System Identification

1 First exercise

Exercise

Assume to have at your disposal an infinite amount of sample taken from the zero-mean stochastic process with the following auto-covariance function:

$$\gamma_y(\tau) = \begin{cases} 4 & \text{if } \tau = 0 \\ -2 & \text{if } |\tau| = 1 \\ 1 & \text{if } |\tau| = 2 \\ 0 & \text{if } |\tau| > 2 \end{cases}$$

Consider the following two classes of models:

$$\begin{aligned} \mathcal{M}_1(a) &: y(t) = \eta_1(t) + a \cdot y(t-1) & \eta_1 \sim WN\left(0, \lambda_1^2\right) \\ \mathcal{M}_2(b,c) &: y(t) = \eta_2(t) + b \cdot y(t-2) + c \cdot y(t-3) & \eta_2 \sim WN\left(0, \lambda_2^2\right) \end{aligned}$$

1. Identify the parameter a and λ_1^2 of the family $\mathcal{M}_1$
2. Identify the parameters b, c and λ_3^2 of the family $\mathcal{M}_2$

Solution

First point The predictor of the model $\mathcal{M}_1$ is:

$$\hat{y}(t|t-1;\vartheta) = a \cdot y(t-1)$$

therefore, the prediction error is:

$$\begin{aligned} \varepsilon_1(t;\vartheta) &= y(t) - \hat{y}(t|t-1;\vartheta) \\ &= y(t) - a \cdot y(t-1) \end{aligned}$$

and the PEM cost function is:

$$\begin{aligned} \bar{V}_1(\vartheta) &= \mathbb{E}\left[\varepsilon_1(t;\vartheta)^2\right] \\ &= \mathbb{E}\left[(y(t) - a \cdot y(t-1))^2\right] \\ &= \mathbb{E}\left[y(t)^2 - 2 \cdot a \cdot y(t) \cdot y(t-1) + a^2 \cdot y(t-1)^2\right] \\ &= \mathbb{E}\left[y(t)^2\right] - 2 \cdot a \cdot \mathbb{E}\left[y(t) \cdot y(t-1)\right] + a^2 \cdot \mathbb{E}\left[y(t-1)^2\right] \\ &= \gamma_y(0) - 2 \cdot a \cdot \gamma_y(1) + a^2 \cdot \gamma_y(0) \\ &= 4 - 2 \cdot a \cdot (-2) + a^2 \cdot 4 \\ &= 4 \cdot a^2 + 4 \cdot a + 4 \end{aligned}$$

in order to find the stationary point, we can compute its gradient:

$$\begin{aligned} \nabla \bar{V}_1(\vartheta) &= \frac{d}{da}\left(4 \cdot a^2 + 4 \cdot a + 4\right) \\ &= 8 \cdot a + 4 \end{aligned}$$

and search for its roots:

$$\begin{aligned} \nabla \bar{V_1}\left(\vartheta\right) &= 0 \\ 8 \cdot a + 4 &= 0 \\ a &= -\frac{1}{2} \end{aligned}$$

therefore:

$$\hat{\vartheta} = \hat{a} = -\frac{1}{2}$$

The estimate of λ_1^2 can be found by evaluating the cost function in $\hat{\vartheta}$:

$$\begin{aligned} \hat{\lambda}_1^2 &= \bar{V_1}\left(\hat{\vartheta}\right) \\ &= 4 \cdot \frac{1}{4} + 4 \cdot \left(-\frac{1}{2}\right) + 4 \\ &= 1 - 2 + 4 \\ &= 3 \end{aligned}$$

therefore, the estimated model is:

$$y\left(t\right) = \eta_1\left(t\right) - \frac{1}{2} \cdot y\left(t-1\right)$$

with $\eta_1\left(t\right) \sim WN\left(0, 3\right)$.

Second point The predictor of the model $\mathcal{M}_2$ is:

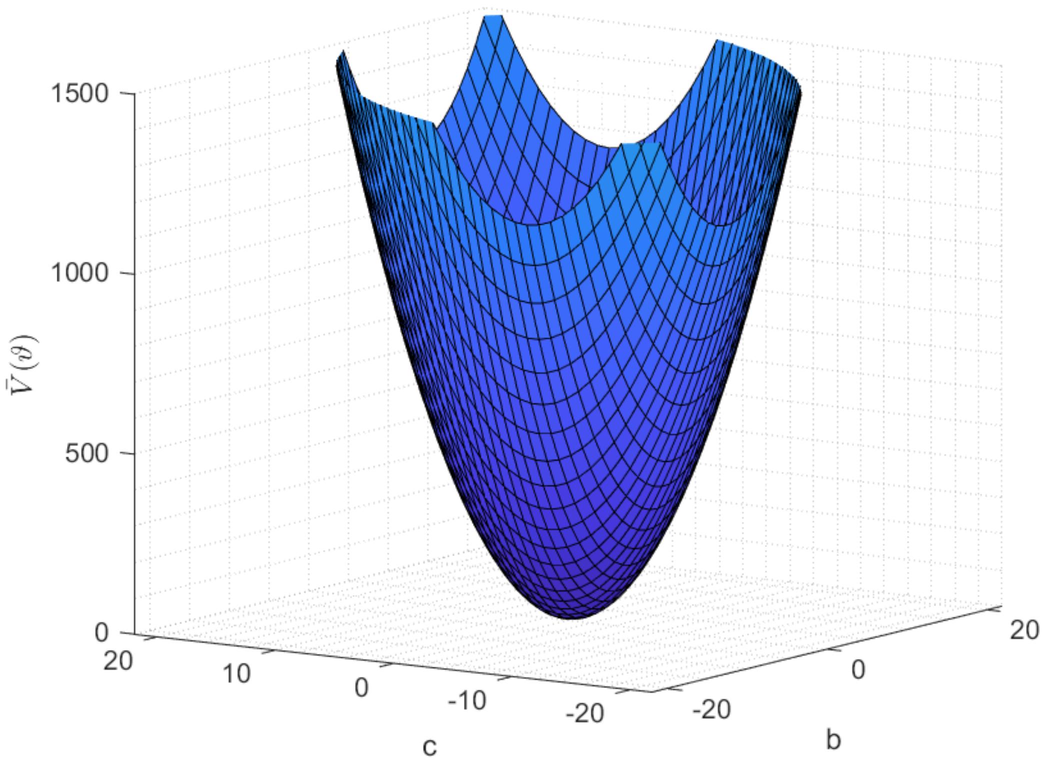
$$\hat{y}\left(t|t-1;\vartheta\right) = b \cdot y\left(t-2\right) + c \cdot y\left(t-3\right)$$

therefore, the prediction error is:

$$\begin{aligned} \varepsilon_2\left(t;\vartheta\right) &= y\left(t\right) - \hat{y}\left(t|t-1;\vartheta\right) \\ &= y\left(t\right) - b \cdot y\left(t-2\right) - c \cdot y\left(t-3\right) \end{aligned}$$

and the PEM cost function is:

$$\begin{aligned} \bar{V_2}\left(\vartheta\right) &= \mathbb{E}\left[\varepsilon_2\left(t;\vartheta\right)^2\right] \\ &= \mathbb{E}\left[\left(y\left(t\right) - b \cdot y\left(t-2\right) - c \cdot y\left(t-3\right)\right)^2\right] \\ &= \mathbb{E}\left[y\left(t\right)^2 + b^2 \cdot y\left(t-2\right)^2 + c^2 \cdot y\left(t-3\right)^2 - 2 \cdot b \cdot y\left(t\right) \cdot y\left(t-2\right) - 2 \cdot c \cdot y\left(t\right) \cdot y\left(t-3\right) + 2 \cdot b \cdot c \cdot y\left(t-2\right) \cdot y\left(t-3\right)\right] \\ &= \mathbb{E}\left[y\left(t\right)^2\right] + b^2 \cdot \mathbb{E}\left[y\left(t-2\right)^2\right] + c^2 \cdot \mathbb{E}\left[y\left(t-3\right)^2\right] - \\ &\quad - 2 \cdot b \cdot \mathbb{E}\left[y\left(t\right) \cdot y\left(t-2\right)\right] - 2 \cdot c \cdot \mathbb{E}\left[y\left(t\right) \cdot y\left(t-3\right)\right] + 2 \cdot b \cdot c \cdot \mathbb{E}\left[y\left(t-2\right) \cdot y\left(t-3\right)\right] \\ &= \gamma_y\left(0\right) + b^2 \cdot \gamma_y\left(0\right) + c^2 \cdot \gamma_y\left(0\right) - 2 \cdot b \cdot \gamma_y\left(2\right) - 2 \cdot c \cdot \gamma_y\left(3\right) + 2 \cdot b \cdot c \cdot \gamma_y\left(1\right) \\ &= 4 + 4 \cdot b^2 + 4 \cdot c^2 - 2 \cdot b \cdot 1 - 2 \cdot c \cdot 0 + 2 \cdot b \cdot c \cdot \left(-2\right) \\ &= 4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4 \end{aligned}$$



In order to find the stationary point, we can compute its gradient:

$$\nabla \bar{V}_2(\vartheta) = \begin{bmatrix} \frac{\partial}{\partial b} \bar{V}_2(b, c) \\ \frac{\partial}{\partial b} \bar{V}_2(b, c) \end{bmatrix}$$

$$\begin{aligned} \frac{\partial}{\partial b} \bar{V}_2(b, c) &= \frac{\partial}{\partial b} \left(4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4\right) \\ &= 8 \cdot b - 4 \cdot c + 0 - 2 + 0 \\ &= 8 \cdot b - 4 \cdot c - 2 \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial c} \bar{V}_2(b, c) &= \frac{\partial}{\partial c} \left(4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4\right) \\ &= 0 - 4 \cdot b + 8 \cdot c - 0 + 0 \\ &= -4 \cdot b + 8 \cdot c \end{aligned}$$

and search for its roots. Obtaining the linear system:

$$\begin{cases} 8 \cdot b - 4 \cdot c - 2 = 0 \\ -4 \cdot b + 8 \cdot c = 0 \end{cases}$$

$$\begin{cases} 4 \cdot b - 2 \cdot c = 1 \\ -b + 2 \cdot c = 0 \end{cases}$$

$$\begin{cases} 4 \cdot b - 2 \cdot c = 1 \\ 2 \cdot c = b \end{cases}$$

$$\begin{cases} 4 \cdot b - b = 1 \\ 2 \cdot c = b \end{cases}$$

$$\begin{cases} b = \frac{1}{3} \\ 2 \cdot c = \frac{1}{3} \end{cases}$$

therefore:

$$\hat{b} = \frac{1}{3}$$

$$\hat{c} = \frac{1}{6}$$

The estimate of λ_2^2 can be found by evaluating the cost function in $\hat{\vartheta}$:

$$\begin{aligned} \hat{\lambda}_1^2 &= \bar{V}_2\left(\hat{b}, \hat{c}\right) \\ &= 4 \cdot \hat{b}^2 - 4 \cdot \hat{b} \cdot \hat{c} + 4 \cdot \hat{c}^2 - 2 \cdot \hat{b} + 4 \\ &= 4 \cdot \left(\frac{1}{3}\right)^2 - 4 \cdot \frac{1}{3} \cdot \frac{1}{6} + 4 \cdot \left(\frac{1}{6}\right)^2 - 2 \cdot \frac{1}{3} \cdot +4 \\ &= \frac{4}{9} - \frac{4}{18} + \frac{4}{36} - \frac{2}{3} + 4 \\ &= \frac{11}{3} \end{aligned}$$

therefore, the estimated model is:

$$y\left(t\right) = \eta_2\left(t\right) + \frac{1}{3} \cdot y\left(t-2\right) + \frac{1}{6} \cdot y\left(t-3\right)$$

with $\eta_2\left(t\right) \sim WN\left(0, \frac{11}{3}\right)$.

2 Second exercise (December 2018, follow up)

Exercise

Consider the following dataset:

t	1	2	3	4	5
$u(t)$	1	0	-1	0	1
$y(t)$	1	-1	0	1	-1

and the following families of model:

$$\mathcal{M}_1 : \quad y(t) = b_1 \cdot u(t-1) + e_1(t) \qquad e_1(t) \sim WN\left(0, \lambda_1^2\right)$$
$$\mathcal{M}_2 : \quad y(t) = a_2 \cdot y(t-1) + b_2 \cdot u(t-1) + e_2(t) \qquad e_2(t) \sim WN\left(0, \lambda_2^2\right)$$

1. According to the *Akaike Information criteria* select the best family between $\mathcal{M}_1$ and $\mathcal{M}_2$ for this dataset

Solution

The Akaike information criteria allows us to select the right complexity of the model by minimizing the cost function:

$$AIC = 2 \cdot \frac{n_\vartheta}{N} + \ln\left(V_N\left(\hat{\vartheta}\right)\right)$$

where n_ϑ is the number of parameters to identify, N is the number of data and $V_N\left(\hat{\vartheta}\right)$ is the cost function used for the identification evaluated using the estimated parameters.

For the first family, we have only $n_\vartheta = 1$ parameters and $N = 4$ data. Recalling that:

$$V_N(\vartheta) = \frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) - b_1 \cdot u(t-1)\right)^2$$
$$\hat{\vartheta} = \hat{b}_1 = -1$$

The AIC cost function assumes the value:

$$\begin{aligned}
AIC_1 &= 2 \cdot \frac{1}{4} + \ln\left(V_N\left(\hat{b}_1\right)\right) \\
&= \frac{1}{2} + \ln\left(\frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) - \hat{b}_1 \cdot u(t-1)\right)^2\right) \\
&= \frac{1}{2} + \ln\left(\frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) + 1 \cdot u(t-1)\right)^2\right) \\
&= \frac{1}{2} + \ln\left(\frac{1}{4} \cdot \left[(-1+1)^2 + (0-0)^2 + (1-1)^2 + (-1+0)^2\right]\right) \\
&= \frac{1}{2} + \ln\left(\frac{1}{4} \cdot 1\right) \\
&= \frac{1}{2} + \ln\left(\frac{1}{4}\right) \simeq -0.886
\end{aligned}$$

For the first family, we have only $n_\vartheta = 2$ parameters and $N = 4$ data. Recalling that:

$$V_N(\vartheta) = \frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) - a_2 \cdot y(t-1) - b_2 \cdot u(t-1)\right)^2$$

$$\hat{\vartheta} = \begin{bmatrix} \hat{a}_2 \\ \hat{b}_2 \end{bmatrix}$$

The AIC cost function assumes the value:

$$\begin{aligned} AIC_2 &= 2 \cdot \frac{2}{4} + \ln\left(V_N\left(\hat{b}_2, \hat{a}_2\right)\right) \\ &= 1 + \ln\left(\frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) - \hat{a}_2 \cdot y(t-1) - \hat{b}_2 \cdot u(t-1)\right)^2\right) \\ &= 1 + \ln\left(\frac{1}{4} \cdot \sum_{t=2}^{5} \left(y(t) + \frac{2}{5} \cdot y(t-1) + \frac{4}{5} \cdot u(t-1)\right)^2\right) \\ &= 1 + \ln\left(\frac{1}{4} \cdot \left[\left(-1 + \frac{2}{5} \cdot 1 + \frac{4}{5} \cdot 1\right)^2 + \left(0 + \frac{2}{5} \cdot (-1) + \frac{4}{5} \cdot 0\right)^2 + \left(1 + \frac{2}{5} \cdot 0 + \frac{4}{5} \cdot (-1)\right)^2 + \left(-1 + \frac{2}{5} \cdot 1 + \frac{4}{5} \cdot 0\right)^2\right]\right) \\ &= 1 + \ln\left(\frac{1}{4} \cdot \left[\left(\frac{1}{5}\right)^2 + \left(-\frac{2}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(-\frac{3}{5}\right)^2\right]\right) \\ &= 1 + \ln\left(\frac{1}{4} \cdot \frac{3}{5}\right) \\ &= 1 + \ln\left(\frac{3}{20}\right) \simeq -0.8971 \end{aligned}$$

since $AIC_2 < AIC_1$ the second family is better.

3 Third exercise

Exercise

Consider a non-null mean WSS process and assume samples from one of its realizations are collected:

t	0	1	2	3	4	5	6
$y(t)$	1	$-\frac{1}{2}$	0	$-\frac{1}{4}$	$\frac{3}{4}$	0	-1

and the following families of model:

$$\mathcal{M}_1 \quad : \quad y(t) = a_1 \cdot y(t-1) + e(t) \qquad\qquad e(t) \sim WN\left(0, \lambda_1^2\right)$$

1. Identify the parameter a_1 and λ_1^2 of the family $\mathcal{M}_1$.
2. Compute the *AIC* of the family$\mathcal{M}_1$.

Solution

The optimal 1-step predictor $\hat{y}(t|t-1;\vartheta)$ of $\mathcal{M}_1$ is:

$$\begin{aligned}
\hat{y}(t|t-r;\vartheta) &= \frac{C(z) - A(z)}{C(z)} \cdot y(t) = \\
&= \frac{1 - \left(1 - a_1 \cdot z^{-1}\right)}{1} \cdot y(t) \\
&= a_1 \cdot z^{-1} \cdot y(t) \\
&= a_1 \cdot y(t-1)
\end{aligned}$$

This predictor can also be found by noting that:

$$y(t) = \underbrace{a_1 \cdot y(t-1)}_{\text{known part}} + \underbrace{e(t)}_{\text{unknown part}}$$

and therefore the predictor is:

$$\hat{y}(t|t-1;\vartheta) = a_1 \cdot y(t-1)$$

Once the predictor is found, it's possible to write the objective function:

$$\begin{aligned}
V_N(\vartheta) &= \frac{1}{6} \cdot \sum_{t=1}^{6} \left(y(t) - \hat{y}(t|t-1;\vartheta)\right)^2 \\
&= \frac{1}{6} \cdot \sum_{t=1}^{6} \left(y(t) - a_1 \cdot y(t-1)\right)^2
\end{aligned}$$

and its gradient:

$$\begin{aligned}
\nabla V_N(\vartheta) &= \frac{1}{6} \cdot \sum_{t=1}^{6} \frac{\partial}{\partial a_1} \left(y(t) - a_1 \cdot y(t-1)\right)^2 \\
&= 2 \cdot \frac{1}{6} \cdot \sum_{t=1}^{6} \left(y(t) - a_1 \cdot y(t-1)\right) \cdot \left(-y(t-1)\right) \\
&= \frac{1}{3} \cdot \sum_{t=1}^{6} \left[-y(t) \cdot y(t-1) + a_1 \cdot y(t-1)^2\right]
\end{aligned}$$

Stationary points can be computed as:

$$\nabla V_N\left(\vartheta\right)\big|_{\vartheta=\hat{\vartheta}} = 0$$

Thus:

$$\begin{gathered}\nabla V_N\left(\vartheta\right) = 0\\ \cancel{\frac{1}{3}}\cdot\sum_{t=1}^{6}\left[-y\left(t\right)\cdot y\left(t-1\right)+a_1\cdot y\left(t-1\right)^2\right] = 0\\ a_1\cdot\sum_{t=1}^{6}y\left(t-1\right)^2 = \sum_{t=1}^{6}y\left(t\right)\cdot y\left(t-1\right)\\ a_1\cdot\left(\left(1\right)^2+\left(-\frac{1}{2}\right)^2+\left(0\right)^2+\left(-\frac{1}{4}\right)^2+\left(\frac{3}{4}\right)^2+\left(0\right)^2\right) = \left(\left(-\frac{1}{2}\right)+\left(0\right)+\left(0\right)+\left(-\frac{3}{16}\right)+\left(0\right)+\left(0\right)\right)\\ a_1\cdot\frac{30}{16} = -\frac{11}{16}\end{gathered}$$

The stationary point is then:

$$\hat{a}_1 = -\frac{11}{30}$$

The Akaike information criteria allows us to select the right complexity of the model by minimizing the cost function:

$$AIC = 2\cdot\frac{n_\vartheta}{N}+\ln\left(V_N\left(\hat{\vartheta}\right)\right)$$

where n_ϑ is the number of parameters to identify, N is the number of data and $V_N\left(\hat{\vartheta}\right)$ is the cost function used for the identification evaluated using the estimated parameters.

For the first family, we have only $n_\vartheta = 1$ parameters and $N = 6$ data. The AIC cost function assumes the value:

$$\begin{aligned}AIC &= 2\cdot\frac{1}{6}+\ln\left(V_N\left(\hat{a}_1\right)\right)\\ &= \frac{1}{3}+\ln\left(\frac{1}{6}\cdot\sum_{t=1}^{6}\left(y\left(t\right)-\hat{a}_1\cdot y\left(t-1\right)\right)^2\right)\\ &= \frac{1}{3}+\ln\left(\frac{1}{6}\cdot\sum_{t=1}^{6}\left(y\left(t\right)+\frac{11}{30}\cdot y\left(t-1\right)\right)^2\right)\\ &= \frac{1}{3}+\ln\left(\frac{1}{6}\cdot\left[\left(-\frac{1}{2}+\frac{11}{30}\cdot 1\right)^2+\left(0-\frac{11}{30}\cdot\frac{1}{2}\right)^2+\left(-\frac{1}{4}+\frac{11}{30}\cdot 0\right)^2+\left(\frac{3}{4}-\frac{11}{30}\cdot\frac{1}{4}\right)^2+\left(0+\frac{11}{30}\cdot\frac{3}{4}\right)^2+\left(-1+\frac{11}{30}\cdot 0\right)^2\right]\right)\\ &= \frac{1}{3}+\ln\left(\frac{1}{6}\cdot\left[\left(-\frac{4}{30}\right)^2+\left(-\frac{11}{60}\right)^2+\left(-\frac{1}{4}\right)^2+\left(\frac{79}{120}\right)^2+\left(\frac{33}{120}\right)^2+\left(-1\right)^2\right]\right)\\ &= \frac{1}{3}+\ln\left(\frac{1}{6}\cdot\frac{779}{480}\right)\simeq -0.974\end{aligned}$$