Laboratory exercise 1

Linear Estimation in Static Systems

Name: Luka Vukelic **JMBAG:** 0036530728

- **Exercise goals**
  Understanding linear estimation in static systems. Implementation and examination of key properties of widely used least squares and minimum mean square error estimators.
- **Preparation**
  Exercise is done in the Matlab computing environment. In order to successfully complete the exercise, study the corresponding lectures for static estimation and use the provided Matlab script.
- **Graphs**
  Attach pictures of figures as comments below corresponding tasks. (For Adobe Reader: Tools->Comment & Markup->Attach a File as a Comment)
- **Useful matlab functions**
  `mvnrnd, plot, plot3, trace, diag, mean`

Introduction

In 1795, Carl F. Gauss conducted the first known instance of the least squares estimation in order to predict future positions of the Ceres asteroid. In this laboratory exercise, a similar problem will be studied. Consider the orbit of a celestial body that can be parameterized as an ellipse in the 3D space given sufficient gravitational pull. With known speed and measurement timestamps, the simplified parametric orbit equation is

$$\mathbf{z}_{t_i} = \mathbf{p}_0 + \mathbf{d}_1 \cos(t_i) + \mathbf{d}_2 \sin(t_i), \tag{1}$$

where $\mathbf{z}_{t_i}$ is a 3D position in the Euclidean space at time instant t_i, $\mathbf{p_0}$ is the center of gravity with $\mathbf{d}_1$ and $\mathbf{d}_2$ being conjugate diameters of the ellipse. In this exercise, measurements will be modeled as 3D position observations corrupted with independent identically distributed zero-mean additive Gaussian noise $\mathbf{v}_{t_i}$. Assuming that $\mathbf{x} = [\mathbf{p}_0^\intercal, \mathbf{d}_1^\intercal, \mathbf{d}_2^\intercal]^\intercal$ represents the parameter vector being estimated, the parametric measurement model for a single measurement can be written as

$$\mathbf{z}_{t_i} = \mathbf{H}_{t_i}\mathbf{x} + \mathbf{v}_{t_i}, \mathbf{v}_{t_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{t_i}). \tag{2}$$

To write this model in the matrix form, the N measurements for time instants t_1 to t_N can be organized into a larger linear system, $\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}$, $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, where

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_{t_1} \\ \vdots \\ \mathbf{z}_{t_\mathrm{N}} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} \mathbf{H}_{t_1} \\ \vdots \\ \mathbf{H}_{t_\mathrm{N}} \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{t_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Sigma}_{t_\mathrm{N}} \end{bmatrix}. \tag{3}$$

Assignments

Task 1 : Least squares (LS) estimator

a) Measurements in this exercise will be artificial, i.e., you will need to generate them by adding noise sampled from a Gaussian distribution. Generate noisy measurements by building a multiple measurements system as in (3). Use the following parameters: $\mathbf{p}_0 = [200, 300, 300]^\intercal$, $\mathbf{d}_1 = [0, 200, 0]^\intercal$, $\mathbf{d}_2 = [100, 0, 200]^\intercal$, $[t_1, ..., t_{50}] = \texttt{linspace}(0, 2\pi, 50)$, $\mathbf{\Sigma}_{t_i} = \texttt{diag}([4, 4, 4])$. By taking into consideration equations (1) and (2), parametrically construct the matrix $\mathbf{H}_{t_i}$ and report it below. Additionally, upload a graphic comparing the orbital trajectory generated without noise to the trajectory generated with the described noise.

Note: Use the provided Matlab script as a starting point.

[1, 0, 0, cos(t), 0, 0, sin(t), 0, 0;
0, 1, 0, 0, cos(t), 0, 0, sin(t), 0;
0, 0, 1, 0, 0, cos(t), 0, 0, sin(t)]

b) Implement a least squares estimator $\hat{\mathbf{x}} = \arg\min_x (\mathbf{z} - \mathbf{Hx})^\intercal(\mathbf{z} - \mathbf{Hx})$. Repeat the estimation process M times, each time generating measurements with different noise sampled from the Gaussian distribution and estimating $\hat{\mathbf{x}}_\mathbf{j}$ given the corresponding measurements. Estimate the variance of your LS estimator $\hat{P}$ with the following equation: $\frac{1}{M}\sum_j^M(\hat{\mathbf{x}}_j - \mathbf{x})^\intercal(\hat{\mathbf{x}}_j - \mathbf{x})$.

Note: You are estimating the variance with an additional estimator that is only asymptotically unbiased and consistent!

Compare the estimated variance with the Cramer-Rao lower bound that represents the theoretically lowest possible estimator variance for a particular problem. For measurement models modeled with linear systems and Gaussian noise, such as this problem, you can calculate it analytically as CRLB $= \text{trace}((\mathbf{H}^\intercal\mathbf{\Sigma}^{-1}\mathbf{H})^{-1})$. Use $[t_1, ..., t_N] = \texttt{linspace}(0, 2\pi, N)$, $\mathbf{\Sigma}_{t_i} = \texttt{diag}([2, 2, 2])$ and other parameters as in the subproblem a).

Repeat this process for different number of measurements $N \in \texttt{linspace}(5, 50, 10)$. Use the `plot` function to draw the CRLB and the estimated variance for different N on the x axis. Draw 3 plots on the same figure; CRLB, estimated $\hat{P}$ for $M = 5$ and estimated $\hat{P}$ for $M = 50$ (this may take some time). Upload the graphic and explain the results below. How do the theoretical CRLB and practical estimation of the variance change with the increasing number of measurements? Does the variance estimation converge to the CRLB value for higher M and why? If so, what can you conclude about the efficiency of the least squares estimator?

With the increasing number of measurements CRLB and variance estimations are reducing exponentially. Variance estimation converge to the CRLB value for higher M because with more measurements, the LS estimator provides increasingly precise and reliable parameter estimate if our dataset is generated with normal, gaussian, distribution (which is our case). LS estimator is efficient when we use big enough M parameter.

c) Instead of varying the number of measurements, repeat the variance estimation and CRLB calculation for different noise variances σ_z^2, $\mathbf{\Sigma}_{t_i} = \texttt{diag}([\sigma_z^2, \sigma_z^2, \sigma_z^2])$. Use $\sigma_z^2 \in \texttt{linspace}(1, 5, 10)$ and draw two plots on the same figure with σ_z^2 on the x axis, one for the CRLB and the other for the estimated variance. Use $N = 10$ for the number of measurements, i.e., $[t_1, ..., t_N] = \texttt{linspace}(0, 2\pi, 10)$ and $M = 10$ for variance estimation. Comment the results and upload your graphic below.

When the estimated variance curve closely follows the CRLB curve, it suggests that the LS estimator is performing close to optimally, achieving efficiency similar to the theoretical lower bound.
CRLB demonstrates a linear increase with noise variance while the estimated variance is zigzag pattern due to the complexities of estimation processes.

Task 2 : Weighted least squares (WLS) estimator

a) Repeat the process from the task 1.a) with the noise variance that depends exponentially on the parameter t_i, i.e., $\mathbf{\Sigma}_{t_i} = \texttt{diag}([e^{t_i}, e^{t_i}, e^{t_i}])$. Again, draw two plots on the same figure for $N \in \texttt{linspace}(5, 50, 10)$, one for the CRLB and other for the estimated variance with M = 20. Comment the results and upload the graphic below. Does the LS estimator attain the CRLB value? Why?

Variance estimation is constantly larger then CRLB but as number of measurements increase the difference between these two decrease. LS estimator does not attain CRLB value because of our more complex covariance matrix. Using a covariance matrix with elements that vary exponentially with a parameter "t" can introduce complexities into the noise model that may impact the efficiency of the LS estimator.

b) Implement the weighted least squares estimator, i.e., the maximum likelihood estimator, $\hat{\mathbf{x}} = \arg\min_x (\mathbf{z} - \mathbf{H}\mathbf{x})^\intercal \mathbf{W} (\mathbf{z} - \mathbf{H}\mathbf{x})$ with the optimal $\mathbf{W}$. Draw the same plot from the teask 2.a), this time using the WLS estimations for variance calculation. Comment the results and upload the graphic below. Compare with the figure from the task 2.a). What is the optimal $\mathbf{W}$?

Using weighted least squares estimator we give more importance for estimate calculation to data with narrower variance. Because of that LS estimator can attain CRLB values. W optimal is inverse of sigma matrix.

Task 3 : Minimum mean square error (MMSE) estimator

a) Now consider that you have some accurate prior knowledge about $\mathbf{x}$, i.e., that the distribution you use to model your prior exactly matches the true distribution. Implement the MMSE estimator, assuming that $\mathbf{x} \sim \mathcal{N}(\overline{\mathbf{x}}, \mathbf{P}_\mathbf{x})$, with $\overline{\mathbf{x}}$ being equal to the parameters defined in the task 1.a). Thus, in addition to the noise, for each of the M iterations of the variance estimation, you need to sample $\mathbf{x}$ from a Gaussian distribution with the `mvnrnd` function. Estimate the variance of both the MMSE and LS estimators for varying variance of the prior distribution σ_x^2, i.e., $\mathbf{P_x} = \sigma_x^2\mathbf{I}$. Use $N = 10$, $M = 50$ and $\mathbf{\Sigma}_{t_i} = \texttt{diag}([2, 2, 2])$. Draw the plots with estimated variances of MMSE and LS estimators for $\sigma_x^2 \in \texttt{linspace}(0.01,\ 5,\ 40)$. Comment the results and upload the graphic below.

MMSE is constantly better because of prior knowledge.

b) Construct the same figure as in the task 3.a), this time varying the measurement noise variance instead of the prior variance. Define the varying measurement noise as $\mathbf{\Sigma}_{t_i} = \texttt{diag}([\sigma_z^2, \sigma_z^2, \sigma_z^2])$ and draw the plots with the estimated variances of MMSE and LS estimators for $\sigma_z^2 \in \texttt{linspace}(0.01,\ 10,\ 40)$. Use $N = 10$, $M = 50$ and $\mathbf{P_x} = \sigma_x^2\mathbf{I}$ with $\sigma_x^2 = 2$. Comment the results and upload the graphic below.

For sigma of z larger then 10, variance of MMSE remains on the same level while variance of LS keeps increasing.

Exercise submission

Create a zip archive containing **this pdf with the filled out answers** and **all other source code files**. Upload on Moodle.