14.3 Prediction Error Methods

14.3.1 Exercises

Exercise 3.1: On the use of cross-correlation test for the LS model.

Consider an ARX model
\[ A(q^{-1})y_t = B(q^{-1})u_t + \epsilon_t \]
with parameters \( \theta = (a_1, \ldots, a_{na}, b_1, \ldots b_{nb})^T \). Assume that the estimate \( \hat{\theta} \) is found by LS, show that
\[ \frac{1}{n} \sum_{t=1}^{n} u_t \epsilon_t = 0 \]
for all \( \tau = 1, \ldots, n_b \).

Exercise 3.2: Identifiability results for ARX models.

Consider the system (with \( A_0, B_0 \) coprime, and degrees \( n_{a0}, n_{b0} \))
\[ A_0(q^{-1})Y_t = B_0(q^{-1})u_t + D_t \]
where \( D_t \) is zero mean, white noise. Let \( \{y_t\}_t \) be realizations of the stochastic process \( \{Y_t\} \). Use the LS in the model structure
\[ A(q^{-1})y_t = B(q^{-1})u_t + \epsilon_t \]
with degrees \( n_a \geq n_{a0} \) and \( n_b \geq n_{b0} \). Assume the system operates in open-loop and the input \( u_t \) is PE of order \( n_b \). Let \( \epsilon_t(\theta) \) denote the random variables modeling the residuals \( \epsilon_t \) for given parameters \( \theta = (A, B) \). Prove the following results:

(a) The asymptotic cost function \( E[\epsilon_t^2(\theta)] \) has a unique minimum.

(b) The estimated polynomials are coprime.

Exercise 3.3

Given the predictor
\[ \hat{y}(t+1|t) = \frac{1 - \alpha}{1 - \alpha q^{-1}} y_t, \]
where \( |\alpha| < 1 \) is a constant.

1. For which ARMA-model is this the optimal predictor, and what assumptions do we make on the noise process so that this is \( L_2 \) optimal?

2. What is the optimal two-step ahead predictor \( \hat{y}(t|t-2) \)?

3. If using a numerical optimization procedure for implementing PEM, one wants to use the gradient of \( \epsilon(t+1|t) = y_t - \hat{y}(t|t-1) \) with respect to \( \alpha \). Give a closed form expression of this gradient.
Exercise 3.4: An illustration of the Parsimony Principle.

Consider the following AR(1) process:

\[ Y_t + a_0 Y_{t-1} = D_t, \]

with \(|a_0| < 1\) unknown, and where \(\{D_t\}_t\) is a white noise stochastic process with zero mean and variance \(\lambda^2\). Let us have a realization \(\{y_t\}_{t=1}^n\) of length \(n\) of this process. Then assume we fit this system with the following two candidate models:

\[ M_1 : y_t + a y_{t-1} = \epsilon_t \]

and

\[ M_2 : y_t + a_1 y_{t-1} + a_2 y_{t-2} = \epsilon'_t \]

Let \(\hat{a}\) denote the LS estimate of \(a\) in \(M_1\), and let \(\hat{a}_1, \hat{a}_2\) be the LS estimate in \(M_2\). What are the asymptotic variances \(\sqrt{n}(\hat{a} - a_0)\) and \(\sqrt{n}(\hat{a}_1 - a_0)\)?

180