Exercises for Lecture Course on Modelling and System Identification (MSI)
Albert-Ludwigs-Universität Freiburg – Winter Term 2014

Exercise 4: Covariance Estimation in a Single Experiment

(to be returned on Nov 18, 2014, 8:15 in HS 26, or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl and Robin Verschueren

Please remember to provide a solution on paper (written or typed) including all the necessary graphs from MATLAB. The MATLAB code (.m-files) should be sent to `robin.verschueren@gmail.com` and `giovanni@ampyxpower.com`

Aim of this sheet is to learn how to get an estimate of the confidence ellipsoid of a least-squares estimator out of a single experiment.

Exercise Tasks

1. **Confidence Ellipsoids for Single Experiments in LS-Estimation:** Download the data `data4_2.txt` from `http://bit.do/MSI_ex`, and import it in MATLAB.

   The data is formatted as follows: `| i1 | u1 | u2 |`, each column contains $N_1 = 20$ samples. The data come from current-voltage measurements similar as in Exercise Sheet 2. In this question, however, we used two different voltage meters with different (but unknown) additive noise levels. The regressor `i1` is the same for both tests. Let us first only regard one experiment with data set `| i1 | u1 |`.

   1st Experiment: (4 points)

   (a) Compute a linear least squares fit through the data set `| i1 | u1 |`, using the estimator $\theta^* = (\Phi_N^\top \Phi_N)^{-1} \Phi_N^\top y_N$, and plot the regression line as well as the data points into an (i, u)-diagram.

   (b) Compute the residuals (=measurement errors), compute their mean and variance, and plot a histogram of them.

   (c) Now we come to the main task of this exercise: compute and plot the (one-sigma) confidence ellipsoid of the estimated parameters. Note the difference with Exercise Sheet 2: today, we only have one experiment of each measurement setup ($M = 1$)! Think about the meaning of the confidence ellipsoid.

   (d) What can you say about the real value (E, R)? How probable is it that the estimated ellipsoid contains the true value? Tip: Use the χ^2-distribution given in Fig. 3.4 in the script.

   (e) How probable is it that the true value for R is in the interval $[\hat{R} - 2\sigma_R, \hat{R} + 2\sigma_R]$, where $\hat{R}$ is the estimated value and σ_R the corresponding standard deviation? Give the upper and lower values of this interval in Ohm. Tip: you can find σ_R^2 in the covariance matrix.

   2nd Experiment: (3 points)

   (f) Now perform the same steps as above (a)-(e) on the second experiment, where you take the data set `| i1 | u2 |`.

   (g) Which of the two experiments has the smaller noise and thus a smaller covariance matrix?

   3rd Experiment: (2 bonus points)

   (h) Download the data `data4_3.txt` from `http://bit.do/MSI_ex`, and import it in MATLAB.

   The data is formatted as follows: `| i2 | u3 |`, each column contains $N_2 = 100$ samples. In this question, we use the same voltage meter (with the same noise level) as in the first measurements `u1` above, but we take more samples ($N_2 > N_1$).

   Perform the same steps as above (a)-(e).

   (i) Compare the confidence ellipsoids from all three experiments and plot them all into a (E, R) diagram

(j) Suppose now that someone tells you that the real value for (E, R) is an integer value (in Ohm and Volt). Which value do you think could be the true value?

2. **Maximum-Likelihood Estimation:** (3 points)

   In a high precision telescope looking into outer space, we want to estimate the position of an extremely far star as exactly as possible. Unfortunately, only very few photons arrive every hour to the cells of our light detector. We model the problem in one dimension only. On an line of N detector cells (which also has a length of N millimeters, i.e. each cell is a millimeter wide) we have counted the number of photons $y(k) \in \mathbb{N}$ that arrived in one hour in each cell. We know that the true but unknown light intensity $\lambda(x;\theta)$ of the star is spread out and bell-shaped, and therefore we use bell-shaped function for describing it. The function has three unknown parameters: the center-point θ_1, the overall scaling θ_2, and the spread θ_3, and is given by

$$\lambda(x;\theta) = \theta_2 \exp\left(-\frac{(x-\theta_1)^2}{\theta_3}\right)$$

   The interesting fact is that the number of photons arriving in each hour for a given light intensity follows a Poisson-distribution, i.e. the probability to count $y \in \mathbb{N}$ photons is given by

$$P(y) = \frac{\lambda^y e^{-\lambda}}{y!}$$

   Formulate the negative log-likelihood function for the maximum-likelihood estimator to estimate θ for given N integer numbers $y(1), \ldots, y(N)$. Assume that the midpoint of the k-the detector cell is located at $x = k - 0.5$ [mm].

This sheet gives in total 10 points and 2 bonus points

Some Announcements on the Course Organization

- The first microexam will take place on November 18, 2014, and starts at 8:15 in the lecture hall. The exam ist multiple-choice, and only pen and paper are allowed.
- The lecture on November 19 is cancelled because the lecture room is needed by others.
- The lectures on November 12 and November 25, 2014, will be given by Robin Verschueren as "exercise lectures". He will not present new topics, but use the time to present and discuss the solutions to the exercise sheets. Aim is that everyone has understood the solutions to all sheets so far.