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

Exercise 5: Ill-Posed Linear Least-Squares

(to be returned on Dec 4th, 2019, 8:30 in HS 00 036 (Schick-Saal), or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl, Tobias Schöls, Naya Baslan, Jakob Harzer, Bryan Ramos

Exercise Tasks

1. ON PAPER: We would like to estimate a constant $\theta_0 \in \mathbb{R}$ that is corrupted by additive zero-mean Gaussian noise, i.e. we assume the following model

$$y = \theta_0 + \epsilon$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. To this end, we use *regularized* linear least-squares, i.e. we compute the estimate $\hat{\theta}_\mathrm{R}$ given by

$$\hat{\theta}_\mathrm{R} = \arg\min_{\theta \in \mathbb{R}} \frac{1}{2}\|y - \Phi\theta\|_2^2 + \frac{\alpha}{2}\|\theta\|_2^2$$

where $\theta \in \mathbb{R}$, $\Phi = (1, \dots, 1)^\top \in \mathbb{R}^{N\times 1}$ and $\alpha \geq 0$. From the lecture, we know that the solution to this optimization problem is given by

$$\hat{\theta}_\mathrm{R} = \left(\Phi^\top\Phi + \alpha\mathbb{I}\right)^{-1}\Phi^\top y$$

   (a) Calculate the expected value $\mathbb{E}\left\{\hat{\theta}_\mathrm{R}\right\}$ of $\hat{\theta}_\mathrm{R}$. Is the estimator unbiased and/or asymptotically unbiased?
   Hint: Check Section 4.5.1. of the lecture notes. (2 points)

   (b) Calculate the variance $\mathrm{var}\left(\hat{\theta}_\mathrm{R}\right)$ of $\hat{\theta}_\mathrm{R}$. Compare with the unregularized case, i.e. $\alpha = 0$.
   Hint: Check Section 4.5.2. of the lecture notes. (2 points)

   (c) What value takes the Cramer-Rao bound on the variance in this specific case? Does the result from (b) contradict this lower bound? (1 point)

2. You are given the following ill-posed Linear Least-Squares problem:

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}} \frac{1}{2}\|y - \Phi\theta\|_2^2 \qquad y = \begin{bmatrix} 1 \\ \vdots \\ 9 \end{bmatrix} \in \mathbb{R}^9 \qquad \Phi = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \vdots & \vdots \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \in \mathbb{R}^{9\times 2} \qquad \theta \in \mathbb{R}^2$$

On Grader you will find a template for this problem. If you copy the code to your computer, you can view the minimization problem in 3D.

   (a) ON PAPER: Why is this an ill-posed problem? What issue do you run into when following the usual LLS approach of $\hat{\theta} = (\Phi^\top\Phi)^{-1}\Phi^\top y$? (0.5 points)

   (b) ON PAPER: Which two approaches do you know to solve this issue? (0.5 points)

   (c) ON PAPER: What are advantages and disadvantages of using one approach instead of the other? State one each! (1 point)

   (d) MATLAB: Find a $\hat{\theta}$ using both methods from (b). Use $\alpha = 0.1$.
   Hint: Useful Matlab commands are: `inv()`,`pinv()`,`eye()` (1 point)

(e) ON PAPER: The original minimization problem is visualized in a figure with the two solutions (your $\hat{\theta}$ from the previous example) as red x. Why do the solutions end up where they are? Give a reason for each solution! (1 point)

3. In this exercise task, you compare LLS and regularized LLS. As before, the regularized linear least-squares estimator is defined as

$$\hat{\theta}_{\mathrm{R}} = \arg\min_{\theta \in \mathbb{R}^n} \frac{1}{2}\|y - \Phi\theta\|_2^2 + \frac{\alpha}{2}\|\theta\|_2^2$$

where $\alpha \geq 0$. Note that $\alpha = 0$ corresponds to the ordinary linear least-squares estimator. We provide data from $N_e = 10$ experiments each comprising $N_m = 9$ measurements.

(a) MATLAB: For $\alpha \in \{0, 10^{-6}, 10^{-5}, 1\}$, fit a polynomial of order 7 to the data of the first experiment. Plot the data and the fitted polynomials. (1 point)

(b) MATLAB: For experiment 1 and for each α, compute the L_2-norm of the estimated parameters. ON PAPER: Compare the results. Do they match your expectation? (1 point)

(c) MATLAB: To compare the goodness of fit, compute the R^2 values for each of the three fits obtained for experiment 1.
ON PAPER: Compare the results. (1 point)

(d) MATLAB: For each α and each experiment, fit a polynomial of order 7. For each α, plot the fitted polynomials in a subplot.

Compute the average parameter vector for each α and plot the polynomial obtained from the averaged parameter vector.

ON PAPER: What do you observe? How does this relate to the result from Task 1b? (2 points)

This sheet gives in total 14 points.