Sensor Data Fusion

Exercise 6

Prof. Dr.-Ing. Marcus Baum
M.Sc. Kolja Thormann

www.fusion.informatik.uni-goettingen.de

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

DATA
FUSION Lab

Today

- Lecture review
- Questions
- Homework 5 solution
- Example - Woodbury Identity proof
- Problem 6 - Silly Estimator with Prior
- Homework 6 presentation

Questions

What is the difference between Bayesian and Fisher approach?

- We have $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{e}$
- Fisher: $\mathbf{x}$?, $\mathbf{e} \in \mathcal{N}(\mathbf{0}, \mathbf{C}_{ee})$
- Bayesian: $\mathbf{x} \in \mathcal{N}(\boldsymbol{\mu}_x, \mathbf{C}_{xx})$, $\mathbf{e} \in \mathcal{N}(\mathbf{0}, \mathbf{C}_{ee})$
- In general
  - Prior: $p(\mathbf{x})$
  - Likelihood: $p(\mathbf{y}|\mathbf{x})$
  - $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$

Questions

How can you measure the quality of an estimator in a Bayesian setting?

Bayesian Mean Square Error (BMSE)

$$
\begin{aligned}
\mathsf{BMSE}(\theta_y) &= \mathsf{E}_{x,y}(||\underbrace{\theta_y - x}_{\epsilon}||^2) \\
&= \mathsf{Tr}(\mathsf{Cov}(\epsilon)) + \mathsf{Tr}(\mathsf{E}_{x,y}(\epsilon)\mathsf{E}_{x,y}(\epsilon)^{\mathrm{T}})
\end{aligned}
$$

Questions

What is the optimal Bayesian estimator and how can you compute it?

- $\theta_y^{opt} = \mathrm{E}[\mathbf{x}|\mathbf{y}]$
- Given $p(\mathbf{x})$, $p(\mathbf{y}|\mathbf{x})$
- $\mathrm{E}[\mathbf{x}|\mathbf{y}] = \int \mathbf{x}\underline{p(\mathbf{x}|\mathbf{y})}\mathrm{d}\mathbf{x}$

likelihood prior

- $p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}$

posterior

Homework 5: Solution

Consider a 1D state x and estimates x_1 and x_2. The errors of the estimates are correlated with

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.8\sqrt{1 \cdot 4} \\ 0.8\sqrt{1 \cdot 4} & 4 \end{bmatrix}\right)$$

a) Write a function which visualizes the joint covariance matrix. Use it to draw the matrix along with the column space in the measurement space. Then assume the correlation coefficient 0.8 from the description is unknown and define a covariance which ignores it. Draw that covariance and calculate and visualize the solution of the BLUE.

b) Now write a function which calculates a matrix which is bigger than the input matrix, determined by a parameter α as in the lecture. Calculate the BLUE estimate for different α values and determine which works best. Visualize the results.

Example – Woodbury Identity Proof

From the Woodburry Matrix Identity it follows:
For matrices A, C and V with suitable dimensions, the following holds:

$$\left(A + V^T C V\right)^{-1} = A^{-1} - A^{-1} V^T \left(C^{-1} + V A^{-1} V^T\right)^{-1} V A^{-1},$$

Example – Woodbury Identity Proof

As $(A+UCV)^{-1} \overset{!}{=} \left(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\right)$, we must show that

$$(A+UCV)\left(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\right) = I$$

$$\begin{aligned}
&(A+UCV)\left(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\right)\\
=&\left(I - \underline{U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}}\right) + \left(\underline{UCVA^{-1}} - UCVA^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\right)\\
=&(I+UCVA^{-1}) - \left(U\underline{(C^{-1}+VA^{-1}U)^{-1}VA^{-1}} + UCVA^{-1}U\underline{(C^{-1}+VA^{-1}U)^{-1}VA^{-1}}\right)\\
=&I+UCVA^{-1} - (U+\underline{UCV}A^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\\
=&I+UCVA^{-1} - UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\\
=&I+UCVA^{-1} - UCVA^{-1}\\
=&I
\end{aligned}$$

Problem 6 – Silly Estimator with Prior

Consider the Silly Estimator from last exercise. This time, assume x is distributed according to

$$x \sim \mathcal{N}(\mu_x, \sigma_x^2).$$

a) Which estimator is now better, the natural or the silly estimator?

b) Is there a better estimator?

Given is a random variable $y \sim \mathcal{N}(x, 1)$. The objective is to estimate the deterministic mean x based on a single observation y. We consider two estimators:

- Natural estimator: $\theta_n(y) = y$
- Silly estimator: $\theta_s(y) = 0$

a) Calculate the bias, variance, and MSE of both estimators.

b) In which cases outperforms the silly estimator the natural estimator?

Problem 6: Solution

a)

You can write y as $y = x + e$ with $e \sim \mathcal{N}(0,1)$. Then you get for

- Natural

$$\beta = 0; Var[\theta_n(y) - x] = Var[e] = 1$$
$$\text{BMSE}(\theta_n(y)) = 0 + 1 = 1$$

- Silly

$$\beta = \mathrm{E}[\theta_s(y) - x] = -\mu_x; Var[\theta_s(y) - x] = \sigma_x^2$$
$$\text{BMSE}(\theta_s(y)) = \sigma_x^2 + \mu_x^2$$

- Silly is better if $\mu_x^2 + \sigma_x^2 < 1$

b)

As we are in Bayesian setting (statistical information about x and e given), the best estimator is the optimal Bayesian estimator

$$\theta_{opt}(y) = \mathrm{E}[x|y] \quad .$$

Homework 6 – Bayesian estimators

Consider again the setup of the Silly Estimator exercise with $x \sim \mathcal{N}(1,1)$ and $y = x + e$ with $e \sim \mathcal{N}(0,1)$.

a) Write a function which calculates a possible value for x based on its distribution and then takes a measurement from that value using the measurement noise.

b) Consider the natural estimator $\theta_n(y) = y$. Calculate the empirical mean square error of 1000 runs using this estimator and compare it with the analytic solution.

c) Repeat the process with the silly estimator $\theta_s(y) = 0$.