EXERCISES (3): identification

Exercise 1

Consider the data generation mechanism ("real" system):

$$S:\ y(t) = e(t) + \frac{1}{2}e(t-1), \quad e(t) \sim WN(0,1)$$

and the classes of models:

$$M_1:\ y(t) = ay(t-1) + \eta(t), \quad \eta(t) \sim WN(0, \lambda_1^2)$$

$$M_2:\ y(t) = -ay(t-1) - by(t-2) + \eta(t), \quad \eta(t) \sim WN(0, \lambda_2^2)$$

$$M_3:\ y(t) = \eta(t) + a\eta(t-1), \quad \eta(t) \sim WN(0, \lambda_3^2)$$

Compute the value $\boldsymbol{\vartheta}^*$ of the vector of parameters $\boldsymbol{\vartheta}$ which minimizes the loss function: $\bar{J}(\boldsymbol{\vartheta}) = E\left[\left(y(t) - \hat{y}(t|t-1;\vartheta)\right)^2\right]$

First of all, notice that, since the data generation mechanism is given, we can study the ASYMPTOTICAL BEHAVIOUR of the identification method.

So, when THE DATA GENERATION SYSTEM IS GIVEN WE CAN ASSUME THAT WE HAVE COLLECTED AN INFINITE AMOUNT OF DATA:

$$N \to \infty$$

This means that we study the asymptotic case:

$$J_N(\boldsymbol{\vartheta}) = \frac{1}{N}\sum_{t=1}^{N}\left(y(t) - \hat{y}(t|t-1;\boldsymbol{\vartheta})\right)^2 \xrightarrow[N\to\infty]{} \bar{J}(\boldsymbol{\vartheta})$$
$$= E\left[\left(y(t) - \hat{y}(t|t-1;\boldsymbol{\vartheta})\right)^2\right]$$

Recall that, asymptotically, the estimated parameters can be retrieved solving the formula:

$$\boldsymbol{\vartheta}^* = \arg\min_{\boldsymbol{\vartheta}}\{\bar{J}(\boldsymbol{\vartheta})\}$$

<u>ANALYSIS OF M_1</u>

Notice that the system S is a MA(1), while the class M_1 is that of AR(1) models. So:

$$S \notin M_1$$

So we can state that, asymptotically, the identified model provided by the PEM method will be the "best" approximant of the system S in the class of models M_1.

The optimal 1-step predictor of M_1 (which is a function of the unknown parameter a) is obviously:

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

Notice that, after we compute the value a^* of a which minimizes the loss function, we have to check whether the model $M_1(a^*)$ is expressed in a canonical form. In this case we have to verify:

$$|a^*| < 1$$

We apply the formula straightforwardly:

$$\begin{aligned}
a^* = \arg\min_a\{var[y(t) - \hat{y}(t|t-1;a)]\} &= \arg\min_a\{var[y(t) - ay(t-1)]\} \\
&= \arg\min_a\left\{E\left[\left(e(t) + \frac{1}{2}e(t-1) - ae(t-1) - \frac{1}{2}ae(t-2)\right)^2\right]\right\} \\
&= \arg\min_a\left\{E\left[\left(e(t) + \left(\frac{1}{2} - a\right)e(t-1) - \frac{1}{2}ae(t-2)\right)^2\right]\right\} \\
&= \arg\min_a\left\{1 + \left(\frac{1}{2} - a\right)^2 + \frac{1}{4}a^2\right\}
\end{aligned}$$

Notice that the variance of the 1-step prediction error:

$$var\left[\left(y(t)-\hat{y}(t|t-1)\right)\right]=1+\left(\frac{1}{2}-a\right)^2+\frac{1}{4}a^2$$

is a convex parabola (its second derivative is positive).

Thus, this function has just one minimum point, which is equal to the vertex of the parabola. We can easily compute it:

$$a^*:\ \left.\frac{\partial\, var[y(t)-\hat{y}(t|t-1;a)]}{\partial a}\right|_{a=a^*}=0$$

$$-2\left(\frac{1}{2}-a^*\right)+\frac{1}{2}a^*=0$$

$$2a^*+\frac{1}{2}a^*=1$$

$$a^*=\frac{2}{5}$$

Check whether the model is expressed in a canonical form:

$$|a^*|<1\quad ??\quad \textbf{YES}$$

Thus the model:

$$M_1(a^*): y(t)=\frac{2}{5}y(t-1)+\eta(t),\quad \eta(t)\sim WN(0,\lambda_1^2)$$

is the best approximant of the system S in the class of AR(1) models (notice that this is a S.S.P.).

Since $S\notin M_1$, the error $\varepsilon(t)$ is NOT A WHITE NOISE, indeed:

$$\begin{aligned}\varepsilon(t;\vartheta^*)=y(t)-\hat{y}(t|t-1;a^*)&=e(t)+\frac{1}{2}e(t-1)-\frac{2}{5}y(t-1)\\&=e(t)+\frac{1}{2}e(t-1)-\frac{2}{5}\left(e(t-1)+\frac{1}{2}e(t-2)\right)\\&=e(t)+\frac{1}{10}e(t-1)-\frac{1}{5}e(t-2)\end{aligned}$$

which is an MA(2) process. Notice that:

$$\eta(t) = y(t) - \hat{y}(t|t-1; a^*) = \varepsilon(t; a^*)$$

$$\begin{aligned}
\Rightarrow \lambda_1^2 &= var[\eta(t)] = var[\varepsilon(t; a^*)] = \bar{J}(a^*) \\
&= E\left[\left(e(t) + \frac{1}{10}e(t-1) - \frac{1}{5}e(t-2)\right)^2\right] = 1 + \frac{1}{100} + \frac{1}{25} \\
&= \frac{100 + 1 + 4}{100} = \frac{21}{20} > var\big(e(t)\big) = 1
\end{aligned}$$

Since the class of models is an AR(1), we could have computed the optimal value for *a* from this equation:

$$a^* = \frac{\gamma_y(1)}{\gamma_y(0)}$$

In fact:

$$\begin{aligned}
a^* &= \arg\min_a \left\{E\left[\big(y(t) - ay(t-1)\big)^2\right]\right\} \\
&= \arg\min_a \{E[y(t)^2 - 2ay(t)y(t-1) + a^2 y(t-1)^2\} \\
&= \arg\min_a \{\gamma_y(0) - 2a\gamma_y(1) + a^2\gamma_y(0)\}
\end{aligned}$$

So:

$$a^*: \frac{\partial\,(\gamma_y(0) - 2a\gamma_y(1) + a^2\gamma_y(0))}{\partial a} = 0$$

$$-2\gamma_y(1) + 2a^*\gamma_y(0) = 0$$

$$\Rightarrow \quad a^* = \frac{\gamma_y(1)}{\gamma_y(0)}$$

This formula is correct only for AR(1) models!!

If we apply it, we have:

$$\gamma_y(0) = E\left[\left(e(t) + \frac{1}{2}e(t-1)\right)^2\right] = 1 + \frac{1}{4} = \frac{5}{4}$$

$$\gamma_y(1) = E[y(t)y(t-1)] = \frac{1}{2}$$

$$\Rightarrow a^* = \frac{\gamma_y(1)}{\gamma_y(0)} = \frac{2}{5}$$

ANALYSIS OF M_2

Notice that the system S is a MA(1), while the system M_2 is an AR(2). So:

$$S \notin M_2$$

Again we can state that, asymptotically, the identified model provided by the PEM method will be the "best" approximant of the system S in the class of models M_2.

Moreover, notice that, in this case:

$$\boldsymbol{\vartheta} = \begin{bmatrix} a \\ b \end{bmatrix}$$

is a vector with dimensionality equal to 2.

The optimal 1-step predictor of M_2 (which is a function of the unknown parameters a and b) is obviously:

$$\hat{y}(t|t-1; \boldsymbol{\vartheta}) = -ay(t-1) - by(t-2)$$

We write the operatorial representation of M_2:

$$y(t) = \frac{1}{1 + az^{-1} + bz^{-2}} \eta(t), \quad \eta(t) \sim WN(0, \lambda_2^2)$$

Notice that, after we compute the value $\boldsymbol{\vartheta}^* = [a^* \;\; b^*]^T$ of $\boldsymbol{\vartheta}$ which minimizes the loss function, we have to check whether the model $M_2(\boldsymbol{\vartheta}^*)$ is expressed in a canonical form. In this case we have to verify that the poles of the transfer function are inside the unit circle:

$$z^2 + a^* z + b^* = 0 \quad \Rightarrow \quad z_{1,2} \qquad \textbf{CHECK } |z_{1,2}| < 1$$

We apply the formula straightforwardly:

$$\boldsymbol{\vartheta}^* = \begin{bmatrix} a^* \\ b^* \end{bmatrix} = \arg\min_{\boldsymbol{\vartheta}} \{var[y(t) - \hat{y}(t|t-1;\boldsymbol{\vartheta})]\}$$

$$= \arg\min_{a,b} \{var[y(t) + ay(t-1) + by(t-2)]\}$$

$$= \arg\min_{a,b} \left\{E\left[\left(y(t) + ay(t-1) + by(t-2)\right)^2\right]\right\}$$

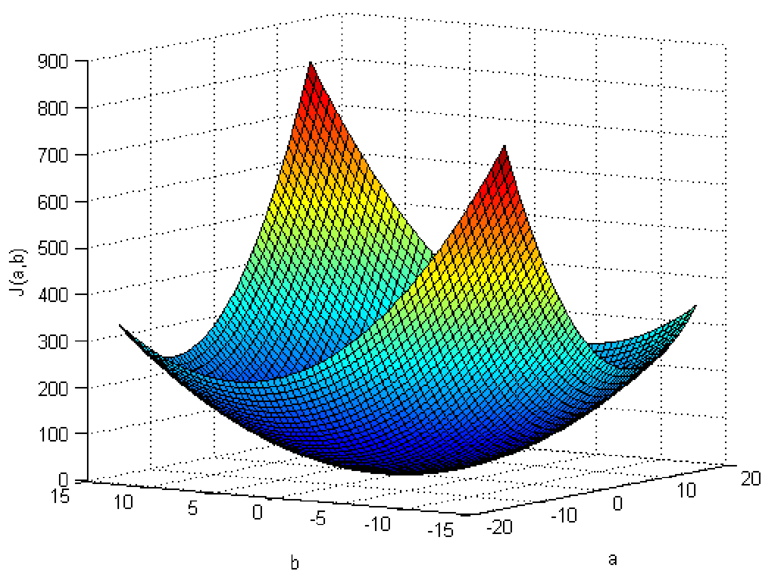
$$= \arg\min_{a,b} \left\{E\left[\left(e(t) + \frac{1}{2}e(t-1) + a\left(e(t-1) + \frac{1}{2}e(t-2)\right) + b\left(e(t-2) + \frac{1}{2}e(t-3)\right)\right)^2\right]\right\}$$

$$= \arg\min_{a,b} \left\{E\left[\left(e(t) + \left(\frac{1}{2} + a\right)e(t-1) + \left(\frac{1}{2}a + b\right)e(t-2) + \frac{1}{2}be(t-3)\right)^2\right]\right\}$$

$$= \arg\min_{a,b} \left\{1 + \frac{1}{4} + a^2 + a + \frac{1}{4}a^2 + b^2 + ab + \frac{1}{4}b^2\right\}$$

$$= \arg\min_{a,b} \left\{\frac{5}{4} + \frac{5}{4}a^2 + a + \frac{5}{4}b^2 + ab\right\}$$

The function we have to minimize with respect to *a*,*b* is a paraboloid:



This function has just one single minimum and can be easily computed.

First of all we define the gradient of $f(\boldsymbol{\vartheta}) = var[y(t) - \hat{y}(t|t-1;\boldsymbol{\vartheta})]$:

$$\nabla f(a,b) = \begin{bmatrix} \dfrac{\partial f(a,b)}{\partial a} \\ \dfrac{\partial f(a,b)}{\partial b} \end{bmatrix} = \begin{bmatrix} \dfrac{5}{2}a + 1 + b \\ \dfrac{5}{2}b + a \end{bmatrix}$$

The stationary point $\boldsymbol{\vartheta}^* = [a^* \;\; b^*]^T$ of $f(\boldsymbol{\vartheta})$ (i.e. $\boldsymbol{\vartheta}^* = \arg\min \bar{J}(\boldsymbol{\vartheta})$) can be computed as:

$$(a^*, b^*) \colon \nabla f(a^*, b^*) = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} \dfrac{5}{2}a^* + 1 + b^* \\ \dfrac{5}{2}b^* + a^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus, we have to solve a system with 2 equations and 2 unknowns (i.e. a^*,b^*):

$$\begin{cases} \dfrac{5}{2}a^* + 1 + b^* = 0 \\ \dfrac{5}{2}b^* + a^* = 0 \end{cases}$$

$$\begin{cases} \dfrac{5}{2}a^* + 1 + b^* = 0 \\ a^* = -\dfrac{5}{2}b^* \end{cases}$$

$$\begin{cases} -\dfrac{25}{4}b^* + 1 + b^* = 0 \\ a^* = -\dfrac{5}{2}b^* \end{cases}$$

$$\begin{cases} b^* = \dfrac{4}{21} \\ a^* = -\dfrac{10}{21} \end{cases}$$

Is $M_2(\boldsymbol{\vartheta}^*)$ expressed in a canonical form? Recall that we have to check whether this condition holds:

$$z^2 + a^* z + b^* = 0 \quad \Rightarrow \quad z_{1,2} \qquad \textbf{CHECK } |z_{1,2}| < 1$$

So:

$$z^2 - \frac{10}{21}z + \frac{4}{21} = 0$$

$$z_{1,2} = \frac{\frac{10}{21} \pm \sqrt{\frac{100}{441} - \frac{16}{21}}}{2} = \frac{5}{21} \pm j\sqrt{\frac{59}{441}}$$

We have two complex conjugate poles. So:

$$|z_1| = |z_2| = \sqrt{\frac{25}{441} + \frac{59}{441}} = \sqrt{\frac{84}{441}} < 1$$

So the poles are strictly inside the unit circle, $M_2(\boldsymbol{\vartheta}^*)$ is a SSP and it is expressed in the canonical form.

Again, notice that:

$$\eta(t) = \varepsilon(t; \boldsymbol{\vartheta}^*) = y(t) - \hat{y}(t|t-1; \boldsymbol{\vartheta}^*)$$

So:

$$\lambda_2^2 = var[\eta(t)] = var[\varepsilon(t; \boldsymbol{\vartheta}^*)] = \bar{J}(\boldsymbol{\vartheta}^*) = \frac{5}{4} + \frac{5}{4}{a^*}^2 + a^* + a^* b^* + \frac{5}{4}{b^*}^2$$
$$= \frac{595}{588} > var[e(t)] = 1$$

The variance of η is closer to the variance of e with respect to the model M_1: this is due to the fact that an MA(1) can be seen as an AR(∞); the greater is the m order of the AR model, the closer is the identified AR(m) model to the system S.

ANALYSIS OF M_3

Notice that the system S is a MA(1) and the system M_3 is a MA(1) too. So:

$$S \in M_3$$

By recalling the theory, we can conclude that, asymptotically:

$$\hat{a}_N \xrightarrow[N\to\infty]{} a^\circ = \frac{1}{2}$$

Thus the identified model is "asymptotically equivalent" to the system S. Thus, in this case, we have that:

$$a^* = \frac{1}{2}$$

However, if we do not recognize that $S \in M_3$, we can apply the definition:

$$\hat{y}(t|t-1;a) = \frac{az^{-1}}{1+az^{-1}} y(t)$$

$$\begin{aligned} a^* &= \arg\min_a \{var[y(t) - \hat{y}(t|t-1;a)]\} \\ &= \arg\min_a \left\{ var\left[y(t) - \frac{az^{-1}}{1+az^{-1}} y(t) \right] \right\} = \\ &= \arg\min_a \left\{ var\left[\frac{(1+az^{-1}-az^{-1})y(t)}{1+az^{-1}} \right] \right\} = \\ &= \arg\min_a \left\{ var\left[\frac{1+\frac{1}{2}z^{-1}}{1+az^{-1}} e(t) \right] \right\} \end{aligned}$$

Notice that, under the condition $|a^*| < 1$, the minimum variance we can obtain is the variance of the noise *e(t)*: by inspection, we can see that this is possible if and only if $a = \frac{1}{2}$, which leads to:

$$a^* = \frac{1}{2}$$

This way, the digital transfer function turns out to be 1:

$$\frac{1+\frac{1}{2}z^{-1}}{1+a^{*}z^{-1}}e(t)=1\,e(t)=e(t)$$

However, we can always proceed in the classical way:

$$\varepsilon(t;a)=\frac{1+\frac{1}{2}z^{-1}}{1+az^{-1}}e(t)$$

$$\varepsilon(t;a)=-a\varepsilon(t-1,a)+e(t)+\frac{1}{2}e(t-1)$$

So:

$$var[\varepsilon(t;a)]=\gamma_{\varepsilon}(0,a)=a^{2}\gamma_{\varepsilon}(0,a)+1+\frac{1}{4}-a$$

$$\gamma_{\varepsilon}(0,a)=\frac{\frac{5}{4}-a}{1-a^{2}}$$

Thus:

$$a^{*}=\arg\min_{a}\gamma_{\varepsilon}(0,a)$$

$$a^{*}:\left.\frac{\partial\gamma_{\varepsilon}(0,a)}{\partial a}\right|_{a=a^{*}}=0$$

So:

$$\frac{\partial\gamma_{\varepsilon}(0,a)}{\partial a}=\frac{-(1-a^{2})+2a\left(\frac{5}{4}-a\right)}{(1-a^{2})^{2}}=0$$

$$\Rightarrow\ a^{2}-\frac{5}{2}a+1=0$$

$$a_{1,2}=\frac{\frac{5}{2}\pm\sqrt{\frac{25}{4}-4}}{2}=\frac{\frac{5}{2}\pm\frac{3}{2}}{2}$$

$$a_1 = 2 \rightarrow \text{ MAXIMUM}$$

$$a_2 = \frac{1}{2} \rightarrow \text{ MINIMUM}$$

Note that these are just local extrema: in fact, $\gamma_\varepsilon(0, a_1) = \frac{1}{4}$ and $\gamma_\varepsilon(0, a_2) = 1$. Nevertheless, a_1 does not respect the condition on the canonical form of $M_3(a_1)$. Thus, only $a^* = a_2$ is acceptable, and it represents the minimum point of $\gamma_\varepsilon(0, a)$ for $|a| < 1$.

Obviously, in this case:

$$\eta(t) = \varepsilon(t; a^*) = y(t) - \hat{y}(t|t-1; a^*) = e(t)$$

$$\lambda_3^2 = var[\eta(t)] = var[e(t)] = 1$$

Exercise 2

Consider the system:

$$S\colon\ y(t) = 3e(t) + e(t-1), \quad e(t) \sim WN(0,1)$$

and the class of models:

$$M(b)\colon\ y(t) = \eta(t) + b\eta(t-1),\ \ \eta(t) \sim WN(0, \lambda^2)$$

1) Compute the asymptotical result of the PEM method for the identification of the parameter b.

2) Compute the value of λ for the identified model.

1)

First of all, notice that the system S is not a canonical representation of the process, since $C(z)$ is not monic:

$$y(t) = (3 + z^{-1})e(t)$$

$$C(z) = 3 + z^{-1}$$

$$A(z) = 1$$

The canonical representation can be derived as follows:

$$y(t) = (3 + z^{-1})\frac{3}{3}e(t) = \left(1 + \frac{1}{3}z^{-1}\right)3e(t)$$

Define the noise:

$$\tilde{e}(t) = 3e(t)$$

$$\tilde{e}(t) \sim WN(0,9)$$

So:

$$y(t) = \tilde{e}(t) + \frac{1}{3}\tilde{e}(t)$$

Notice that S is a MA(1) and M(b) is a MA(1) too. So:

$$S \in M(b)$$

In this case, the value b^* of the parameter b estimated using the PEM method (i.e. minimization of the 1-step prediction error variance) tends asymptotically (i.e. for $N \to \infty$) to the true value $b°$ of the system:

$$b^* = b° = \frac{1}{3}$$

2)

Obviously:

$$\lambda^2 = var[\eta(t)] = var[\varepsilon(t)] = var[y(t) - \hat{y}(t|t-1; b^*)] = var[\tilde{e}(t)] = 9$$

Exercise 3

Given the data generation mechanism:

$$S:\ y(t) = 0.25y(t-2) + 2e(t), \qquad e(t) \sim WN(0,1)$$

and the class of models:

$$M(a):\ y(t) = (0.5-a)y(t-1) + 0.5ay(t-2) + \eta(t), \qquad \eta(t) \sim WN(0,\lambda^2)$$

Compute the value of the parameter *a* which minimizes the loss function:

$$\bar{J}(a) = E\left[\left(y(t) - \hat{y}(t|t-1;a)\right)^2\right]$$

We assume that $N \to \infty$ (i.e. we have an infinite amount of data).

The AR(2) data generation mechanism is not a canonical representation of the process, since *C(z)* is not monic:

$$y(t) = \frac{2}{1-0.25z^{-2}}e(t)$$

$$C(z) = 2$$

$$A(z) = 1 - 0.25z^{-2}$$

Define the noise:

$$\tilde{e}(t) = 2e(t)$$

$$\tilde{e}(t) \sim WN(0,4)$$

Thus the canonical representation of S is:

$$y(t) = \frac{1}{1-0.25z^{-2}}\tilde{e}(t)$$

The model *M(a)* is an AR(2) too.

The one step predictor of M is:

$$\hat{y}(t|t-1;a) = (0.5-a)y(t-1) + 0.5a\, y(t-2)$$

So:

$$\begin{aligned}\varepsilon(t;a) &= y(t) - \hat{y}(t|t-1;a) = y(t) - (0.5-a)z^{-1}y(t) - 0.5az^{-2}y(t) \\ &= (1-(0.5-a)z^{-1} - 0.5az^{-2})y(t)\end{aligned}$$

But:

$$y(t) = \frac{1}{1-0.25z^{-2}}\tilde{e}(t)$$

If we substitute *y(t)* in the previous expression of $\varepsilon(t)$, we have:

$$\varepsilon(t;a) = \frac{1-(0.5-a)z^{-1} - 0.5az^{-2}}{1-0.25z^{-2}}\tilde{e}(t) = F(z;a)\tilde{e}(t)$$

Notice that the variance of $\varepsilon(t)$ is minimum when $\varepsilon(t) = \tilde{e}(t)$. This is true if the denominator and the numerator of $F(z;a)$ are equal to each other (i.e. $F(z;a) = 1$). This is possible if a satisfies the system of equations:

$$\begin{cases} -(0.5-a) = 0 \\ 0.5a = 0.25 \end{cases}$$

This system (2 equations and 1 unknown) has 1 solution, which is the value provided asymptotically by the PEM method for a:

$$a^* = 0.5$$

In this case:

$$\lambda^2 = var[\eta(t)] = var[\varepsilon(t;a^*)] = \bar{J}(a^*) = var[\tilde{e}(t)] = 4$$

Exercise 4

We want to identify the system S using the class of models (i.e. estimate the parameter a):

$$M(a):\ y(t) = ay(t-1) + e(t), \quad e(t) \sim WN(0, \lambda^2)$$

based on 4 measurements taken on the system S:

$$y(0) = 1$$

$$y(1) = 0$$

$$y(2) = -\frac{1}{2}$$

$$y(3) = \frac{1}{2}$$

$$y(4) = -1$$

Estimate $\hat{a}_4$ adopting the PEM method and then compute $\hat{y}(5|4)$.

In this case the general difference equation of the system S is not given. So we can not compute the asymptotical estimated value of *a*. But… we can compute the value of *a* based on the measurements up to *t=4*, i.e. $\hat{a}_4$.

Operatorial representation of the model M:

$$y(t) = \frac{1}{1 - az^{-1}} e(t)$$

The model *M(a)* is expressed in a canonical form if and only if $|a| < 1$.

The 1-step predictor of *M(a)* is:

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

Notice that the sample estimate of the mean is:

$$\widehat{m}_{y,N} = \frac{1}{5}\sum_{t=0}^{4} y(t) = \frac{1}{5}\left(1 + 0 - \frac{1}{2} + \frac{1}{2} - 1\right) = 0$$

Since the sample estimate of the mean is 0, we can use the formula:

$$\hat{a}_5 = \arg\min_a \left\{ \frac{1}{5} \sum_{t=0}^{4} \big(y(t) - \hat{y}(t|t-1)\big)^2 \right\}$$

if and only if we consider $y(-1) = m_y = 0$. However, it will be better to use:

$$\hat{a}_4 = \arg\min_a \left\{ \frac{1}{4} \sum_{t=1}^{4} \big(y(t) - \hat{y}(t|t-1)\big)^2 \right\}$$

So:

$$\hat{a}_4 = \arg\min_a \{J_4(a)\} = \arg\min_a \left\{ \frac{1}{4} \sum_{t=1}^{4} \big(y(t) - ay(t-1)\big)^2 \right\} =$$

$$= \arg\min_a \left\{ \frac{1}{4} \left[(0-a)^2 + \left(-\frac{1}{2} - 0 \cdot a\right)^2 + \left(\frac{1}{2} + \frac{1}{2}a\right)^2 + \left(-1 - \frac{1}{2}a\right)^2 \right] \right\} =$$

$$= \arg\min_a \left\{ \frac{1}{4} \left[a^2 + \frac{1}{4} + \frac{1}{4} + \frac{1}{2}a + \frac{1}{4}a^2 + 1 + a + \frac{1}{4}a^2 \right] \right\} =$$

$$\arg\min_a \left\{ \frac{1}{4} \left[\frac{3}{2}a^2 + \frac{3}{2}a + \frac{3}{2} \right] \right\}$$

The derivative of $J_4(a)$ with respect to a is:

$$\frac{\partial J_4(a)}{\partial a} = \frac{1}{4} \frac{\partial \left[\frac{3}{2}a^2 + \frac{3}{2}a + \frac{3}{2} \right]}{\partial a} = \frac{1}{4}\left(3a + \frac{3}{2}\right)$$

And so the estimate for the parameter a given 4 measurements of the system is:

$$\frac{1}{4}\left(3\hat{a}_4 + \frac{3}{2}\right) = 0$$

$$\Rightarrow \hat{a}_4 = -\frac{1}{2}$$

We can conclude that:

$$\hat{y}(t|t-1) = -\frac{1}{2}y(t-1)$$

Thus:

$$\hat{y}(5|4) = -\frac{1}{2}y(4) = \frac{1}{2}$$

Exercise 5

We want to identify the model (i.e. estimate the parameter a):

$$M(a):\ y(t) = \eta(t) + a\eta(t-1), \quad \eta(t) \sim WN(0, \lambda^2)$$

based on 3 measurements taken on the system S:

$$y(1) = 2$$

$$y(2) = 0$$

$$y(3) = 2$$

Compute $\hat{a}_3$.

The model *M(a)* is a canonical representation of the process for $|a| < 1$ (zero inside the unit circle).

1-step predictor of the model *M(a)*:

$$\hat{y}(t|t-1;a) = \frac{az^{-1}}{1+az^{-1}} y(t)$$

$$\hat{y}(t|t-1;a) = -a\hat{y}(t-1|t-2;a) + ay(t-1)$$

We fill the table:

Time Instant (t)	$y(t)$	$\hat{y}(t\|t-1)$
0	0	0
1	2	0
2	0	$2a$
3	2	$-2a^2$

So:

$$J_3 = \frac{1}{3}\sum_{t=1}^{3}\big(y(t) - \hat{y}(t|t-1)\big)^2 = \frac{1}{3}[(2-0)^2 + (0-2a)^2 + (2+2a^2)^2]$$

$$= \frac{1}{3}[4 + 4a^2 + 4 + 8a^2 + 4a^4] = \frac{4}{3}[a^4 + 3a^2 + 2]$$

The derivative of $J_3(a)$ with respect to a is:

$$\frac{\partial J_3(a)}{\partial a} = \frac{4}{3}\frac{\partial\ [a^4 + 3a^2 + 2]}{\partial a} = \frac{4}{3}(4a^3 + 6a)$$

And so the estimate for the parameter a given 3 measurements of the system is:

$$\frac{4}{3}(4\hat{a}_3^3 + 6\hat{a}_3) = 0$$

$$\hat{a}_3(2\hat{a}_3^2 + 3) = 0$$

Three solutions:

1) $\hat{a}_3 = 0$

2) $\hat{a}_3 = j\sqrt{\frac{3}{2}}$ → Not feasible ($a \in \mathbf{R}$)

3) $\hat{a}_3 = -j\sqrt{\frac{3}{2}}$ → Not feasible ($a \in \mathbf{R}$)

So the feasible solution is $\hat{a}_3 = 0$. Is this a minimum?

Study of the SECOND DERIVATIVE of $J_3(a)$:

$$\frac{\partial^2 J_3(a)}{\partial a^2} = \frac{4}{3}(12a^2 + 6)$$

When $a = \hat{a}_3 = 0$:

$$\frac{4}{3}(12\hat{a}_3^2 + 6) = \frac{24}{3} > 0 \quad \Rightarrow \quad \text{MINIMUM POINT}$$

Exercise 6

Suppose we have 10 measurements taken from a system S:

$$S: \{y(1), y(2), \dots, y(10)\}$$

We use these data to identify the system S adopting the class of models:

$$M(\boldsymbol{\vartheta}): y(t) = a_1 y(t-1) + a_2 y(t-2) + e(t), \quad e(t) \sim WN(0, \lambda^2)$$

Write the formula used by the PEM method to estimate the parameters a_1 and a_2 (the notation used for the estimated value of these parameters is $\hat{a}_{1,10}$, $\hat{a}_{2,10}$).

The 1-step predictor of $M(\boldsymbol{\vartheta})$ is:

$$\hat{y}(t|t-1) = a_1 y(t-1) + a_2 y(t-2)$$

The vector of parameters is:

$$\boldsymbol{\vartheta} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

The PEM method minimizes the loss function:

$$J_N(\boldsymbol{\vartheta}) = \frac{1}{N-h} \sum_{t=1+h}^{N} \big(y(t) - \hat{y}(t|t-1; \boldsymbol{\vartheta})\big)^2$$

with respect to $\boldsymbol{\vartheta}$. N is the number of measurements and $h \triangleq \max(m, p+d)$, where m and p are the order of the AR part and X part of the ARX model $M(\boldsymbol{\vartheta})$ respectively, and d is the delay of the X part. This means that:

$$\hat{\boldsymbol{\vartheta}}_N = \arg\min_{\boldsymbol{\vartheta}} J_N$$

In this case, $M(\boldsymbol{\vartheta})$ is an AR(2) (i.e. ARX(2,0,0)) and we have 10 measurements, so:

$$N = 10$$
$$h = \max(2,0) = 2$$

Thus:

$$J_{10}(a_1, a_2) = \frac{1}{8}\sum_{t=3}^{10}\big(y(t) - a_1 y(t-1) - a_2 y(t-2)\big)^2$$

and:

$$\begin{bmatrix}\hat{a}_1 \\ \hat{a}_2\end{bmatrix} = \arg\min_{a_1, a_2} J_{10}(a_1, a_2)$$

This means that the estimate of the parameters a_1, a_2 solves the equation:

$$\left.\frac{\partial J_{10}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|_{\widehat{\boldsymbol{\vartheta}}} = \mathbf{0}$$

Notice that:

$$\frac{\partial J_{10}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \boldsymbol{\nabla} J_{10}(\boldsymbol{\vartheta})$$

is the gradient of $J_{10}(\boldsymbol{\vartheta})$. In this case this is a vector:

$$\frac{\partial J_{10}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \begin{bmatrix}\dfrac{\partial J_{10}(a_1, a_2)}{\partial a_1} \\ \dfrac{\partial J_{10}(a_1, a_2)}{\partial a_2}\end{bmatrix}$$

Thus the equation:

$$\boldsymbol{\nabla} J_{10}(\widehat{\boldsymbol{\vartheta}}) = \mathbf{0}$$

is actually a system of 2 equations with 2 unknowns:

$$\begin{bmatrix}\dfrac{\partial J_{10}(a_1, a_2)}{\partial a_1} \\ \dfrac{\partial J_{10}(a_1, a_2)}{\partial a_2}\end{bmatrix}_{\widehat{\boldsymbol{\vartheta}}} = \begin{bmatrix}0 \\ 0\end{bmatrix}$$

We compute separately the two elements of $\nabla J_{10}(\boldsymbol{\vartheta})$:

$$\frac{\partial J_{10}(a_1,a_2)}{\partial a_1} = -\frac{2}{8}\sum_{t=3}^{10}\left[\big(y(t)-a_1y(t-1)-a_2y(t-2)\big)y(t-1)\right]$$

$$\frac{\partial J_{10}(a_1,a_2)}{\partial a_2} = -\frac{2}{8}\sum_{t=3}^{10}\left[\big(y(t)-a_1y(t-1)-a_2y(t-2)\big)y(t-2)\right]$$

So:

$$\begin{bmatrix}\sum_{t=3}^{10}\left[\big(y(t)-\hat{a}_1y(t-1)-\hat{a}_2y(t-2)\big)y(t-1)\right]\\ \sum_{t=3}^{10}\left[\big(y(t)-\hat{a}_1y(t-1)-\hat{a}_2y(t-2)\big)y(t-2)\right]\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \quad (*)$$

Notice that the system of equations (*) is linear with respect to the parameters a_1, a_2, and can be expressed using the 2x2 matrix K and the 2x1 vector *b*:

$$\boldsymbol{K}\widehat{\boldsymbol{\vartheta}} + \boldsymbol{b} = \boldsymbol{0} \quad \Rightarrow \quad \begin{bmatrix}k_{11} & k_{12}\\ k_{21} & k_{22}\end{bmatrix}\begin{bmatrix}\hat{a}_1\\ \hat{a}_2\end{bmatrix} = -\begin{bmatrix}b_1\\ b_2\end{bmatrix}$$

The first row of the matrix **K** and of the vector ***b*** refers to the first equation of (*), while the second row is relative to the second equation of (*). It is important to stress the fact that the matrix **K** and the vector ***b*** are totally known, since they depend on measurements.

Then, the estimates of the parameters a_1, a_2 (i.e. $\widehat{\boldsymbol{\vartheta}} = [\hat{a}_1 \;\; \hat{a}_2]^T$) can be obtained by:

$$\widehat{\boldsymbol{\vartheta}} = -\boldsymbol{K}^{-1}\boldsymbol{b}$$

The terms $k_{i,j}$ and $b_{1,2}$ can be easily computed:

FIRST EQUATION OF (*)

$$k_{11} = -\sum_{t=3}^{10} y(t-1)^2$$

$$k_{12} = -\sum_{t=3}^{10} y(t-1)y(t-2)$$

$$b_1 = \sum_{t=3}^{10} y(t)y(t-1)$$

SECOND EQUATION OF (*)

$$k_{21} = -\sum_{t=3}^{10} y(t-1)y(t-2)$$

$$k_{22} = -\sum_{t=3}^{10} y(t-2)^2$$

$$b_2 = \sum_{t=3}^{10} y(t)y(t-2)$$

The matrix **K** is symmetric, since $k_{12} = k_{21}$.

Notice that the matrix **K** and the vector ***b*** can be written as:

$$\boldsymbol{K} = -\boldsymbol{\Phi}^T\boldsymbol{\Phi}$$

$$\boldsymbol{b} = \boldsymbol{\Phi}^T\boldsymbol{y}$$

where:

$$\boldsymbol{\Phi} = \begin{bmatrix} y(2) & y(1) \\ \vdots & \vdots \\ y(9) & y(8) \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y(3) \\ \vdots \\ y(10) \end{bmatrix}$$

So the matrix **K** has the structure we have studied in theory: the four blocks (in this case the four elements) which span the rows and the columns are the sample estimates of the autocorrelation and the cross-correlation. Since, in this particular case, there are no exogenous inputs and the class of models is the AR(2), these elements are the sample estimate of the output variance and the sample estimate of the covariance function for $\tau = 1$.

Finally:

$$\hat{\boldsymbol{\vartheta}} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = -\boldsymbol{K}^{-1}\boldsymbol{b}$$

$$= \begin{bmatrix} \sum_{t=3}^{10} y(t-1)^2 & \sum_{t=3}^{10} y(t-1)y(t-2) \\ \sum_{t=3}^{10} y(t-1)y(t-2) & \sum_{t=3}^{10} y(t-2)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=3}^{10} y(t)y(t-1) \\ \sum_{t=3}^{10} y(t)y(t-2) \end{bmatrix}$$

Exercise 7

Consider a periodic system whose output is:

$$S:\ \begin{cases} y(i) = 1, & \text{for } i \text{ even} \\ y(i) = -1, & \text{for } i \text{ odd} \end{cases}$$

Suppose we want to identify this system using the class of models:

$$M(a):\ y(t) = ay(t-1) + \frac{1}{2}y(t-2) + \eta(t), \quad \eta(t) \sim WN(0, \lambda^2)$$

based on 10 measurements taken from the system.

Compute $\hat{a}_{10}$.

1-step predictor of *M(a)*:

$$\hat{y}(t|t-1) = ay(t-1) + \frac{1}{2}y(t-2)$$

Notice that:

$$\varepsilon(t; a) = y(t) - \hat{y}(t|t-1) = y(t) - ay(t-1) - \frac{1}{2}y(t-2)$$

When:

- *t* even:

$$\varepsilon(t; a) = 1 + a - \frac{1}{2} = a + \frac{1}{2}$$

- *t* odd:

$$\varepsilon(t; a) = -1 - a + \frac{1}{2} = -\left(a + \frac{1}{2}\right)$$

Recall that:

$$J_{10}(a) = \frac{1}{10}\sum_{t=1}^{10} \varepsilon(t; a)^2$$

But:

$$\varepsilon(t;a)^2 = \left(a + \frac{1}{2}\right)^2 = \varepsilon(a)^2 \qquad \forall t$$

In this particular case, the variance of the prediction error does not depend on *t*, so we have:

$$J_{10}(a) = \frac{1}{10}\sum_{t=1}^{10}\varepsilon^2(a) = \frac{10}{10}\varepsilon^2(a) = \left(a + \frac{1}{2}\right)^2$$

Then we take the derivative of the loss function with respect to the unknown parameter:

$$\frac{\partial J_{10}}{\partial a} = 2\left(a + \frac{1}{2}\right)$$

Thus:

$$\hat{a}_{10} + \frac{1}{2} = 0$$

$$\Rightarrow \quad \hat{a}_{10} = -\frac{1}{2}$$

Exercise 8

Consider system:

$$S1:\ y(t) = e(t), \quad e(t) \sim WN(0,1)$$

and the class of models:

$$M(a):\ y(t) = -ay(t-1) + \eta(t) + \frac{1}{2}\eta(t-1), \quad \eta(t) \sim WN(0, \lambda^2)$$

Compute the value of a which minimizes the loss function:

$$\bar{J}(a) = E\left[\left(y(t) - \hat{y}(t|t-1;a)\right)^2\right]$$

What happens if the system S is as follows?

$$S2: y(t) = e(t) + 2e(t-1), \quad e(t) \sim WN(0,1)$$

The 1-step predictor of the ARMA model *M(a)* can be simply evaluated.

In fact:

$$C(z) = 1 + \frac{1}{2}z^{-1}$$

$$A(z;a) = 1 + az^{-1}$$

So:

$$\hat{y}(t|t-1;a) = \frac{C(z) - A(z;a)}{C(z)}y(t) = \frac{1 + \frac{1}{2}z^{-1} - 1 - az^{-1}}{1 + \frac{1}{2}z^{-1}}y(t)$$

$$= \frac{\left(\frac{1}{2} - a\right)z^{-1}}{1 + \frac{1}{2}z^{-1}}y(t)$$

The prediction error is:

$$\varepsilon(t;a) = y(t) - \hat{y}(t|t-1;a) = y(t) - \frac{\left(\frac{1}{2} - a\right)z^{-1}}{1 + \frac{1}{2}z^{-1}}y(t)$$

$$= \frac{1 + \frac{1}{2}z^{-1} - \frac{1}{2}z^{-1} + az^{-1}}{1 + \frac{1}{2}z^{-1}}y(t) = \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}}y(t)$$

But:

$$y(t) = e(t)$$

Thus:

$$\varepsilon(t;a) = \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}}e(t)$$

Notice that the variance of $\varepsilon(t;a)$ is minimum when the value a^* of a is such that $\varepsilon(t;a^*) = e(t)$. By inspecting the previous formula, one can see that:

$$a = a^* = \frac{1}{2}$$

is such that:

$$\varepsilon(t;a^*) = \frac{1 + \frac{1}{2}z^{-1}}{1 + \frac{1}{2}z^{-1}}e(t) = e(t)$$

So the value of a which minimizes the loss function $\bar{J}(a)$ is:

$$a = a^* = \frac{1}{2}$$

However, if the system S is described by the difference equation:

$$S2: y(t) = e(t) + 2e(t-1), \quad e(t) \sim WN(0,1)$$

everything changes.

In fact:

$$\begin{aligned}\varepsilon(t;a) &= \frac{1+az^{-1}}{1+\frac{1}{2}z^{-1}}y(t) = \frac{1+az^{-1}}{1+\frac{1}{2}z^{-1}}(1+2z^{-1})e(t) \\ &= (1+az^{-1})\,2\,\frac{1}{2}\frac{1+2z^{-1}}{1+\frac{1}{2}z^{-1}}e(t) = (1+az^{-1})2e(t) \\ &= 2e(t) + 2ae(t-1)\end{aligned}$$

So the value of a which minimizes the loss function $\bar{J}(a)$ is:

$$a = a^* = 0$$