Exercises for Course on Modeling and System Identification (MSI)
Albert-Ludwigs-Universität Freiburg – Winter Term 2018-2019

Exercise 4: Weighted Linear Least-Squares

(to be returned on Nov 30, 2018, 10:00 in SR 00-010/014, or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl, Tobias Schöls, Katrin Baumgärtner, Alexander Petrov

In this exercise, you will learn some basic properties about quadratic functions and how they relate to weighted linear least-squares.

Exercise Tasks

1. Suppose $W \in \mathbb{R}^{n\times n}$ is symmetric and positive definite. We consider the following optimization problem

$$\underset{\theta\in\mathbb{R}^d}{\text{minimize}} \quad \frac{1}{2}r^\top W r \tag{1}$$

where the vector of residuals $r \in \mathbb{R}^n$ denotes the difference between the measurement vector $y \in \mathbb{R}^n$ and the linear model $M(\theta) = \Phi\theta$ with $\Phi \in \mathbb{R}^{n\times d}$, $\theta \in \mathbb{R}^d$, i.e. we have

$$r = y - M(\theta) = y - \Phi\theta = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix} - \begin{bmatrix} \phi^{\mathrm{T}}(1) \\ \vdots \\ \phi^{\mathrm{T}}(N) \end{bmatrix}\theta.$$

   (a) Expand the objective $f(\theta) = \frac{1}{2}r^\top W r$ as a quadratic function of θ, using the following notation:

$$f(\theta) = \frac{1}{2}\theta^{\mathrm{T}}H\theta + g^{\mathrm{T}}\theta + c.$$

   Please identify H, g, c with the given matrices y and Φ in the objective after expansion. (1 point)

   (b) Calculate the the gradient and the Hessian of the objective. (1 point)

   (c) Analytically solve the optimization problem. (1 point)

   (d) Is it a convex problem? Prove it. (1 point)

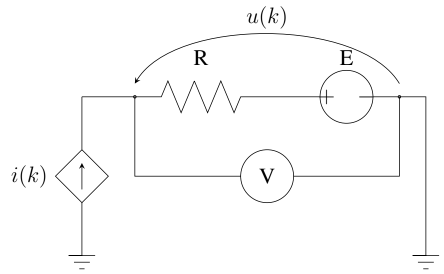
   (e) The weighted least-squares problem (WLS) given by (1) can be formulated as an unweighted least-squares problem with rescaled measurements and rescaled regressor matrix.

   Please re-write the WLS optimization problem (1) as an unweighted LLS problem, i.e. specify $\tilde{y}$ and $\tilde{\Phi}$ such that

$$\frac{1}{2}\|\tilde{y} - \tilde{\Phi}\theta\|_2^2 \;=\; \frac{1}{2}r^\top W r.$$

   (1 point)

2. Recall the resistance estimation example from the last exercise sheet. Again, we consider the following experimental setup:



We assume that only our measurements of the voltage are corrupted by noise, i.e. we make the following model assumption:

$$u(k) = R_0 i(k) + E_0 + n_u(k)$$

where $n_u(k)$ follows a zero-mean Gaussian distribution.

You are given the data of N_e students, each of them performed the same experiment where they measured the voltage $u(k)$ for increasing values of $i(k)$, $k = 1, \ldots, N_m$.

Unfortunately, the fan of your measuring device is broken. Thus, it starts heating up over the course of the experiment which decreases the accuracy of your measurements such that later measurements are much noisier than earlier ones.

(a) ON PAPER: We already provided a plot showing the measurements from all students. What do you observe?

To account for the decreasing accuracy of your measuring device, you decide to assume that the noise variance $\text{var}\,(n_u(k))$ is proportional to the timestep k, i.e.

$$\text{var}\,(n_u(k)) = c \cdot k, \; k = 1, \ldots, N_m,$$

for some fixed value of c (you may simply choose $c = 1$). How do you make use of this assumption when applying weighted linear least-squares?

(1 point)

(b) MATLAB: For student 1, perform both linear least-squares (LLS) and weighted linear least-squares (WLS) to obtain estimates of the parameter $\theta_0 = [R_0, E_0]^\top$. Plot the data of student 1, as well as the fit obtained from LLS and WLS in a single figure.

ON PAPER: What do you observe? (1 point)

(c) MATLAB: For each student $d = 1, \ldots, N_e$, compute $\theta_{\text{LLS}}^{(d)}$ and $\theta_{\text{WLS}}^{(d)}$. (1 point)

(d) MATLAB: Estimate the mean and covariance matrix of the random variables θ_{LLS} and θ_{WLS} by calculating the sample mean $\bar{\theta}_{*\text{LS}} = \frac{1}{N_e}\sum_{d=1}^{N_e} \theta_{*\text{LS}}^{(d)}$ and the sample covariance matrix $\Sigma_{*\text{LS}}$ that is given by

$$\Sigma_{*\text{LS}} = \frac{1}{N_e - 1}\sum_{d=1}^{N_e}\left(\theta_{*\text{LS}}^{(d)} - \bar{\theta}_{*LS}\right)\left(\theta_{*\text{LS}}^{(d)} - \bar{\theta}_{*LS}\right)^\top.$$

Here $*$LS refers to LLS and WLS. (1 point)

(e) MATLAB: Plot $\theta_{\text{LLS}}^{(d)}$ and $\theta_{\text{WLS}}^{(d)}$, $d = 1, \ldots, N_e$, where the x-axis corresponds to the estimated R values and the y-axis corresponds to the estimated E values.

Plot the mean and 1σ-confidence ellipsoids for both θ_{LLS} and θ_{WLS} in the same figure.

ON PAPER: What do you observe? (1 point)

(f) ON PAPER: In part (b) we assumed that the measurement noise is proportional to k. Does θ_{WLS} depend on the choice of the proportionality factor? Why (not)? (1 point)

This sheet gives in total 11 points.