ID3

ARX Identification

3.16 EXAMPLE 3.2

The ARX process considered now has been generated using the model

$$y(t) = 0.5\,y(t-1) - 0.08\,y(t-2) - 0.096\,y(t-3) + 0.3413\,u(t-1) + 0.0683\,u(t-2) - 0.1638\,u(t-3) + e(t). \tag{3.16.1a}$$

for $t \leq 150$ and the model

$$y(t) = 0.4\,y(t-1) + 0.2\,y(t-2) - 0.048\,y(t-3) - 0.3413\,u(t-1) - 0.3413\,u(t-2) - 0.0307\,u(t-3) + e(t). \tag{3.16.1b}$$

for $t > 150$. The whole lenght of the available input–output sequences is $L = 320$. The input sequence, reported in Figure 3.16.1, has null mean value and variance, computed, as subsequent ones, on 300 samples, $\sigma_u^2 = 1$.

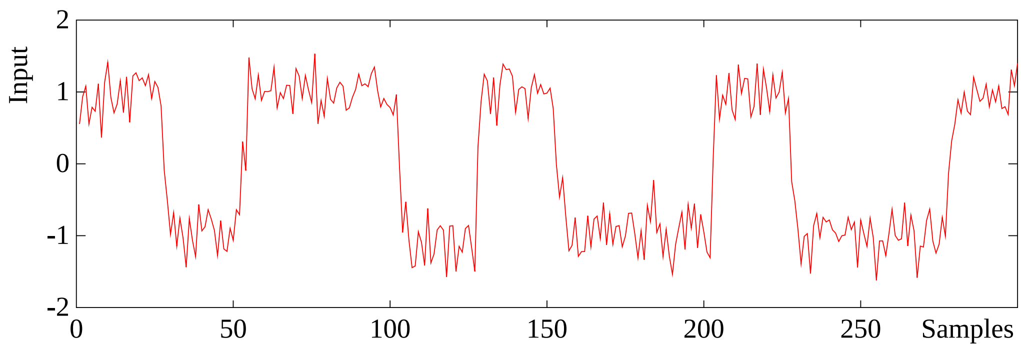


Figure 3.16.1 – Input sequence of process (3.16.1)

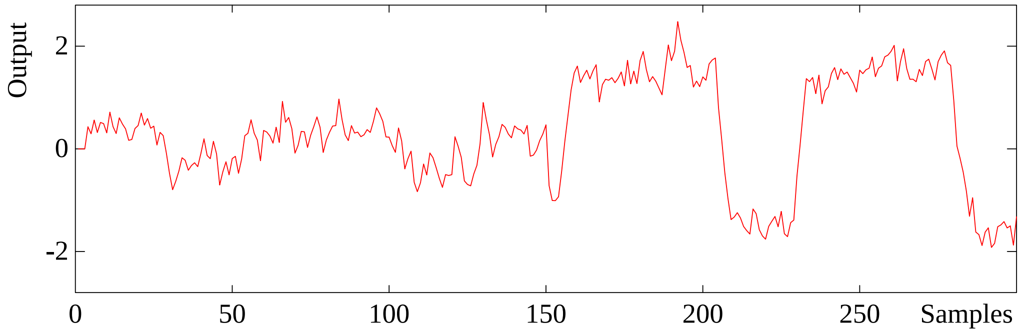


Figure 3.16.2 – Output sequence of process (3.16.1)

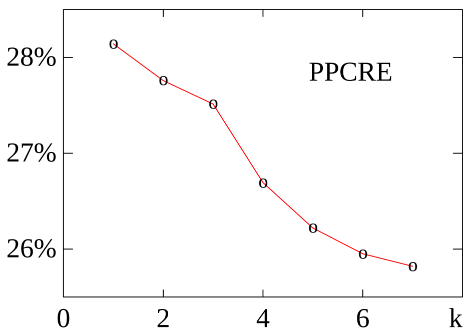
The variance of the output sequence (Figure 3.16.2) is, in absence of noise ($e(t) = 0$), $\sigma^2_{y*} = 1$; $e(\cdot)$ is a stationary and Gaussian process with null expected value and variance $\sigma^2_e = 0.04$. With reference to the decomposition of Figure 3.1.2, the colored noise $v(t)$ has a variance, computed, as previous ones, on the first 300 samples, given by $\sigma^2_v = 0.0509$; these values can be interpreted as the presence, on the data, of an amount of noise of

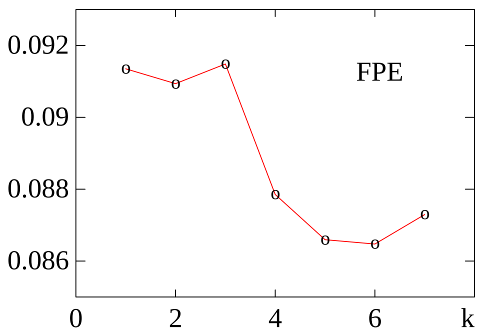
$$100\,\frac{\sigma_v}{\sigma_{y*}} = 22.6\,\%.$$

The data of this process will be used to describe some of the problems that can be met in modelling non stationary processes and to compare the results obtainable with on–line identification algorithms.

3.16.1 Determination of the model order

Computing the values of the PPCRE for $N = 300$ and $k = 1, \ldots, 7$ we obtain the results reported in Figure 3.16.3. This plot does not give any useful indication on the choice of a suitable model order; the only indication concerns a modest stabilization corresponding to $n = 6$.

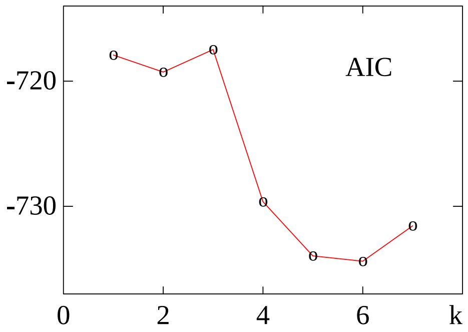


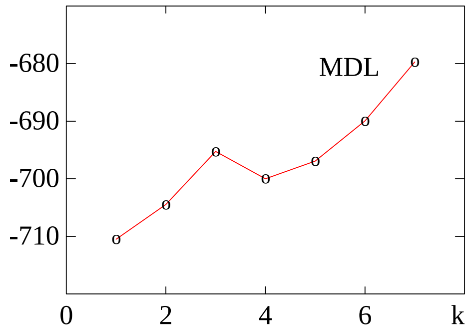


Figures 3.16.3 and 3.16.4 – PPCRE and FPE criteria for $N = 300$

FPE and AIC criteria, reported in Figures 3.16.4 and 3.16.5, show similar behaviors indicating 6 as most suitable order and 5 as acceptable. The MDL criterion (Figure

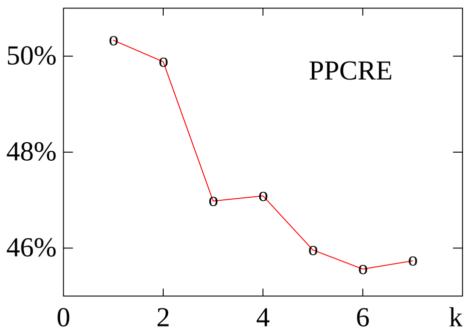
3.16.6) does not give, in this case useful indications having its absolute minimum at $k = 1$ and a local minimum at $k = 4$.

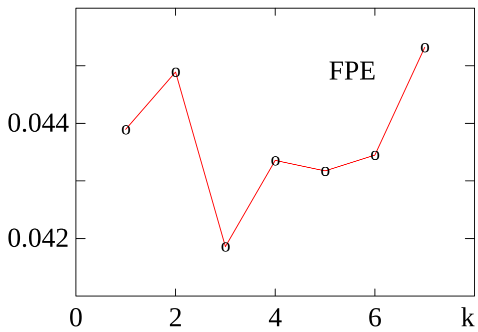




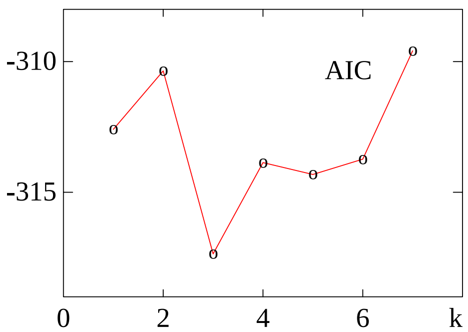
Figures 3.16.5 and 3.16.6 – AIC and MDL criteria for $N = 300$

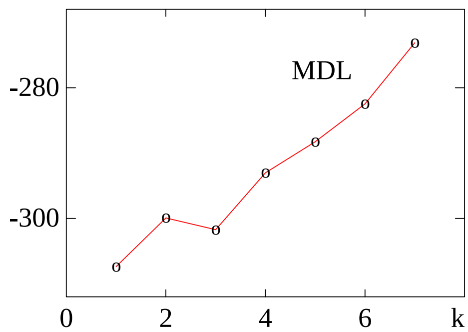
The analysis of the results given by previous tests could lead to select a model with order 6 and this indication cannot be considered completely wrong since, in fact, the sequences have been generated using two models with order 3 without common poles; they are thus influenced by six different time constants. The discordances between the tests and the lack of a real stabilization in the PPCRE indicate, however, problems in fitting the data with models belonging to the considered class.





Figures 3.16.7 and 3.16.8 – PPCRE and FPE criteria for $N = 100$





Figures 3.16.9 and 3.16.10 – AIC and MDL criteria for $N = 100$

In situations of this kind it can be advisable to test a possible time dependence of the process performing tests on reduced–length sequences. In our case the tests have been applied to the first 100 samples obtaining the results reported in Figures 3.16.7, 3.16.8, 3.16.9 and 3.16.10. The PPCRE criterion shows that the prediction error increases passing from a model with order 3 to an order 4 one. Also FPE and AIC criteria indicate 3 as optimal order for the model while MDL has only a local minimum for $k = 3$. The conclusion of previous analysis is that the process exhibits a time dependent (or, possibly, nonlinear) behavior and that it cannot be properly described by means of a single stationary (or linear) model.

It can look surprising, comparing Figures 3.16.3 and 3.16.7 to observe PPCRE values that are higher on the first 100 samples than on 300 samples, while computing the variance of the equation error with (3.14.16) we would find higher values for $N = 300$ (approximately double than for $N = 100$). The explanation is linked to the definition of the PPCRE as ratio between the standard deviations of the prediction error and of the output; in this example the standard deviation of the first 100 output samples is considerably lower than that of the whole sequence as can be observed in Figure 3.16.2.

3.16.2 Parameter estimate

For a better evaluation of the opportunity of using, in this case, on–line identification techniques, a first set of parameters has been estimated from the whole sequence ($N = 300$) using an order 3 model. The obtained parameters are

$$\begin{aligned}
\alpha_1 &= -0.0713 \\
\alpha_2 &= 0.0188 \\
\alpha_3 &= 0.9826 \\
\beta_1 &= 0.0355 \\
\beta_2 &= -0.1243 \\
\beta_3 &= 0.0214.
\end{aligned}$$

The one–step–ahead prevision given by this model is compared in Figure 3.16.11 (black line) with the observed output. The residuals are plotted in Figure 3.16.12.

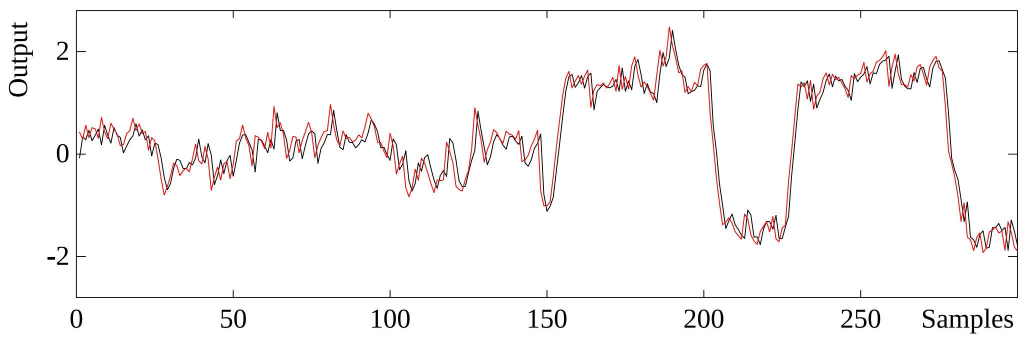


Figure 3.16.11 – Model prevision (black line) and observed output

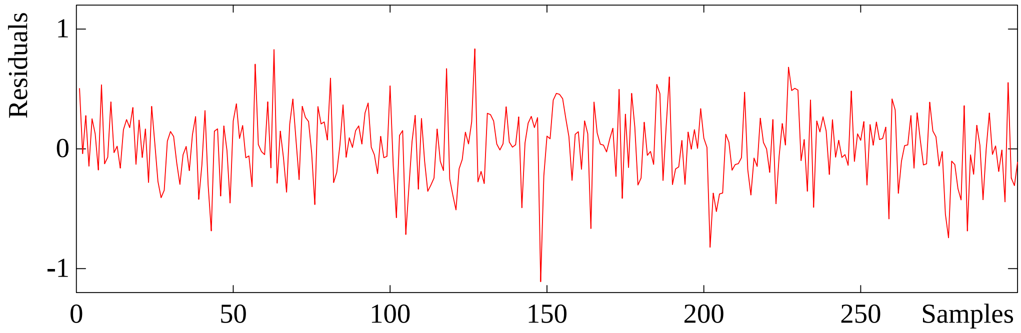


Figure 3.16.12 – Residuals of a single order 3 model

The mean square prevision error $J(\theta^\circ_{300})$ of this model is

$$J(\theta^\circ_{300}) = 0.0879$$

and the corresponding estimate of the variance of $e(t)$ given by (3.10.5) is $\hat{\sigma}_e^2 = 0.0897$. Using this value we obtain, by means of (3.8.5), the following covariance matrix

$$\operatorname{cov}\theta^\circ_{300} = 10^{-3}\begin{bmatrix} 3.261 & -3.286 & 0.164 & 0.123 & -0.103 & -0.068 \\ -3.286 & 6.569 & -3.306 & -0.283 & 0.287 & -0.029 \\ 0.164 & -3.306 & 3.380 & 0.325 & -0.152 & 0.076 \\ 0.123 & -0.283 & 0.325 & 2.008 & -1.244 & -0.550 \\ -0.103 & 0.287 & -0.152 & -1.244 & 2.588 & -1.255 \\ -0.068 & -0.029 & 0.076 & -0.550 & -1.255 & 1.921 \end{bmatrix}.$$

The standard deviations of the single parameters are thus,

$$\begin{aligned} \operatorname{std}\alpha_1 &= 0.057 & \qquad \operatorname{std}\beta_1 &= 0.045 \\ \operatorname{std}\alpha_2 &= 0.081 & \qquad \operatorname{std}\beta_2 &= 0.051 \\ \operatorname{std}\alpha_3 &= 0.058 & \qquad \operatorname{std}\beta_3 &= 0.044 \end{aligned}$$

and with the exception of α_1, are somehow high with respect to the parameter values. The unsatisfactory behavior of the model is highlighted by the plot of its previsions; in fact, looking at Figure 3.16.11, we see that the prevision reproduces the observed output well but with a delay equal to one sampling interval ($\alpha_1 \simeq 1$); this prevision is thus useless.

3.16.3 On–line least squares identification

A first on–line identification has been performed using recursive weighted least squares with a forgetting factor $\kappa = 0.97$. The obtained one–step–ahead prevision (black line) is compared in Figure 3.16.13 with the observed output. The corresponding residuals are reported in Figure 3.16.14.

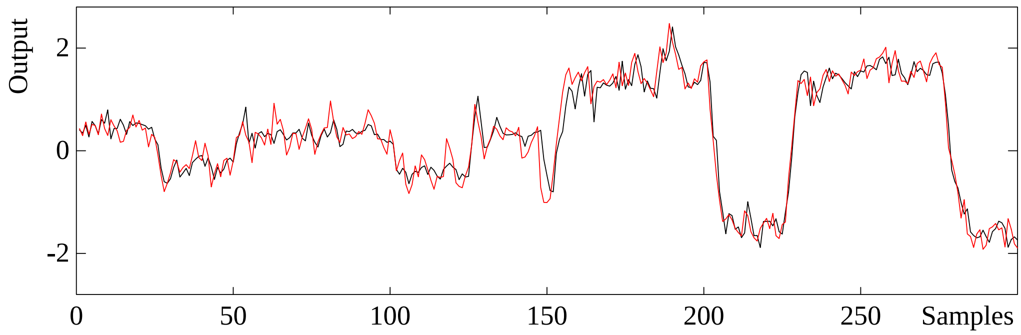


Figure 3.16.13 – Model prevision (black line) and observed output

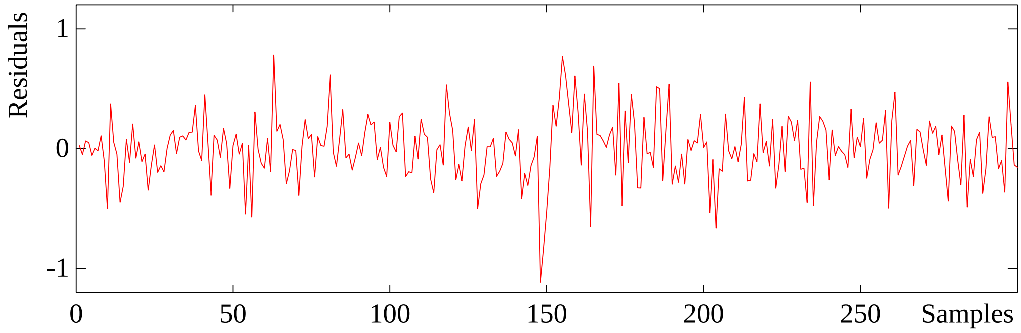
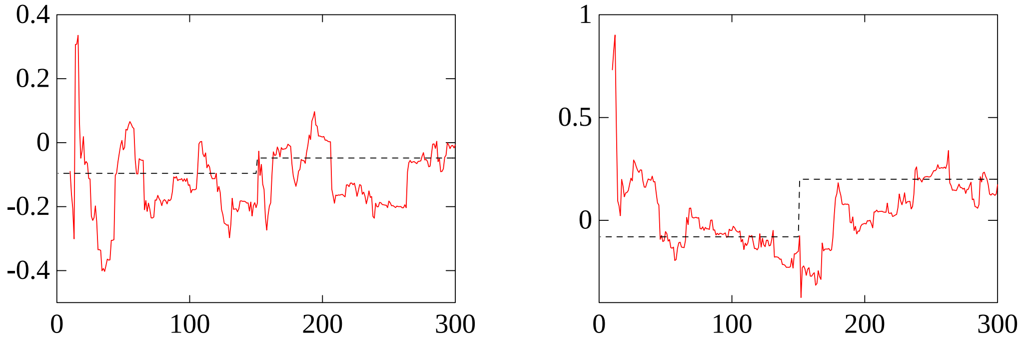


Figure 3.16.14 – Residuals of weighted least squares on–line identification

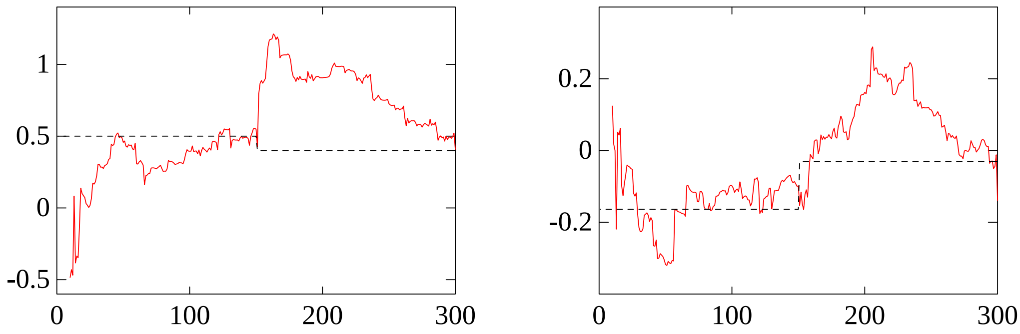
We can observe an increase in the prediction error in correspondence with the switch of the process from the first model to the second one at $t = 150$ and the subsequent adaptive behavior of the model. It can be of some interest to plot the values of the parameters at every identification step; Figures 3.16.15 – 3.16.20 report the estimated values of α_1, α_2, α_3, β_1, β_2 and β_3 against true ones (dashed line). The mean square prevision error is

$$J(\theta^\circ_{300}) = 0.0685,$$

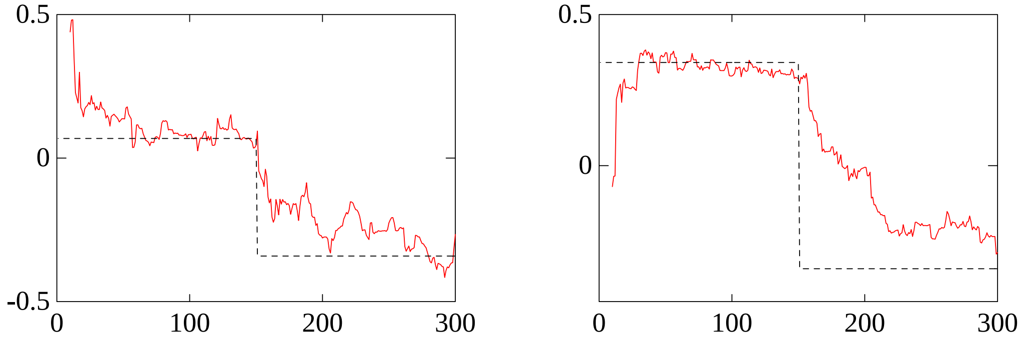
lower, as expected, than the value associated with a single model (0.0879).



Figures 3.16.15 and 3.16.16 – Weighted least squares: estimates of α_1 and α_2



Figures 3.16.17 and 3.16.18 – Weighted least squares: estimates of α_3 and β_1



Figures 3.16.19 and 3.16.20 – Weighted least squares: estimates of β_2 and β_3

3.16.4 On–line identification by Kalman filtering

A second on–line identification has been performed using a Kalman filter, using the procedure described in module ID3.13. The initial parameter estimate inserted in the algorithm has been $\theta(3) = [0\ 0\ 0\ 0\ 0\ 0]^T$, the covariance of the initial estimate has been set at $P(3) = 0.1\, I$ and the covariance matrix of the noise on the filter state has been set equal to $Q = 0.0005\, I$.

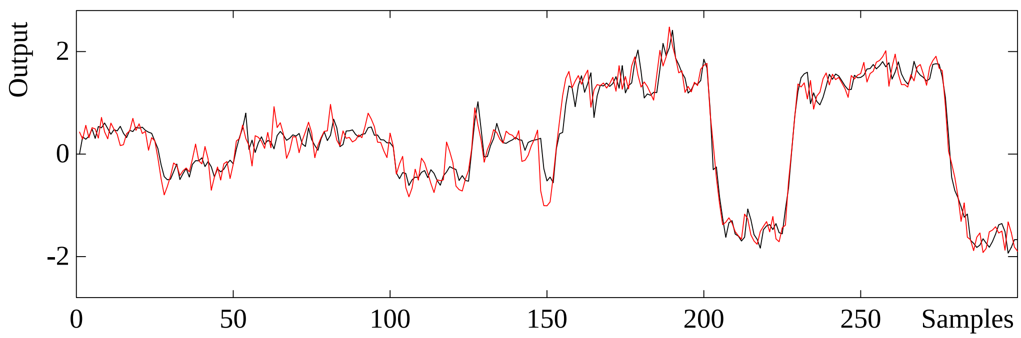


Figure 3.16.21 – Model prevision (black line) and observed output

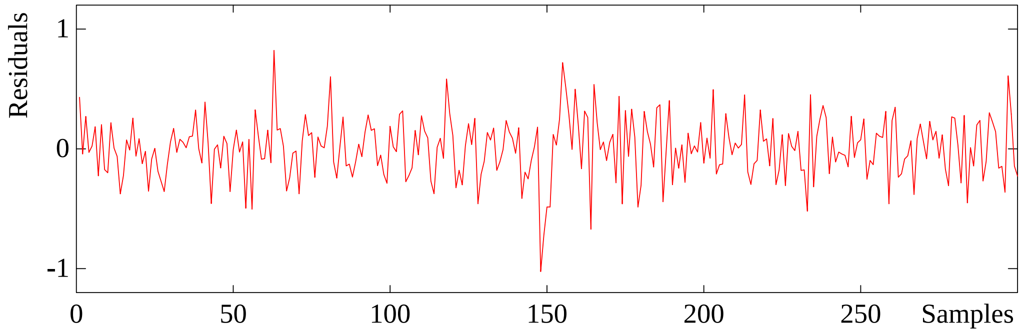
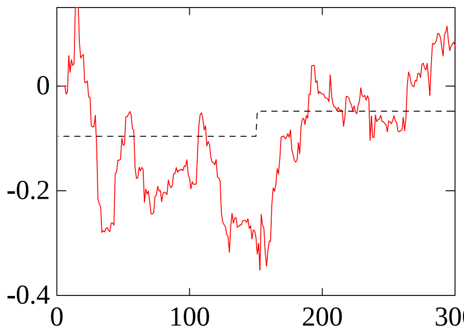
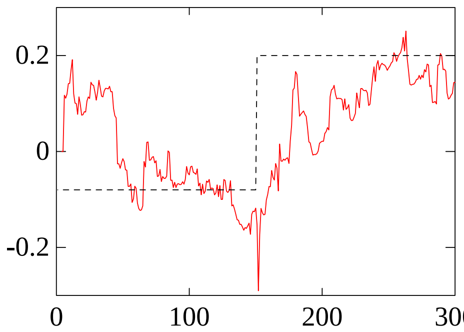


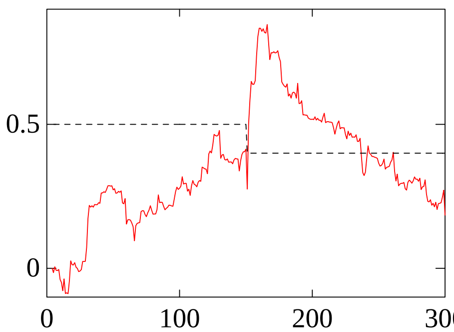
Figure 3.16.22 – Residuals of on–line identification performed by Kalman filtering

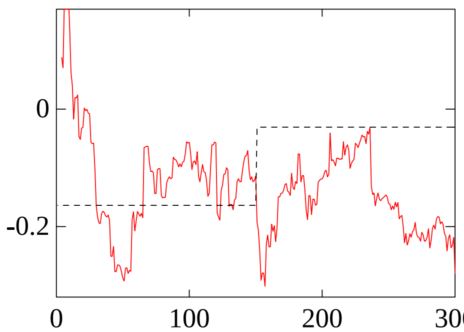
The variance of the filter observation noise has been obtained, by means of (3.14.16), from the first 100 samples and is given by $\hat{\sigma}_e^2 = 0.0395$. The one–step–ahead prevision given by the filter is reported (black line) in Figure 3.16.21 where it is compared with the observed output. The corresponding residuals are reported in Figure 3.16.22. It is interesting to note, by comparing Figure 3.16.22 with Figure 3.16.14, that, the residuals obtained using Kalman filtering are quite similar to those given by weighted least squares; in both cases the residuals constitute a good approximation of the equation error $e(t)$.



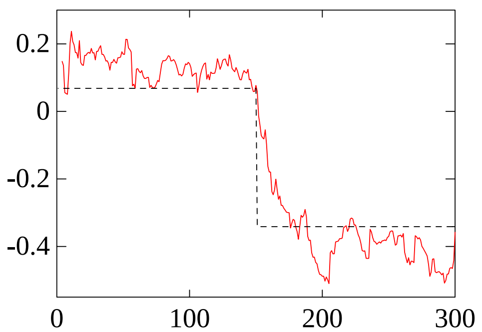


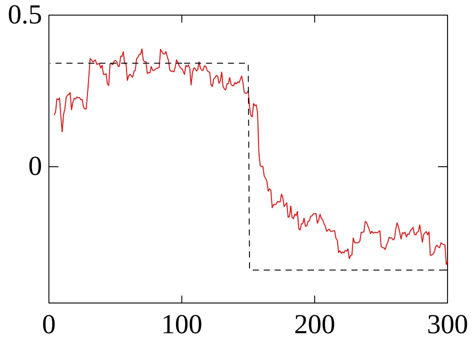
Figures 3.16.23 and 3.16.24 – Kalman filtering identification: estimates of α_1 and α_2





Figures 3.16.25 and 3.16.26 – Kalman filtering identification: estimates of α_3 and β_1





Figures 3.16.27 and 3.16.28 – Kalman filtering identification: estimates of β_2 and β_3

Also in this case it is possible to observe how the prediction error increases in correspondence with the commutation of the process from the first model to the second and how the filter rapidly adapts to this change. The mean square prevision error is given by

$$J(\theta^\circ_{300}) = 0.0619,$$

better than the value obtained with weighted least squares. The estimates of α_1, α_2, α_3, β_1, β_2 and β_3 at every step are reported in Figures 3.16.23 – 3.16.28 where they are compared with actual values (dashed lines).