ESS101 Mod & Sim
SP1 2019
Assignment 2 - System Identification
Due date: 2019-10-04

In this assignment, we will explore various aspects of parameter estimation (including least-squares and maximum likelihood estimation) and system identification. Write a small report providing the answers to the questions. For the last question (compute exercise), provide the answers in the report and a file with the Matlab code, too.

Instructions

The assignments comprise an important part of the examination in this course. Hence, it is important to comply with the following rules and instructions:

- The assignment is pursued in groups of two students. For group formation procedures, see Canvas.
- The findings from each assignment are described in a short report, written by each group independently.
- The report should provide clear answers to the questions, including your motivations, explanations, observations from simulations etc. Figures included in the report should have legends, and axes should be labelled.
- Since the assignments are part of the examination in the course, plagiarism is of course not allowed. If we observe that this happens anyway, it will be reported.
- The report should be uploaded to Canvas *before the deadline*. Matlab code is uploaded as separate files. If the deadline is not met, we cannot guarantee that you will be able to complete the assignment during the course.
- The written report is graded PASS, REVISE or RESUBMIT. With a decision REVISE, you will get a new due date to provide a revised report with changes highlighted. With a decision RESUBMIT, you will need to submit a completely rewritten report.

Please be environment-friendly: avoid printing these pages!

First part - Estimators, Maximum Likelihood and Least Squares.

1. **Unbiased estimate.** Consider the data samples $\{x[0], \cdots, x[N-1]\}$, where each sample is distributed as $\mathscr{U}[0,\theta]$ (uniform distribution in the interval $[0,\theta]$) and the samples are IID.
   (a) Find an unbiased estimator for θ. The range of θ is $0 < \theta < \infty$.

   Hint: Think about the relation between the mean value of the data samples and the parameter you want to estimate!

2. **Sample mean and Maximum Likelihood** Consider the following data samples

$$x[k] = A + w[k], \quad k = 0, \cdots, N-1 \tag{3}$$

   with $w[k]$ being white Gaussian noise, $w[k] \sim \mathscr{N}(0, \sigma^2)$. In Exercise 4.9, it can be seen that a possible estimator for the parameter A is the sample mean,

$$\hat{A} = \frac{1}{N} \sum_{k=0}^{N-1} x[k]. \tag{4}$$

   (a) Show that the sample mean is a Maximum Likelihood estimator of the parameter A.

3. **Linear least squares.** Consider the linear system

$$y[k] = \theta u[k] + e[k], \quad k = 0, \cdots, N-1 \tag{16}$$

with θ being a parameter to estimate and $e[k] \sim \mathcal{N}(1, \sigma^2)$.

   (a) Compute the least squares estimate of θ.

   (b) Is the estimate biased? If yes, give the expression of the bias.

   (c) What happens to the estimation of θ if $u[k] = 0, \forall k$?

4. **Invariance property of the maximum likelihood estimate.** Consider the data samples

$$y[k] = A + e[k], \quad k = 0, \cdots, N-1 \tag{18}$$

with $e[k]$ being WGN with variance σ^2. An estimate for a transformed parameter $\alpha = \mathrm{e}^A$ is required.

   (a) Show that the maximum likelihood estimate of α is

$$\hat{\alpha} = \mathrm{e}^{\hat{A}} \tag{19}$$

   where $\hat{A}$ is the maximum likelihood estimate of A.
   Hint: Apply the transformation of parameters in (18), and compute the max-likelihood estimate for α. Then compare it with the max-likelihood estimate of A.

Second part - Identification of linear models for dynamical systems.

1. **One-step ahead predictor**. Consider the following model structure

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t) + c_1 e(t-1) \tag{25}$$

   (a) What kind of model structure is it (ARX, FIR, ARMAX, OE)?

   (b) Which are the expressions of the plant model G and noise model H?

   (c) Find the 1-step-ahead predictor for Model (25).

   (d) Is the 1-step-ahead predictor you found a linear function of the parameters?

2. **Prediction or simulation?** Consider the following model structure

$$y(t) + a_1 y(t-1) = b_0 u(t) + e(t) + a_1 e(t-1) \tag{29}$$

   (a) What kind of model structure is it (ARX, FIR, ARMAX, OE, BJ)?

   (b) Find the 1-step-ahead predictor for the model in (a). Something weird is happening! Can you explain what and why?

3. **Computer exercise. Identification of an ARX model.** In Matlab, load the data files `input.mat` and `output.mat` from Canvas. The data contained in the file are generated using an ARX model with additive white Gaussian noise ($\sigma^2 = 0.01$). Consider the following candidate ARX model structures,

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t) \tag{32a}$$

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + b_1 u(t-1) + e(t) \tag{32b}$$

$$y(t) + a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) = b_1 u(t-1) + e(t) \tag{32c}$$

   (a) Split the data in estimation and validation data, and identify the parameters of each candidate ARX model, using the estimation data. *Hint: ARX models provide a 1-step-ahead predictor which is a linear regressions in the parameters. Hence, compute the predictor and use linear least squares to find a close form of the expression for the parameters to find.*

   (b) For each estimated model, write a small script for computing the simulated and predicted (1-step) output, using the validation data. Compare those outputs with the true data (you can use Root Mean Square Error as a measure of the goodness of the comparison). Which model is the best in simulation? And which one in prediction?

   (c) Compute the covariance of the parameters. Which one is the best model describing the system generating the data, according to the covariance analysis?