Vision-Aided Navigation (086761)
Homework #2
Solution

January 28, 2024

Submission in pairs. Please email your submission (a single file in the format ID1-ID2.zip) to andreyz@campus.technion.ac.il by 11 February 2024, 23:59.

Basic probability and Bayesian Inference

1. Consider a random variable $x \in \mathbb{R}^n$ with a Gaussian distribution, written in covariance form as $x \sim \mathcal{N}(\mu, \Sigma)$. Show the corresponding information form $x \sim \mathcal{N}^{-1}(\eta, \Lambda)$ is

$$\mathcal{N}^{-1}(x; \eta, \Lambda) = \frac{e^{\left(-\frac{1}{2}\eta^T \Lambda^{-1} \eta\right)}}{\sqrt{\det(2\pi\Lambda^{-1})}} e^{\left(-\frac{1}{2}x^T \Lambda x + \eta^T x\right)}. \tag{1}$$

2. Consider a standard observation model involving a random variable $x \in \mathbb{R}^n$

$$z = h(x) + v, \quad v \sim \mathcal{N}(0, \Sigma_v), \tag{2}$$

and assume the initial belief regarding the state x is a Gaussian with mean $\hat{x}_0$ and covariance Σ_0.

   (a) Write expressions for the prior $p(x)$ and the measurement likelihood $p(z|x)$.

   (b) A measurement z_1 is acquired. Assuming the measurement was generated by the measurement model (2), write an expression for the posterior probability $p(x|z_1)$ in terms of $p(x)$ and the measurement likelihood.

   (c) Derive expressions for the a posteriori mean $\hat{x}_1$ and covariance Σ_1 such that $p(x|z_1) = \mathcal{N}(\hat{x}_1, \Sigma_1)$.

   (d) A second measurement, z_2 , is obtained. Assuming $p(x|z_1) = \mathcal{N}(\hat{x}_1, \Sigma_1)$ from last clause is given (i.e. $\hat{x}_1$ and Σ_1 are known), derive expressions for $p(x|z_1, z_2) = \mathcal{N}(\hat{x}_2, \Sigma_2)$.

3. Consider a multivariate random variable $x_k \in \mathbb{R}^n$ with the following state transition model

$$x_{k+1} = f(x_k, u_k) + w_k, \quad w_k \sim \mathcal{N}(0, \Sigma_w),$$

and a standard observation model as in exercise 2.

   (a) Write an expression for the motion model $p(x_k|x_{k-1}, u_{k-1})$.

   (b) Assume the robot executes action (or control) u_0 and then acquires a measurement z_1. Write an expression for the a posteriori pdf $p(x_1|z_1, u_0)$ in terms of the prior, motion and observation models.

   (c) In the same setting (given action u_0 and measurement z_1), consider the a posteriori pdf over the joint state

$$x = x_{0:1} \triangleq \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \text{ with } x_i \in \mathbb{R}^n.$$

   Show that calculating the maximum a posteriori (MAP) estimate for x is equivalent to solving a non-linear least squares problem.

   (d) Assume the a posteriori pdf over the joint state $x_{0:1}$ is given in covariance and information forms as

$$p(x_{0:1}|u_0, z_1) = \mathcal{N}(\hat{x}_{0:1}; \Sigma_{0:1}) = \mathcal{N}^{-1}(\hat{\eta}_{0:1}, I_{0:1}) \tag{3}$$

   where

$$\Sigma_{0:1} = \begin{pmatrix} \Sigma_{00} & \Sigma_{01} \\ \Sigma_{01}^T & \Sigma_{11} \end{pmatrix}, \quad I_{0:1} = \begin{pmatrix} I_{00} & I_{01} \\ I_{01}^T & I_{11} \end{pmatrix}$$

Indicate the dimensionality of the covariance matrix $\Sigma_{0:1}$ and of its components. We are interested in the marginal pdf over the state x_1, i.e. $p(x_1|u_0, z_1)$, while marginalizing out the past state x_0. Write expressions (using $\hat{x}_{0:1}, \Sigma_{0:1}, \hat{\eta}_{0:1}, I_{0:1}$) for the marginal covariance and information matrices, Σ'_1 and I'_1, over the state x_1 such that

$$p(x_1|u_0, z_1) = \mathcal{N}(\times, \Sigma'_1) = \mathcal{N}(\times, I'_1)$$

where $\times$ indicates a non-zero vector that is not of interest in this exercise.

Hands-on Exercises - Please print and submit your code.

1. Projection of a 3D point. Consider a camera with focal length of $f = 480$ pixels in each axis and principal point $(320, 270)$ pixels. Camera pose is described by the following rotation and translation with respect to the global frame (as in Homework #1):

$$R_G^C = \begin{bmatrix} 0.5363 & -0.8440 & 0 \\ 0.8440 & 0.5363 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$t_{C\to G}^G = (-451.2459, 257.0322, 400)^T \tag{5}$$

The camera observes a 3D point with known coordinates in a global frame: $X^G = (350, -250, -35)^T$. The corresponding image observation representing this 3D point is $(241.5, 169)$.

   (a) Write an expression for the camera projection matrix P. Write also numerical values of this matrix according to the given data in this exercise.

   (b) Write the underlying equations for projecting a 3D point using the projection matrix P. Calculate the image coordinates by projecting the 3D X^G.

   (c) Calculate the re-projection error.

2. Calibrate your own camera (any camera will do).

   (a) Download the camera calibration toolbox for Matlab .

   (b) Capture images of a calibration pattern and follow the described steps to calibrate your own camera.

   (c) Submit snapshots of the process.

   (d) Write an expression for the estimated camera calibration matrix K. Indicate the principal point, and focal length in each axis.

3. In this exercise we will perform basic image feature extraction and matching. You are free to choose what implementation and programming language to use - one recommended alternative for the latter is the vlfeat library, which provides also interface to Matlab and useful tutorials.

   (a) Use a camera to capture 2 images of yourself from relatively close viewpoints (not more than 30 degrees change in yaw angle)

   (b) Extract SIFT features in each image. Please attach the two images with the extracted features. Indicate feature scale and orientation for a representative feature.

   (c) Calculate putative matches by matching SIFT descriptors as explained in class. Please show, on a separate figure, the calculated matches between the two images. Among these matches, indicate representative inlier and outlier matches.