Statistical Machine Learning 2018
Assignment 3
Deadline: 25th of November 2018

Christoph Schmidl
s4226887
c.schmidl@student.ru.nl

Mark Beijer
s4354834
mbeijer@science.ru.nl

November 10, 2018

Exercise 1 - The faulty lighthouse (weight 5)

A lighthouse is somewhere off a piece of straight coastline at a position α along the shore and a distance β out to sea. Due to a technical fault, as it rotates the light source only occasionally and briefly flickers on and off. As a result it emits short, highly focused beams of light at random intervals. These pulses are intercepted on the coast by photo-detectors that record only the fact that a flash has occurred, but not the angle from which it came. So far, N flashes have been recorded at positions $\mathcal{D} = x1, ..., xN$. Where is the lighthouse?

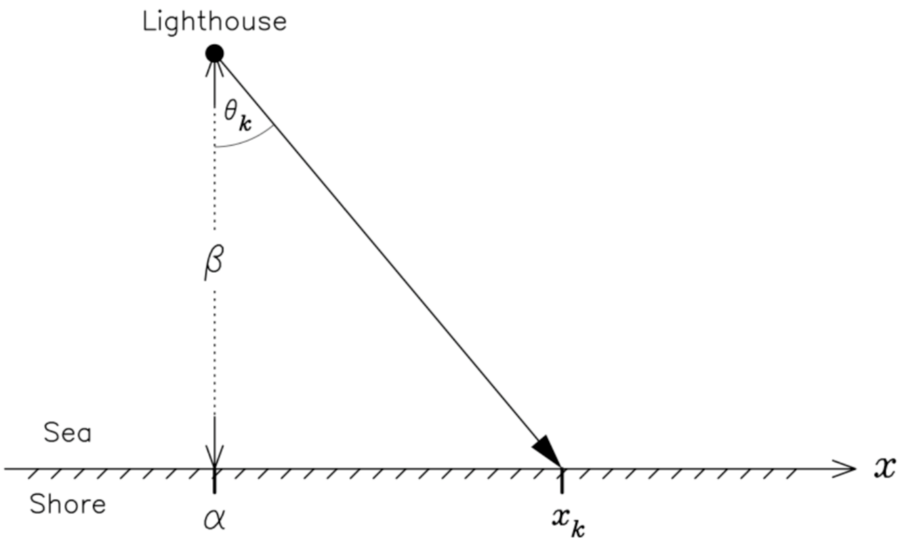

Figure 1: Geometry of the lighthouse problem.

Part 1 - Constructing the model

1.1.1

Let θ_k be the (unknown) angle for the k-th recorded flash, see fig.1. Argue why

$$p(\theta_k|\alpha, \beta) = \frac{1}{\pi} \tag{1}$$

would be a reasonable distribution over θ_k between $\pm\frac{\pi}{2}$ (zero otherwise).

Answer:

1.1.2

We only have the position x_k of the detector that recorded flash k, but we can relate this to the unknown θ_k via elementary geometry as

$$\beta \tan(\theta_k) = x_k - \alpha \tag{2}$$

Show that the expected distribution over x given α and β can be written as

$$p(x_k|\alpha, \beta) = \frac{\beta}{\pi[\beta^2 + (x_k - \alpha)^2]} \tag{3}$$

by using (2) to substitute variable x_k for θ_k in the distribution (1). Plot the distribution for $\beta = 1$ and a particular value of α.
Hint: use the Jacobian $|\frac{d\theta}{dx}|$ (Bishop, p.18) and the fact that $(tan^{-1}x)' = \frac{1}{1+x^2}$.

Answer:

1.1.3

Inferring the position of the lighthouse corresponds to estimating α and β from the data $\mathcal{D}$. This is still quite difficult, but if we assume that β is known, then from Bayes' theorem we know that $p(\alpha|\mathcal{D}, \beta) \propto p(\mathcal{D}|\alpha, \beta)p(\alpha|\beta)$. We have no a priori knowledge about the position α along the coast other than that it should not depend on the distance out at sea.

Show that with these assumptions the log of the posterior density can be written as

$$L = \ln(p(\alpha|\mathcal{D}, \beta)) = \texttt{constant} - \sum_{k=1}^{N} \ln[\beta^2 + (x_k - \alpha)^2] \tag{4}$$

and give an expression for the value $\hat{\alpha}$ that maximizes this posterior density.

Answer:

1.1.4

Suppose we have a data set (in km) of $\mathcal{D} = \{3.6, 7.7, -2.6, 4.9, -2.3, 0.2, -7.3, 4.4, 7.3, -5.7\}$. We also assume that the distance β from the shore is known to be 2 km. As it is difficult to find a simple expression for the value of $\hat{\alpha}$ that maximizes (4), we try an alternative approach instead.

Plot $p(\alpha|\mathcal{D}, \beta = 2)$ as a function of α over the interval $[-10, 10]$. What is your most likely estimate for $\hat{\alpha}$ based on this graph? Compare with the mean estimate of the dataset. Can you explain the difference?

Answer: