Exercises for Lecture Course on Modeling and System Identification (MSI)
Albert-Ludwigs-Universität Freiburg – Winter Term 2017

Exercise 5: Least Squares with a Single Experiment

Prof. Dr. Moritz Diehl, Robin Verschueren, Alexander Resch

In this exercise you get to run a single experiment, and apply linear-least squares to your data. You also get to explore further the motivation behind linear least squares.

For the the MATLAB exercises, create a MATLAB script called `main.m` with your code, possibly calling other functions/scripts. From running this script, all the necessary results and plots should be clearly visible. Compress all the files/functions/scripts necessary to run your code in a `.zip` file. You should either run the code on your or on your tutors computer.

Exercise Tasks

1. Assume you have a nonlinear model f that relates a measurement y to a series of independent random variables x_i for $i = 1, ..., N$:

$$y = f(x_1, x_2, ..., x_N)$$

The mean of y is μ_y. Each of the x_i has a mean μ_i and a (small) variance σ_i^2. Show that the variance of y can be approximated as:

$$\sigma_y^2 \approx \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\bigg|_{\mu_i} \right)^2 \sigma_i^2$$

where the derivatives are evaluated at the collected means of the random variables x_i. (Hint: use a Taylor expansion.)

2. Please prove that the maximum likelihood estimator and the linear least squares estimator are the same for a relation that is linear-in-the-parameters with noise-free regressors and additive i.i.d. Gaussian measurement noise. (Hint: Please expand on the steps.)

Computer Exercise

1. In order to illustrate the process of estimation with only one experiment, we will estimate the power used by your water kettle[1]. Collect the following:

   - At least three time-keeping devices, each accurate to the seconds, and people to run the devices;
   - A water-heating kettle with specified volume measurements;
   - A volume (liquid) of tap water, at least equal to the kettle's minimum safe heating volume;
   - Pieces of ice.

   As water boils at 100^o C, we need to ensure the starting temperature. For this: mix the ice and the liquid water and let it sit for a short time so that some but not all of the ice pieces dissolve. Pour (with a strainer if necessary) a quantity of the liquid equivalent to the minimum safe heating volume

[1] question is also applicable for a stovetop, if needed...

into the kettle. Turn the kettle on, and time how long it runs before the kettle automatically switches off. Make as many measurements of this time as you have timing devices.

Please assume that the walls of the kettle are very well insulated, so that no heat flows between the kettle and the surrounding air. Further, assume that the power usage of the kettle is constant, and that the kettle turns off exactly when all of the water is at 100^{o}C. Also, assume that this water is similar enough in composition to pure water that its specific heat at constant pressure $c_p \approx 4.19$kJ/kg/K, and that its density $\rho \approx 1$kg/L.

Consider the model for water-heating to be:

$$t = c_p \, \Delta T \, V \rho \, \mathcal{P} + \epsilon_t,$$

where t is the time until boiling, ΔT is the 100K change in temperature, V is the volume of water being boiled, $\mathcal{P} = 1/P$ is the inverse of the kettle power P, and ϵ_t is the measurement noise.

Please combine your procedures for (1c), (1d), (1f), and (1i) in a `powerEstimator.m` function:

```
function [ inverse_power_estimator, inverse_power_variance,
inverse_power_upper_bound, inverse_power_lower_bound,
inverse_power_Rsquared ] = powerEstimator( datafile_name )
```

which takes as argument the name of the data-file `powerData.dat`, which contains one column. The first entry of the column should be the volume of water heated V [L], the subsequent entries of the column should be the measured t_i [s] values. Please have BOTH files prepared. (5 pts)

(a) Please formulate the optimization problem for estimating the inverse power $\mathcal{P}$ as a linear least squares problem. (.5 pt)

(b) Are the assumptions required for linear least squares valid for this experiment? Why or why not? (If not, please make the assumptions anyway for illustrative purposes.) Do you expect the estimator to be unbiased? Why or why not? (.5 pt)

(c) Please use linear least squares to estimate $\hat{\mathcal{P}}$. From this, what is the estimated kettle power $\hat{P}$? (.5 pt)

(d) Further, estimate $\sigma^2_{\hat{\mathcal{P}}}$, the variance of the estimated $\hat{\mathcal{P}}$. (1 pt)

(e) Is the variance of the estimator the same as the variance of the measurement noise? Why? What is the relationship between the two? (.5 pt)

(f) What are the bounds of the one-sigma confidence interval for $\mathcal{P}$? To what values of P do these bounds correspond? Plot the bounds on P and the estimate $\hat{P}$. (.5 pt)

(g) Please give an interpretation of the confidence interval, given the assumptions in (1b). (.5 pt)

(h) Somewhere (probably on the underside) of your kettle, there is likely a sticker[2] that gives a range of operating power. Please add this power range to the plot of (1f). How does this range compare to your estimate and the one-sigma confidence interval? Why is this? (.5 pt)

(i) What is the R-squared value for this estimator? What does this value mean? (.5 pt)

Do not use the predefinded MatLab methods *var()* and *inv()* for your computations.

[2]If this sticker is missing, please assume that this range would be 2kW $\pm$ 200W.