Exercises for Lecture Course on Modeling and System Identification (MSI)
Albert-Ludwigs-Universität Freiburg – Winter Term 2016

Exercise 2: Linear Least Squares and the Freiburg Atmosphere

(to be returned on Nov. 7, 2016, 8:15 in SR 00-010/014, or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl, Robin Verschueren, Rachel Leuthold, Tobias Schöls, Mara Vaihinger

In this exercise, you will discover the linear least squares estimation method. For 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 and send it to `msi.syscop@gmail.com`. Please state your name and the names of your team members in the e-mail.

Exercise Tasks

1. The covariance matrix of a vector-valued random variable X in $\mathbb{R}^n$ with mean $\mathbb{E}\{X\} = \mu_X$ is defined by
$$\operatorname{cov}(X) := \mathbb{E}\left\{(X - \mu_X)(X - \mu_X)^{\mathrm{T}}\right\}.$$
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 (1 point)
$$\operatorname{cov}(Y) = A \operatorname{cov}(X) A^{\mathrm{T}}.$$

2. Consider again the measurement data describing the thermal expansion of the steel bar. The bar expands, from some initial length L_0 [cm] at an initial temperature T_0 [K], to a length L [cm] when exposed to a change in temperature ΔT [K].

ΔT [K]	5	15	35	60
L [cm]	6.55	9.63	17.24	29.64

Let's assume that the relationship between L and ΔT is a polynomial relationship, such that:
$$L = \theta_0 + \theta_1 \Delta T + \theta_2 \Delta T^2$$
where $\theta \in \mathbb{R}^3$ is an unknown constant. (3 points)

   (a) Consider the summation form of the optimization problem that estimates the linear least squares coefficients, which reads as:
$$\hat{\theta}_{LS} = \arg\min_{\theta\in\mathbb{R}^d} \frac{1}{2}\sum_{i=1}^{N}(y_i - L(\theta, \phi_i))^2.$$
Let's play matching. What are:

      i. the number of decision variables $d \in \mathbb{Z}^+$?

      ii. the column matrix $y \in \mathbb{R}^N$ that contains the measurements?

      iii. the form of the row matrix $\phi_i \in \mathbb{R}^{1\times d}$ that determines $L(\theta, \phi_i) = \phi_i\theta$?

iv. the squared difference between the measurement y_i and the model $L(\theta, \phi_i)$?

There is an equivalent matrix form of the above optimization problem that simplifies the optimization problem above, which reads as:

$$\hat{\theta}_{LS} = \arg\min_{\theta \in \mathbb{R}^d} f = \arg\min_{\theta \in \mathbb{R}^d} \frac{1}{2} \left|\left| y - \mathbf{\Phi}\theta \right|\right|_2^2 .$$

What is:

v. the regression matrix $\mathbf{\Phi} \in \mathbb{R}^{N \times d}$?

(b) The first order necessary condition[1] for optimality says that if $\hat{\theta}_{LS}$ is a minimizer of this optimization problem, then $\nabla f(\hat{\theta}_{LS}) = 0$. What linear relationship must $\hat{\theta}_{LS}$ satisfy?

(c) What is an analytical expression for $\hat{\theta}_{LS}$?

3. Let's consider atmospheric data taken by radiosonde at the Freiburg airport. This data is available in its complete form[2], or lightly pre-processed on the course page. Take a moment to familiarize yourself with the dataset `GMXUAC00001-preprocessed.csv`, using the (edited to reflect the data pre-processing) readme file `igra2-data-format-preprocessed.txt`.

A radiosonde[3] is a weather-balloon equipped with a barometer to measure air pressure, a resistance thermistor to measure air temperature, and - on older radiosondes - a mechanical switch that connected the thermistor at predetermined intervals of the pressure. The altitude is calculated from the temperature and pressure measurements. (5 points)

(a) Plot the altitude z [m] vs. the air pressure p [Pa] data, with pressure along the x-axis and altitude along the y-axis. When looking at the data for altitude as a function of pressure, what is the lowest order polynomial relation that you would expect to give a meaningful linear least squares fit? What form does this relation take? (Hint: remember that linear least squares does not necessarily require a linear relationship between p and z.)

(b) Which properties should the data fulfil, for linear least squares to be an appropriate estimation method? Make those assumptions for the following sub-questions.

(c) Choose $y \in \mathbb{R}^N$, $\phi_i \in \mathbb{R}^{1\times d}$ and $\mathbf{\Phi} \in \mathbb{R}^{N\times d}$, to correspond to the linear least squares problem form.

(d) Is $\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}$ invertible? Why? Does Matlab return a warning, if you attempt to invert this matrix? Why?

(e) In general, it is not advised to use a linear system solve, using the pseudo-inverse, for problems that have this (see 3d) behavior. Try it anyways. Plot the estimated altitudes as a function of pressure in your data-plot, and consider the coefficients of the fitted polynomial. Do these coefficients appear reasonable? (Hint: use the backslash operator in the linear system solve.)

(f) Reconsider your fitting problem, using units of [10^5 Pa] for p and [km] for z. Apply a linear system solve with the pseudo-inverse to this problem. Add this estimate to the data-plot, and consider the coefficients of the fitted polynomial. Does scaling the measurements improve the performance of linear least squares? Why?

[1] You're welcome to demonstrate the convexity of this problem for your own happiness.

[2] http://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive, IGRA2 station-code GMXUAC00001

[3] More information can be found at http://www.aos.wisc.edu/ hopkins/wx-inst/wxi-raob.htm.