).$$

4. From the definition of X we can write: $(1-2B)X_t = (1+4B)\epsilon_t$. Setting the AR process W_t such that $(1-2B)W_t = \epsilon_t$, we have $X_t = (1+4B)W_t$.

Now,

$$\begin{aligned}
W_t &= \frac{1}{1-2B}\epsilon_t = -\frac{1}{2B}\frac{1}{1-\frac{1}{2}B^{-1}}\epsilon_t = -\left(\frac{1}{2}B^{-1}\right)\sum_{k\geq 0}\left(\frac{1}{2}B^{-1}\right)^k \epsilon_t \\
&= -\sum_{k\geq 1}\left(\frac{1}{2}B^{-1}\right)^k \epsilon_t = -\sum_{k\geq 1}\left(\frac{1}{2}\right)^k \epsilon_{t+k} \\
X_t &= W_t + 4W_{t-1} = -\left[\sum_{k\geq 1}\left(\frac{1}{2}\right)^k \epsilon_{t+k}\right] - 4\left[\sum_{n\geq 1}\left(\frac{1}{2}\right)^n \epsilon_{t+n-1}\right] \quad \text{set } \ell = n-1 \\
&= -\left[\sum_{k\geq 1}\left(\frac{1}{2}\right)^k \epsilon_{t+k}\right] - 4\left[\sum_{\ell\geq 0}\left(\frac{1}{2}\right)^\ell \frac{1}{2}\epsilon_{t+\ell}\right] \\
&= -\left[\sum_{k\geq 1}\left(\frac{1}{2}\right)^k \epsilon_{t+k}\right] - 4\left[\frac{1}{2}\epsilon_t + \sum_{\ell\geq 1}\left(\frac{1}{2}\right)^\ell \frac{1}{2}\epsilon_{t+\ell}\right] \\
&= -2\epsilon_t - \left[\sum_{k\geq 1}\left(\frac{1}{2}\right)^k \epsilon_{t+k}\right] - 2\left[\sum_{\ell\geq 1}\left(\frac{1}{2}\right)^\ell \epsilon_{t+\ell}\right] \\
&= -2\epsilon_t - \sum_{k\geq 1}\frac{3}{2^k}\epsilon_{t+k}
\end{aligned}$$

Solution of Exercise 5.7 We have to compute the impulse response of a recursive filter. Since $|\phi| < 1$, a stable, causal solution exists. The weights ψ_k are such that:

$$\sum_{k\in\mathbb{Z}} \psi_k z^k = \frac{1}{1-\phi z} = \sum_{k\geq} \phi^k z^k \Rightarrow \psi_k = \begin{cases} \phi^k & \text{if } k \geq 0 \\ 0 & \text{if } k < 0 \end{cases}$$

Therefore, $X_t = \sum_{k\geq 0} \phi^k \epsilon_{t-k}$

2. We can apply Corollary 3.1.3 on the linear filtering of WN. Therefore, observing that ψ_k is real,

$$\begin{aligned}
\gamma_X(h) &= \sigma_\epsilon^2 \sum_{k\in\mathbb{Z}} \psi_{k+h}\psi_k = \\
&= \begin{cases} \sigma_\epsilon^2 \phi^h \sum_{k\geq 0} \phi^{2k} = \frac{\sigma_\epsilon^2 \phi^h}{1-\phi^2} & \text{if } h \geq 0 \\ \sigma_\epsilon^2 \sum_{k\geq -h} \phi^{k+h}\phi^k = \sigma_\epsilon^2 \sum_{n\geq 0} \phi^n \phi^{n-h} = \sigma_\epsilon^2 \phi^{-h} \sum_{k\geq 0} \phi^{2k} = \frac{\sigma_\epsilon^2 \phi^{-h}}{1-\phi^2} & \text{if } h < 0 \end{cases} \\
&= \frac{\sigma_\epsilon^2 \phi^{|h|}}{1-\phi^2}
\end{aligned}$$

A Annals

A.1 Exam of 2021

Documents autorisés: polycopié, notes de cours/TD.
Durée: 1 heure 30.

Authorized Documents : lecture notes, tutorial notes.
Duration: 1 hour and 30 minutes.

Exercise A.1. Let us consider $\{X_t, t \in \mathbb{Z}\}, \{Y_t, t \in \mathbb{Z}\}$ two L^2, stationary and independent stochastic processes with means μ_X, μ_Y and autocovariance functions γ_X, γ_Y, respectively.

1. Show that $S_t := X_t + Y_t$ is weakly stationary and compute its mean μ_S and autocovariance function γ_S.
2. Assuming that X and Y have spectral densities f_X and f_Y, show that S admits a spectral density f_S and express it using f_X and f_Y.
3. Show that the process $Z_t := X_t Y_t$ is weakly stationnary and compute its mean μ_Z and its autocovariance function γ_Z, first in the case where $\mu_X = \mu_Y = 0$, then in the general case.
4. Show that Z admits a spectral density f_Z and compute it first in the case where $\mu_X = \mu_Y = 0$, then in the general case. Use the convolution of two functions with a period of 2π defined by

$$f \star g\,(x) := \int_{-\pi}^{\pi} f\,(u)\, g\,(x-u)\, du$$

Exercise A.2. Consider a random process $X = (X_t)_{t\in\mathbb{Z}}$ satisfying the following recurrence equation:

$$X_t = \rho X_{t-1} + Z_t - (a + 1/a) Z_{t-1} + Z_{t-2} \tag{13}$$

where Z_t is a zero-mean weak white noise with variance σ^2 and both ρ and a are numbers in $(-1, 1)$ such that $a \neq \rho$ and $a \neq 0$.

5. Justify that this equation admits a weakly stationary solution X and find the expression of the power spectral density $f(\lambda)$ of this solution.
6. Is Eq. (14) an ARMA equation *in canonical form*?
7. Express X in its MA(∞) form, that is, compute (ϕ_k) such that

$$X_t = \sum_{k\in\mathbb{Z}} \phi_k Z_{t-k}\ .$$

8. Find b and c such that, for all $z \in \mathbb{C} \setminus \{a, 1/a\}$,

$$\frac{b}{1-az} + \frac{c}{1-z/a} = \frac{1}{1-(a+1/a)z+z^2}\ .$$

   Compute (ψ_k) using a, b, c and ρ such that

$$Z_t = \sum_{k\in\mathbb{Z}} \psi_k X_{t-k}\ .$$

9. Determine the variance of the innovation process W of X.
10. Compute (α_k) such that

$$W_t = \sum_{k\in\mathbb{Z}} \alpha_k Z_{t-k}\ .$$

11. Express $\mathrm{proj}(X_t|\mathcal{H}_{t-1}^X)$ using X and W.

A.2 Exam of 2020

Documents autorisés: polycopié, notes de cours/TD.
Durée: 1 heure 30.

Authorized Documents : lecture notes, tutorial notes.
Duration: 1 hour and 30 minutes.

Exercise A.3 (Representations of an ARMA(2,1) process). We consider a random process $(X_t)_{t\in\mathbb{Z}}$ satisfying the following recurrence equation:

$$X_t = 6X_{t-1} - 9X_{t-2} + \varepsilon_t + \frac{1}{2}\varepsilon_{t-1} \ , \tag{14}$$

where (ϵ_t) is a zero-mean weak white noise with variance σ^2.

1. Why does Eq. (14) admit a unique weakly stationary solution ? What is the nature of this solution (X_t)?
2. Find the expression of the power spectral density $f(\lambda)$ of the process X.
3. Find a canonical representation of X by using a suitable all-pass filter.
4. What is the innovation process of X? What is its variance?
5. Compute the coefficients $(\phi_k)_{k\geq 1}$ of the AR(∞) representation

$$X_t = \sum_{k\geq 1} \phi_k X_{t-k} + Z_t \ ,$$

where (Z_t) is the innovation process of (X_t).

Exercise A.4 (Linear prediction). Let $\{X_t, t \in \mathbb{Z}\}$ be a weakly stationary, zero-mean, real random process satisfying the equation

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

where $\theta \in]-1,1[$, and $\{Z_t, t \in \mathbb{Z}\}$ is weak noise with $\mathrm{Var}(Z_t) = \sigma^2$. Let $\hat{X}_t$ be a linear predictor of X_t of the form

$$\hat{X}_t = \sum_{k=1}^{P} \alpha_k X_{t-k},$$

with $P \in \mathbb{N}$ being the *order* of the predictor. Finally, we define

$$Y_t = X_t - \hat{X}_t,$$

as the prediction error. We want to compare the variance (power) and the autocorrelation function of the prediction error with those of the original process X. In several applications (*e.g.*, signal compression) it is desirable to have a prediction error with a smaller power than the original process. Also, achieving a white prediction error is desirable.

1. The input signal
   (a) Is X a causal filtering of Z?
   (b) Find the autocorrelation function (ACF) of X, $\rho_X(h)$
   (c) Find the variance of X_t
2. Simple first-order predictor
   (a) Let us consider the simplest predictor: $\hat{X}_t = X_{t-1}$. Find the variance of the prediction error.

(b) In which case the variance of Y is smaller than the variance of X?

(c) Find the ACF of Y

3. Optimal first-order predictor

   (a) The optimal first-order predictor is: $\hat{X}_t = \alpha X_{t-1}$ with $\alpha \in \mathbb{R}$ such that the variance of Y is minimized. Find the optimal value of α.
   Hint: recall that the optimal α is such that $\mathsf{Cov}\left(Y_t, X_{t-1}\right) = 0$

   (b) Find the variance of Y: is it smaller than that of X?

   (c) Find the ACF of Y and justify the name “whitening filter”

4. Optimal second-order predictor

   (a) A second-order predictor is in the form $\hat{X}_t = \alpha X_{t-1} + \beta X_{t-2}$. Show that for the optimal second-order predictor, $\beta = 0$, and conclude.

B Solutions of annals

Solution of Exercise A.1 1. $\mathbb{E}[X_t + Y_t] = \mathbb{E}(X_t) + \mathbb{E}(Y_t) = \mu_X + \mu_Y$, $\gamma_{X+Y}(t, t+h) = \gamma_X(h) + \gamma_Y(h)$ (does not depends on t)

2. By using Herglotz Theorem (2.3.1) + Corollary (2.3.2) : $\gamma_Z(h) = \gamma_X(h) + \gamma_Y(h) = \int e^{ih\lambda}\nu_X(d\lambda) + \int e^{ih\lambda}\nu_Y(d\lambda) = \int e^{ih\lambda}(\nu_X(d\lambda) + \nu_Y(d\lambda))$. Then $\nu_Z = \nu_X + \nu_Y$.
$f_Z(\lambda) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\gamma_Z(h)e^{-ih\lambda} = \frac{1}{2\pi}\left(\sum_{h\in\mathbb{Z}}\gamma_X(h)e^{-ih\lambda}\right) + \frac{1}{2\pi}\left(\sum_{h\in\mathbb{Z}}\gamma_Y(h)e^{-ih\lambda}\right) = f_X(\lambda) + f_Y(\lambda)$

3. Using the independence, we get: $\mathbb{E}[X_tY_t] = \mathbb{E}(X_t)\mathbb{E}(Y_t) = \mu_X\mu_Y$. Moreover:

$$\begin{aligned}
\gamma_{XY}(t, t+h) &= \mathbb{E}(X_tX_{t+h})\mathbb{E}(Y_tY_{t+h}) - \mu_X\mu_Y\mu_X\mu_Y \\
&= \left(\gamma_X(h) + \mu_X^2\right)\left(\gamma_Y(h) + \mu_Y^2\right) - \mu_X^2\mu_Y^2 \\
&= \gamma_X(h)\gamma_Y(h) + \mu_X^2\gamma_Y(h) + \mu_Y^2\gamma_X(h)
\end{aligned}$$

4. Let fix $t \in [-\pi, \pi]$, by using Fubini-Tonelli and Hertglotz theorem, we get:

$$\begin{aligned}
(f_X \star f_Y)(t) &= \int_{-\pi}^{\pi} f_X(u) f_Y(t-u)du \\
&= \int_{-\pi}^{\pi} \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\gamma_X(h)\mathrm{e}^{-ihu}\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\gamma_Y(k)\mathrm{e}^{-ik(t-u)}du \\
&= \frac{1}{2\pi}\sum_{h,k\in\mathbb{Z}}\gamma_X(h)\gamma_Y(k)\mathrm{e}^{-ikt}\underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-iu(h-k)}du}_{\mathbf{1}_{h=k}} \\
&= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\gamma_X(h)\gamma_Y(h)\mathrm{e}^{-iht} \\
&= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\mathrm{e}^{-iht}\left(\gamma_{XY}(h) - \mu_Y^2\gamma_X(h) - \mu_X^2\gamma_Y(h)\right) \\
&= f_{XY}(t) - \mu_Y^2 f_X(t) - \mu_X^2 f_Y(t)
\end{aligned}$$

Meaning that finally:

$$f_Z(t) = (f_X \star f_Y)(t) + \mu_Y^2 f_X(t) + \mu_X^2 f_Y(t)$$

Solution of Exercise A.2 5. Let $\Phi(z) = 1 - \rho z$ and $\Theta(z) = 1 - (a + 1/a)z + z^2$. Then Φ have no roots on the unit disk of $\mathbb{C}$. In this case the ARMA equation $\Phi(B)X = \Theta(B)Z$ has a unique weakly stationary solution X and its spectral density is

$$f(\lambda) = \frac{\sigma^2}{2\pi}\left|\frac{1 - (a+1/a)\mathrm{e}^{-\mathrm{i}\lambda} + 1}{1 - \rho\mathrm{e}^{-\mathrm{i}\lambda}}\right|^2 .$$

6. Since Θ has roots a and a^{-1} it vanishes on the unit disk. Hence (14) is not a canonical ARMA equation.

7. Inverting the filter $\Phi(B)$ leads to

$$\begin{aligned}
X_t &= \sum_{k=0}^{\infty}\rho^k(Z_{t-k} - (a+1/a)Z_{t-k-1} + Z_{t-k-2}) \\
&= Z_t + (\rho - a - 1/a)Z_{t-1} + \sum_{j\geq 2}\rho^{j-2}\left(\rho^2 - \rho(a+1/a) + 1\right)Z_{t-j} .
\end{aligned}$$

It is then easy to identify ϕ_j for all $j \geq 0$.

8. We must have $b + c = 1$ and $b/a + ca = 0$. We find $b = -ca^2$, so

$$c = 1/(1 - a^2) \; ,$$
$$b = -a^2/(1 - a^2) \; .$$

Since $|a| < 1$ and $a \neq 0$, we have for $z \in \mathbb{C}$ with $|z| = 1$,

$$(1 - az)^{-1} = \sum_{k=0}^{\infty} a^k z^k$$
$$(1 - z/a)^{-1} = -a/z(1 - a/z)^{-1} = -\sum_{k=1}^{\infty} a^k z^{-k} \; .$$

Hence we get

$$\frac{1}{1 - (a + 1/a)z + z^2} = \sum_{k=0}^{\infty} b\, a^k z^k - \sum_{-\infty<k<0} c\, a^{-k} z^k \; .$$

This provides the inverse linear filter of $\Theta(B)$ and we obtain

$$\begin{aligned} Z_t &= \sum_{k=0}^{\infty} b\, a^k (X_{t-k} - \rho X_{t-k-1}) - \sum_{-\infty<k<0} c\, a^{-k} (X_{t-k} - \rho X_{t-k-1}) \\ &= \sum_{k=0}^{\infty} b\, a^k X_{t-k} - \sum_{k=1}^{\infty} b\, a^{k-1} \rho X_{t-k} - \sum_{-\infty<k<0} c\, a^{-k} X_{t-k} + \sum_{-\infty<k\leq 0} c\, a^{-k+1} \rho X_{t-k} \\ &= \sum_{-\infty<k<0} c\, a^{-k} (a\, \rho - 1) X_{t-k} + (b + c\, a\, \rho) X_t + \sum_{k=1}^{\infty} b\, a^k (1 - a^{-1}\rho) X_{t-k} \; . \end{aligned}$$

It is then easy to identify ψ_j for all $j \in \mathbb{Z}$.

9. Let $W = F_\alpha(Z)$ with

$$\sum_{k\in\mathbb{Z}} \alpha_k z^k = \frac{1 - z/a}{1 - az} \; , z \in \mathbb{C}, \; |z| = 1 \; ,$$

Then if $|z| = 1$ we have $|\frac{1-z/a}{1-az}| = 1/|a|$ so that $W \sim \mathrm{WN}(0, \sigma^2/|a|^2)$ and, moreover, $Z = F_\beta(W)$ with

$$\sum_{k\in\mathbb{Z}} \beta_k z^k = \frac{1 - az}{1 - z/a} \; ,$$

which gives that

$$\Phi(B)X = \Theta(B) \circ F_\beta(W) = (1 - aB)^2 W \; .$$

The latter equation is canonical, so W is the innovation of X. It has variance $\sigma^2/|a|^2$.

10. Moreover W can be written as

$$W_t = \sum_{k\in\mathbb{Z}} \alpha_k Z_{t-k} \; .$$

The coefficients (α_k) are identified by the equation, for $z \in \mathbb{C}, \; |z| = 1$,

$$\begin{aligned} \sum_{k\in\mathbb{Z}} \alpha_k z^k &= \frac{1 - z/a}{1 - az} \\ &= (1 - z/a) \sum_{k\geq 0} (az)^k \\ &= \sum_{k\geq 0} (az)^k - \sum_{k\geq 0} a^{k-1} z^{k+1} \\ &= 1 + \sum_{k\geq 0} a^k \left(1 - a^{-2}\right) z^k \; . \end{aligned}$$

It is then easy to identify α_j for all $j \in \mathbb{Z}$.

11. Since $\Phi(B)X = (1-aB)^2 W$ is canonical, we have $\mathcal{H}^X_{t-1} = \mathcal{H}^W_{t-1} \perp W_t$ and

$$X_t = \rho X_{t-1} + W_t - 2aW_{t-1} + a^2 W_{t-2}$$

gives that

$$\text{proj}(X_t|\mathcal{H}^X_{t-1}) = \rho X_{t-1} - 2aW_{t-1} + a^2 W_{t-2} \ .$$

Solution of Exercise A.3 1. We have that $\Phi(z) := 1 - 6z + 9z^2 = (1-3z)^2$ dos not vanish on the unit circle, which ensures existence and uniqueness of the solution, which is then called an ARMA(2,1) process.

2. The spectral density is given by

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|1 + \mathrm{e}^{-\mathrm{i}\lambda}/2\right|^2}{\left|1 - 3\mathrm{e}^{-\mathrm{i}\lambda}\right|^4} \ .$$

3. Let F_β denote the all-pass filter with coefficients $(\beta_k) \in \ell^1$ defined by the equation

$$\sum_{k\in\mathbb{Z}} \beta_k z^k = \frac{1 - z^{-1}/3}{1-3z} \ , \qquad z \in \mathbb{C} \ , \ |z| = 1 \ .$$

We apply this filter twice on both sides of (14) and obtain that X is solution of

$$(1 - \mathrm{B}/3)^2 \, X = (1 + \mathrm{B}/2) \, Z \ , \tag{15}$$

where $Z = \mathrm{F}_\beta(\epsilon)$ has spectral density

$$f^Z(\lambda) = \frac{\sigma^2}{2\pi} \left| \frac{1 - \mathrm{e}^{\mathrm{i}\lambda}/3}{1 - 3\mathrm{e}^{-\mathrm{i}\lambda}} \right|^4 = \frac{\sigma^2}{3^4 . 2\pi}$$

Hence Z is a white noise with variance $\sigma^2/3^4$. The representation (15) is a canonical representation of X.

4. From the previous question, we deduce that Z is the innovation process of X. It has variance $\sigma^2/3^4$.

5. From (15), we have

$$Z = \mathrm{F}_\alpha(X) \ ,$$

where $(\alpha)_k$ is the ℓ^1 sequence satisfying

$$\sum_{k\in\mathbb{Z}} \alpha_k z^k = \frac{(1-z/3)^2}{1+z/2} \ , \qquad z \in \mathbb{C} \ , \ |z| = 1 \ .$$

Now, we have, for all $z \in \mathbb{C}$ with $|z| = 1$,

$$\begin{aligned}
\frac{(1-z/3)^2}{1+z/2} &= (1-z/3)^2 \sum_{k\geq 0} (-1/2)^k z^k \\
&= (1-z/3) \left(1 + \sum_{k\geq 1} \left((-1/2)^k - (-1/2)^{k-1}/3\right) \, z^k \right) \\
&= (1-z/3) \left(1 + \frac{5}{3} \sum_{k\geq 1} (-1/2)^k \, z^k \right) \\
&= 1 - \frac{7}{6} z + \frac{5}{3} \sum_{k\geq 2} \left((-1/2)^k - (-1/2)^{k-1}/3\right) \, z^k \\
&= 1 - \frac{7}{6} z + \left(\frac{5}{3}\right)^2 \sum_{k\geq 2} (-1/2)^k \, z^k \ .
\end{aligned}$$

We conclude that $\phi_1 = -\alpha_1 = 7/6$ and, for all $k \geq 2$, $\phi_k = -\alpha_k = -(5/3)^2(-1/2)^k$.

Solution of Exercise A.4 1. The input signal

(a) X is a causal filtering of Z because the only root of the polynomial $\Theta(z) = 1 - \theta z$ is $\frac{1}{\theta}$, outside the unit circle. Therefore, one can write $X_t = \sum_{\ell \geq 0} \theta^\ell Z_{t-\ell}$

(b) For $h \geqslant 0$, the autocovariance function of X, $\gamma_X(h)$ is

$$\begin{aligned}
\gamma_X(h) &= \mathbb{E}\left[\sum_{\ell\geq 0} \theta^\ell Z_{n-\ell} \sum_{k\geq 0} \theta^k Z_{n+h-k}\right] = \sum_{\ell\geq 0}\sum_{k\geq 0} \theta^{\ell+k}\mathbb{E}\left[Z_{n-\ell}Z_{n+h-k}\right] \\
&= \sum_{\ell\geq 0}\sum_{k\geq 0} \theta^{\ell+k}\sigma^2\delta_{k-(\ell+h)} = \sum_{\ell\geq 0}\sigma^2\theta^{2\ell+h} \\
&= \theta^h \frac{\sigma^2}{1-\theta^2}
\end{aligned}$$

By symmetry, we have $\gamma_X(h) = \theta^{|h|}\frac{\sigma^2}{1-\theta^2}$. Thus the ACF reads

$$\rho_X(h) = \gamma_X(h)/\gamma_X(0) = \theta^{|h|}\ , \quad h \in \mathbb{Z}\ .$$

(c) The variance of X_t is

$$\sigma_X^2 = \gamma_X(0) = \frac{\sigma^2}{1-\theta^2}\ .$$

2. Simple first order predictor

(a) The variance of the prediction error is:

$$\begin{aligned}
\mathsf{Var}\,(Y_t) &= \mathbb{E}\left[Y_t^2\right] = \mathbb{E}\left[(X_t - X_{t-1})^2\right] = \mathbb{E}\left[X_t^2 + X_{t-1}^2 - 2X_tX_{t-1}\right] \\
&= 2\gamma_X(0) - 2\gamma_X(1) = \frac{2\sigma^2}{1-\theta^2}(1-\theta) \\
&= \sigma_X^2 2(1-\theta)
\end{aligned}$$

(b) From the previous, the variance of Y is smaller than the variance of X if and only if $2(1-\theta) < 1$, implying $\theta > \frac{1}{2}$. Also, remember that $\theta < 1$ by hypothesis. In conclusion the simple predictor is effective only if consecutive samples of X are correlated enough.

(c) The autocovariance function of Y is computed as follows for $h > 0$:

$$\begin{aligned}
\gamma_Y(h) &= \mathbb{E}\left[Y_tY_{t+h}\right] = \mathbb{E}\left[(X_t - X_{t-1})(X_{t+h} - X_{t-1+h})\right] \\
&= 2\gamma_X(h) - \gamma_X(h-1) - \gamma_X(h+1) = \frac{\sigma^2}{1-\theta^2}\left(2\theta^h - \theta^{h-1} - \theta^{h+1}\right) \\
&= \frac{-\sigma^2}{1-\theta^2}(1-\theta)^2\theta^{h-1} = \frac{-\sigma^2}{1+\theta}(1-\theta)\,\theta^{h-1} = \frac{-(1-\theta)\theta^{h-1}}{2}\sigma_Y^2
\end{aligned}$$

For $h = 0$, $\gamma_Y(h) = \mathsf{Var}\,(Y_t)$ and for $h < 0$, $\gamma_Y(h) = \gamma_Y(-h)$.

3. Optimal first order predictor

(a) The optimal first order predictor is found by setting $\mathsf{Cov}\,(\alpha X_{t-1} - X_t, X_{t-1}) = 0$

$$\begin{aligned}
0 &= \mathsf{Cov}\,(\alpha X_{t-1} - X_t, X_{t-1}) = \alpha\gamma_X(0) - \gamma_X(1) \\
\alpha &= \frac{\gamma_X(1)}{\gamma_X(0)} = \theta \\
\hat{X}_t &= \theta X_{t-1} \\
Y_t &= X_t - \theta X_{t-1} = Z_t
\end{aligned}$$

(b) Since $Y_t = Z_t$, its variance is σ^2, which is smaller than $\sigma_X^2 = \frac{\sigma^2}{1-\theta^2}$ for any $\theta \in]-1,1[$. The variance of Y_t can also be found explicitly as $\mathbb{E}\left[(X_t - \theta X_{t-1})^2\right]$.

(c) The ACF of Y is the one of Z: $\rho_Y(h) = \delta_h$. Therefore Y is white noise. Again, γ_Y can be found by calculating $\mathbb{E}\left[(X_t - \theta X_{t-1})(X_{t+h} - \theta X_{t-1+h})\right]$

4. Optimal second order predictor

(a) The optimal second order predictor is such that:

$$
\begin{aligned}
\mathsf{Cov}\left(\alpha X_{t-1} + \beta X_{t-2} - X_t, X_{t-1}\right) &= 0 & \qquad \alpha\gamma_X(0) + \beta\gamma_X(1) &= \gamma_X(1) \\
\mathsf{Cov}\left(\alpha X_{t-1} + \beta X_{t-2} - X_t, X_{t-2}\right) &= 0 & \qquad \alpha\gamma_X(1) + \beta\gamma_X(0) &= \gamma_X(2)
\end{aligned}
$$

$$
\begin{aligned}
\begin{bmatrix} \gamma_X(0) & \gamma_X(1) \\ \gamma_X(1) & \gamma_X(0) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} &= \begin{bmatrix} \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} \\
\begin{bmatrix} \alpha \\ \beta \end{bmatrix} &= \frac{1}{\gamma_X^2(0) - \gamma_X^2(1)} \begin{bmatrix} \gamma_X(0) & -\gamma_X(1) \\ -\gamma_X(1) & \gamma_X(0) \end{bmatrix} \begin{bmatrix} \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} \\
\beta &= \frac{\gamma_X(0)\gamma_X(2) - \gamma_X^2(1)}{\gamma_X^2(0) - \gamma_X^2(1)}
\end{aligned}
$$

But $\gamma_X(0)\gamma_X(2) - \gamma_X^2(1) = \sigma_X^4\theta^2 - \sigma_X^4\theta^2 = 0$, thus $\beta = 0$.
Conclusion: since X is an AR(1) process, there is no advantage in considering linear predictors of order greater than 1.