Chapter 20

ARX

ARX model name stands for **Auto-Regressive** with e**X**tra input or **Auto-Regressive** e**X**ogenic.

Systems classification Two class of models:

- **Endogenic** system is a system without inputs.
- **Exogenic** is a system with inputs.

Goal: Extension for AR model to ARX model.

The $ARX(p,q)$ model is given by

$$\begin{aligned} y[n] &= a_1 y[n-1] + \cdots + a_p y[n-p] \\ &\quad + b_1 x[n-1] + \cdots + b_k x[n-k] + \epsilon[n] \end{aligned} \tag{20.1}$$

20.1 Cross-Correlation Function

Goal: Analogous to ACF, for two different signals.

ARX(0,1) Model The goal is to predict $y[n]$ from $x[n-k]$,

$$\hat{y}[n] = b_k x[n-k], \tag{20.2}$$

The resulting MSE-based loss function is of the form

$$\mathcal{L}(b) = \frac{1}{2} \sum_n \left(y[n] - b_k x[n-k] \right) \Big)^2 \tag{20.3}$$

with the solution by

$$\frac{d\mathcal{L}(b)}{db} = \sum_n (y[n] - b_k x[n-k])(-x[n-k]) = 0 \tag{20.4}$$

The corresponding solution is

$$b_k = \frac{\sum_n y[n]x[n-k]}{\sum_n x^2[n-k]}. \tag{20.5}$$

Cross-Correlation Function The resulting coefficients are related to the cross-correlation function,

$$R_{\mathbf{xy}}[k] = \sum_n x[n]y[n-k], k = -L+1, \ldots, L-1 \tag{20.6}$$

Similar to the ACF, it exists in three additional modifications: biased, unbiased and normalized,

$$R_{\mathbf{xy},biased}[k] = \frac{1}{L} R_{\mathbf{xy}}[k] \tag{20.7}$$

$$R_{\mathbf{xy},unbiased}[k] = \frac{1}{L-|k|} R_{\mathbf{xy}}[k] \tag{20.8}$$

$$R_{\mathbf{xy},norm}[k] = \frac{R_{\mathbf{xy}}[k]}{\sqrt{R_{\mathbf{x}}[0]R_{\mathbf{y}}[0]}} \tag{20.9}$$

Note, these modification are available only if $x[n]$ and $y[n]$ are of the same length. Otherwise, Eq. (20.6) is used.
This time the normalized cross-correlation function and the correlation coefficient are related by

$$R_{\mathbf{xy},norm}[k] \approx \rho_{\mathbf{xy}}[k] \tag{20.10}$$

Properties:

$$R_{\mathbf{xy}}[k] = R_{\mathbf{yx}}[-k] \tag{20.11}$$

$$R_{\mathbf{xy}}[-k] = R_{\mathbf{yx}}[k] \tag{20.12}$$

$$\left| R_{\mathbf{xy}}[k] \right| \leqslant \sqrt{R_{\mathbf{x}}[0]R_{\mathbf{y}}[0]} \tag{20.13}$$

$$\left| R_{\mathbf{xy}}[k] \right| \leqslant \frac{1}{2} \left[R_{\mathbf{x}}[0] + R_{\mathbf{y}}[0] \right] \tag{20.14}$$

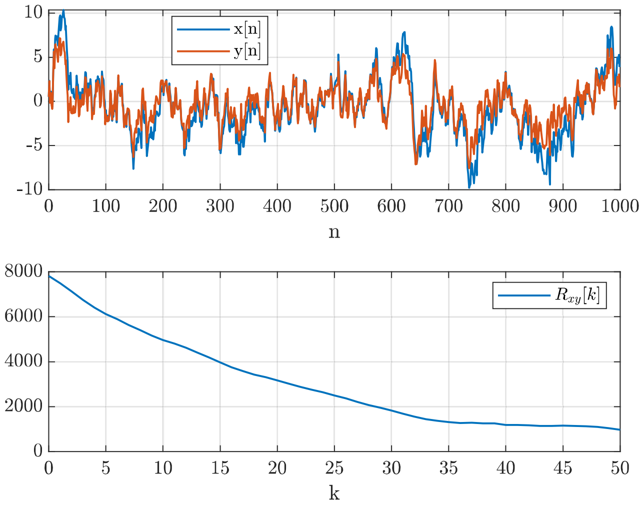
20.1.1 Cross-Covariance Function

For simplicity, a zero-average, $\bar{x}[n] = \bar{y}[n] = 0$, was assumed. When either of the signals is non-zero mean, the subtraction of signal average from the signal before cross-correlation calculation is termed as cross-covariance.
It is similar to auto-correlation and auto-covariance functions.

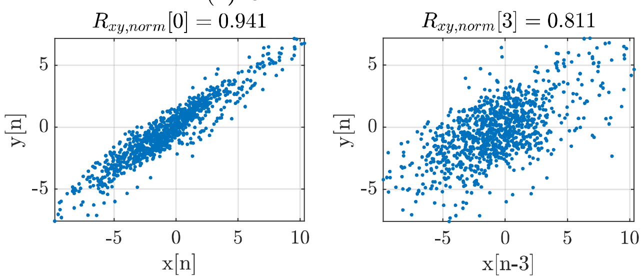
20.2 ARX(0,q) model

An exogenous input model, where the output is a linear combination of the signal values at the different times [7, Example 4.3, pp. 90]

$$\begin{aligned} y[n] &= b_1 x[n-1] + \cdots + b_{m-1} x[n-q] + \epsilon[n] \\ &= \sum_{k=1}^{q} b_k x[n-k] + \epsilon[n] \end{aligned} \tag{20.15}$$



(a) Cross-correlation



(b) $y[n]$ vs. $x[n-k]$

Figure 20.1: Illustration of the linear dependence between $y[n]$ and $x[n-k]$.

In matrix form,

$$\underbrace{\begin{bmatrix} \hat{y}[1] \\ \hat{y}[2] \\ \vdots \\ \hat{y}[L-1] \end{bmatrix}}_{\hat{\mathbf{y}}} = \underbrace{\begin{bmatrix} x[0] & 0 & \vdots & 0 \\ x[1] & x[0] & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x[L-2] & x[L-3] & \vdots & x[L-m-2] \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_{m-1} \end{bmatrix}}_{\mathbf{b}} \tag{20.16}$$

with $\hat{\mathbf{y}} \in \mathscr{R}^{L-1}, \mathbf{X} \in \mathscr{R}^{(L-1)\times q}, \mathbf{b} \in \mathscr{R}^{q}$ Similar ot AR model, the solution is also comprised of the corresponding $R_{\mathbf{xx}}[k]$ and $R_{\mathbf{xy}}[k]$ values.

20.3 General ARX model

Example 20.1: ARX(3,3) model with signals

$$\begin{aligned} x[n] &= x[0], x[1], \ldots, x[7] \\ y[n] &= y[0], y[1], \ldots, y[7] \end{aligned}$$

The required difference equation is

$$\begin{aligned} \hat{y}[n] = {} & a_1 y[n-1] + a_2 y[n-2] + a_3 y[n-2] \\ & + b_1 x[n-1] + b_2 x[n-2] + b_3 x[n-3] \end{aligned} \tag{20.17}$$

Find prediction of $\hat{y}[8]$.

Solution:

$$\underbrace{\begin{bmatrix} x[0] & 0 & 0 & y[0] & 0 & 0 \\ x[1] & x[0] & 0 & y[1] & y[0] & 0 \\ x[2] & x[1] & x[0] & y[2] & y[1] & y[0] \\ x[3] & x[2] & x[1] & y[3] & y[2] & y[1] \\ x[4] & x[3] & x[2] & y[4] & y[3] & y[2] \\ x[5] & x[4] & x[3] & y[5] & y[4] & y[3] \\ x[6] & x[5] & x[4] & y[6] & y[5] & y[4] \end{bmatrix}}_{\mathbf{X}} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y[1] \\ y[2] \\ y[3] \\ y[4] \\ y[5] \\ y[6] \\ y[7] \end{bmatrix}}_{\mathbf{y}} \tag{20.18}$$

The prediction of $\hat{y}[8]$ is straightforward after finding the prediction coefficients by LS minimization. The resulting calculation is comprised of the corresponding $R_{\mathbf{xx}}[k]$ and $R_{\mathbf{xy}}[k]$ values.

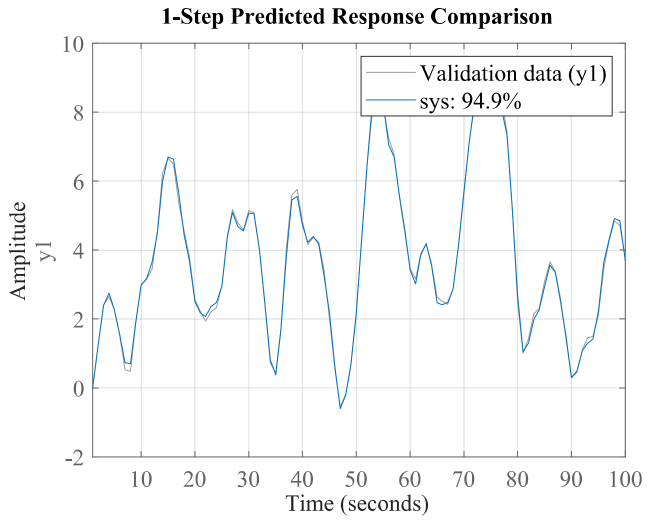


Figure 20.2: Example of ARX model-based prediction. The input is noise-corrupted binary signal that passed through a "synthetic" filter.