TSIA 202a – Exam

Nov 4th, 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 1: Representations of an ARMA(2,1) process

We consider a random process $(X_t)_{t\in\mathbb{Z}}$ satisfying the following recurrence equation:

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

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

1. Why does Eq. (1) 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).

Answer of exercise 1

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 (1) and obtain that X is solution of

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

   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 (2) 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 (2), 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 $k \geq 2$, $\phi_k = -\alpha_k = -(5/3)^2(-1/2)^k$.

Exercise 2: 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, that is

$$\hat{X}_t = \sum_{k=1}^{P} \alpha_p X_t - p,$$

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, $\gamma_X(h)$
   (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: the optimal predictor is such that the prediction error is uncorrelated with the linear past, in particular $\mathrm{Cov}\,(Y_t, X_{t-1}) = 0$
   (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.

Answer of exercise 2

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

   (c) The variance of X_t is easily found as $\gamma_X(0)$:

$$\sigma_X^2 = \frac{\sigma^2}{1-\theta^2}$$

2. Simple first order predictor

   (a) The variance of the prediction error is:

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

   (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 ACF of Y is computed as follows for $h > 0$:

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

   For $h = 0$, $\gamma_Y(h) = \mathrm{Var}\left(Y_t\right)$ 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 $\operatorname{Cov}(\alpha X_{t-1} - X_t, X_{t-1}) = 0$

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

(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: $\gamma_Y(h) = \sigma^2\delta_h$. 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:

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

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

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.