Exercises: inference in the linear Gaussian model

Joaquín Rapela*

October 12, 2022

1 Inferring location of a static submarine from its sonar measurements

A static submarine is located in a 2D planar surfarce deep in the sea. A priori, we model its unknown location with a 2D Gaussian random variable $\mathbf{z}$ with mean μ_z and covariance Σ_z.

We obtain noisy measurements of the location of the submarine with a sonar. We represent a 2D sonar measurement with random variable $\mathbf{y}_n$, with $p(\mathbf{y}_n|\mathbf{z}) = \mathcal{N}(y_n|\mathbf{z}, \Sigma_y)$.

(a) Sample 100 a priory locations of the submarine (i.e., $\mathbf{z}_1, \ldots, \mathbf{z}_{100}$) using a mean $\mu_z = [2, 3]^\intercal$, a standard deviation along the horizontal direction $\sigma_{\mathbf{z}x} = 1.0$, a standard deviation along the vertical direction $\sigma_{\mathbf{z}y} = 2.0$, and a correlation coefficient between the vertical and horizontal directions $\rho_{\mathbf{z}} = 0.7$.

Plot μ_z, the 95% confidence ellipse for $\mathbf{z}$, and verify that approximately 95% of the samples lie inside the 95% confidence ellipse.

You may want to complete the script doExSubmarine_a.py to address this item.

(b) Select a submarine location $\mathbf{z}$ from those generated in the previous item. Sample $N = 5$ sonar measurements, assuming the submarine is

*j.rapela@ucl.ac.uk

at location $\mathbf{z}$ (i.e., sample from $p(\mathbf{y}|\mathbf{z})$ to obtain $\mathbf{y}_1, \ldots, \mathbf{y}_N$). Use a standard deviation of 1.0 for the measurement noise along the horizontal and vertical directions, and assume that this noise is uncorrelated along these directions.

Plot $\mathbf{z}$, the 95% confidence ellipse for the sonar measurements given that the submarine located at $\mathbf{z}$, and verify that approximately 95% of the samples lie inside the 95% confidence band.

You may want to complete the script doExSubmarine_b.py to address this item.

(c) derive a mathematical expression for the posterior of the submarine location, given sonar measurements; i.e., $p(\mathbf{z}|\mathbf{y}_1, \ldots, \mathbf{y}_N)$.

Hints:

- The posterior of the submarine location given sonar measurements is proportional to the joint distribution of the submaring location and sonar measurments; i.e., $p(\mathbf{z}|\mathbf{y}_1, \ldots \mathbf{y}_N) = \frac{p(\mathbf{y}_1, \ldots, \mathbf{y}_N, \mathbf{z})}{p(\mathbf{y}_1, \ldots, \mathbf{y}_N)} = K\, p(\mathbf{y}_1, \ldots, \mathbf{y}_N, \mathbf{z})$, where K is a value that does not depend on $\mathbf{z}$. Thus, to obtain the posterior we can just keep the terms of the joint that depend on $\mathbf{z}$ and normalize the resulting expression to integrate to one.
- The joint is the product of the likelihood and the prior; i.e., $p(\mathbf{y}_1, \ldots \mathbf{y}_N, \mathbf{z}) = p(\mathbf{y}_1, \ldots, \mathbf{y}_N|\mathbf{z})p(\mathbf{z})$. Thus, to keep the terms of the joint that depend on $\mathbf{z}$, we should just keep the term of the likelihood that depend on $\mathbf{z}$ and combine the result with the prior.
- As shown in Claim. 1, the terms of the likelihood that depend on $\mathbf{z}$ are proportional to a Gaussian distribution with mean $\mathbf{z}$ and covariance $\frac{1}{N}\Sigma$; i.e., $p(\mathbf{y}_1, \ldots, \mathbf{y}_N|\mathbf{z}) = K\, \mathcal{N}(\bar{\mathbf{y}}|\mathbf{z}, \frac{1}{N}\Sigma_y)$, where K is a value that does not depend on $\mathbf{z}$.
- From the previous arguments, to obtain the posterior of $\mathbf{z}$ we can multiply $\mathcal{N}(\bar{\mathbf{y}}|\mathbf{z}, \frac{1}{N}\Sigma_y)$ with the prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mu_z, \Sigma_z)$ and normalize the result. To do this we can use the expression for the posterior of the linear Gaussian model described in class.

(d) plot the sample mean of the measurements, the mean of the posterior, its 95% confidence ellipse, and check if the true submarine location, i.e., $\mathbf{z}$, lies within this ellipse.

You may want to complete the script doExSubmarine_d.py to address this item.

(e) repeat (b) and (d) with $N \in \{3, 50, 200, 500\}$ sonar measurements. How do the posterior estimates change as N increases?

(f) write the expressions of the posterior mean and covariances to see that:

- as the number of measurements increase, the relative contribution of the prior to estimates of the posterior mean and covariance decreases,
- in the limit when the number of measurements approaches infinity, the posterior covariance approaches zero and the posterior mean approaches the measurements sample mean. That is, for an infinite number of measurments, the posterior estimate becomes deterministic and the contribution of the prior to this estimate dissappears.

Can you see these points in the previous simulations?

Claim 1. *If* $P(\mathbf{y}_i|\mathbf{z}) = \mathcal{N}(\mathbf{y}_i|\mathbf{z},\Sigma)$, $i = 1,\ldots,N$,
and $P(\mathbf{y}_1,\ldots,\mathbf{y}_N|\mathbf{z}) = \prod_{i=1}^{N} P(\mathbf{y}_i|\mathbf{z})$,
then $P(\mathbf{y}_1,\ldots,\mathbf{y}_N|\mathbf{z}) = K\mathcal{N}(\bar{\mathbf{y}}_N|\mathbf{z},\frac{1}{N}\Sigma)$

Proof. By induction: $P_n = P(\mathbf{y}_1,\ldots,\mathbf{y}_n|\mathbf{z}) = K\mathcal{N}(\bar{\mathbf{y}}_n|\mathbf{z},\frac{1}{n}\Sigma)$

P_1:

$$P(\mathbf{y}_1|\mathbf{z}) = \mathcal{N}(\mathbf{y}_1|\mathbf{z},\Sigma) = \mathcal{N}(\bar{\mathbf{y}}_1|\mathbf{z},\frac{1}{1}\Sigma)$$

$P_n \to P_{n+1}$:

$$\begin{aligned}
P(\mathbf{y}_1,\ldots,\mathbf{y}_n,\mathbf{y}_{n+1}|\mathbf{z}) &= \prod_{i=1}^{n+1} P(\mathbf{y}_i|\mathbf{z}) \\
&= P(\mathbf{y}_1,\ldots,\mathbf{y}_n|\mathbf{z})P(\mathbf{y}_{n+1}|\mathbf{z}) \\
&= \mathcal{N}(\bar{\mathbf{y}}_n|\mathbf{z},\frac{1}{n}\Sigma)\mathcal{N}(\mathbf{y}_{n+1}|\mathbf{z},\Sigma)
\end{aligned}$$

then

$$\begin{aligned}
\log P(\mathbf{y}_1,\ldots,\mathbf{y}_n,\mathbf{y}_{n+1}|\mathbf{z}) =& K - \frac{1}{2}(\bar{\mathbf{y}}_n - \mathbf{z})^\intercal n\Sigma^{-1}(\bar{\mathbf{y}}_n - \mathbf{z}) - \\
& \frac{1}{2}(\mathbf{y}_{n+1} - \mathbf{z})^\intercal \Sigma^{-1}(\mathbf{y}_{n+1} - \mathbf{z}) \\
=& K_1 - \frac{1}{2}\left(\mathbf{z}^\intercal(n+1)\Sigma^{-1}\mathbf{z} - 2\mathbf{z}^\intercal n\Sigma^{-1}\bar{\mathbf{y}}_n - 2\mathbf{z}^\intercal\Sigma^{-1}\mathbf{y}_{n+1}\right) \\
=& K_1 - \frac{1}{2}\left(\mathbf{z}^\intercal(n+1)\Sigma^{-1}\mathbf{z} - 2\mathbf{z}^\intercal\Sigma^{-1}\sum_{i=1}^{n}\mathbf{y}_i - 2\mathbf{z}^\intercal\Sigma^{-1}\mathbf{y}_{n+1}\right) \\
=& K_1 - \frac{1}{2}\left(\mathbf{z}^\intercal(n+1)\Sigma^{-1}\mathbf{z} - 2\mathbf{z}^\intercal\Sigma^{-1}\sum_{i=1}^{n+1}\mathbf{y}_i\right) \\
=& K_1 - \frac{1}{2}\left(\mathbf{z}^\intercal(n+1)\Sigma^{-1}\mathbf{z} - 2\mathbf{z}^\intercal(n+1)\Sigma^{-1}\bar{\mathbf{y}}_{n+1}\right)
\end{aligned}$$

Therefore

$$P(\mathbf{y}_1,\ldots,\mathbf{y}_n,\mathbf{y}_{n+1}|\mathbf{z}) = K_2\mathcal{N}\left(\mathbf{z}\,\middle|\,\bar{\mathbf{y}}_{n+1},\frac{1}{n+1}\Sigma\right) = K_2\mathcal{N}\left(\bar{\mathbf{y}}_{n+1}\,\middle|\,\mathbf{z},\frac{1}{n+1}\Sigma\right)$$

$\square$