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

Exercise 1: Brief Introduction to Statistics

(to be returned on Oct 28, 2014, 8:15 in HS 101-00-026, or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl and Robin Verscheuren

In this exercise we investigate some important facts from mathematical statistics in theory and in numerical experiments.

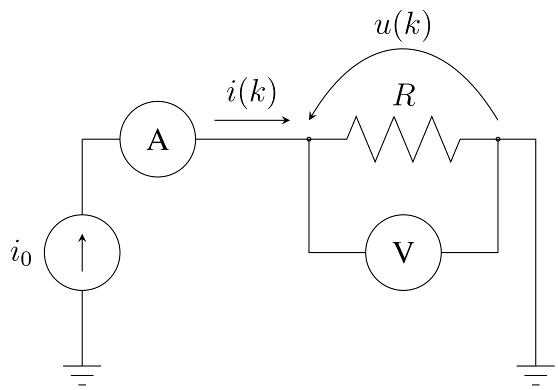
Exercise Tasks

1. Compute the unique minimizer x^* of the function $f(x) = \sum_{i=1}^{N}(\eta_i - x)^2$, where the numbers $\eta_1, \ldots, \eta_N \in \mathbb{R}$ are given. Justify your answer. (2 points)

2. Compute the probability density function $p_Z(z)$ of the scalar random variable $Z = a + bX + cY$, where $a, b, c \in \mathbb{R}$ are constants and X and Y are normally distributed random variables with zero mean and variances σ_X^2 and σ_Y^2. Justify your answer. (4 points)

3. *The covariance matrix of a vector valued random variable X in $\mathbb{R}^n$ with mean $\mathbb{E}\{X\} = \mu_X$ is defined by
$$\mathrm{cov}(X) := \mathbb{E}\{(X - \mu_X) \cdot (X - \mu_X)^\top\}.$$
Prove that the covariance matrix of a vector valued variable $Y = AX + b$ with constant $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ is given by
$$\mathrm{cov}(Y) = A \cdot \mathrm{cov}(X) \cdot A^\top.$$
(2 bonus points)

4. Computer exercise with MATLAB. We consider the following experimental setup:



The resistor R has a current i_0 flowing through it which is determined by the current source. The goal of the experiment is to estimate the value of R using Ohm's law. Since the value of i_0 is not known exactly, we use an ammeter to measure the current $i(k)$ and a voltmeter to measure $u(k)$. Since several measurements are taken, k represents the measurement number. We assume that the measurements are noisy:
$$i(k) = i_0 + n_i(k)$$
and
$$u(k) = u_0 + n_u(k)$$
where u_0 is the true value of the voltage across the resistor and $n_i(k)$ and $n_u(k)$ are the values of the noise.

Let us now investigate the behaviour of the following three different estimators:

- $\hat{R}_{\text{SA}}(N) = \dfrac{1}{N} \sum_{k=1}^{N} \dfrac{u(k)}{i(k)}$
- $\hat{R}_{\text{LS}}(N) = \dfrac{\frac{1}{N} \sum_{k=1}^{N} u(k)i(k)}{\frac{1}{N} \sum_{k=1}^{n} i(k)^2}$
- $\hat{R}_{\text{EV}}(N) = \dfrac{\frac{1}{N} \sum_{k=1}^{N} u(k)}{\frac{1}{N} \sum_{k=1}^{N} i(k)}$

We will use MATLAB to simulate the behavior of these estimators for $i_0 = 5$ and $u_0 = 10$ and the following measurement noise: $n_i(k)$ and $n_u(k)$ are i.i.d. with normal distribution, zero mean, and $\sigma = 2$.

For each of the three estimators, carry out the following tasks (useful MATLAB commands are `help`, `plot`, `for`, `randn`, `mean`, `std`)

(a) Simulate the outcome of M=200 experiments with N_{max}=1000 measurements each, i.e. $k = 1, \ldots, N_{\text{max}}$;

(b) For all 200 experiments plot the function $\hat{R}_*(N)$, $N = 1, \ldots, N_{\text{max}}$ ($*$ can be either SA, LS or EV); for each estimator; to see the stochastic variations, plot all these functions in one graph using `hold on`.

(c) Compute the mean of $\hat{R}_*(N)$ over all experiments and plot it for N from 1 to N_{max};

(d) Compute the standard deviation of $\hat{R}_*(N)$ over all experiments and plot it for N from 1 to N_{max};

(e) *Plot a histogram containing all values of $\hat{R}_*(N_{\text{max}})$;

(f) *Set $\epsilon = 0.01$ and plot the experimental probability $\mathrm{P}(|\hat{R}_*(N) - R| \geq \epsilon)$ w.r.t. N (use $R = u_0/i_0 = 2$).

(4 points and 2 bonus points)

This sheet gives in total 10 points and 4 bonus points (plus 4 bonus points from the previous sheet)