Advanced Scientific Computing: Stochastic Methods for Data Analysis, Inference and Optimization

P. Koumoutsakos
323 Pierce Hall, 29 Oxford Street Cambridge,
MA 02138

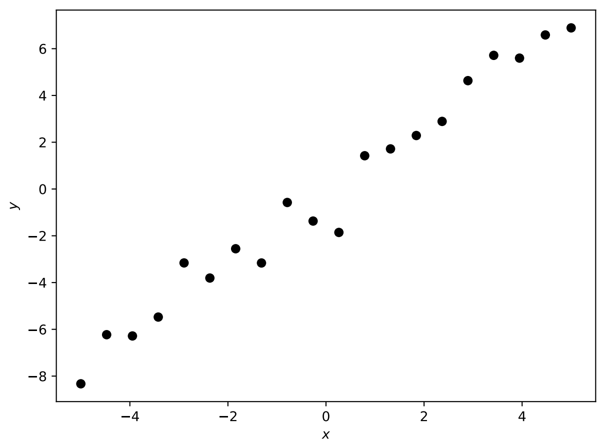
Fall semester 2022

Set 4 - Bayesian inference

Issued: November 2, 2022
Hand in: November 17, 2022 00:00

Question 1: Bayesian inference of linear model

An experimentalist collected data $D = \{x_i, y_i\}_{i=1}^{N}$, reported on the figure below (also available in the file `data.csv`).



A model of this data is given by

$$y_i = ax_i + \sigma\xi_i,$$

where a and σ are scalars and ξ_i are i.i.d. normally distributed random variables, $\xi_i \sim \mathcal{N}(0, 1)$.

a) Write the log-likelihood of the data given the parameters $\theta = (a, \sigma)$. Find the parameters that maximize the log-likelihood.

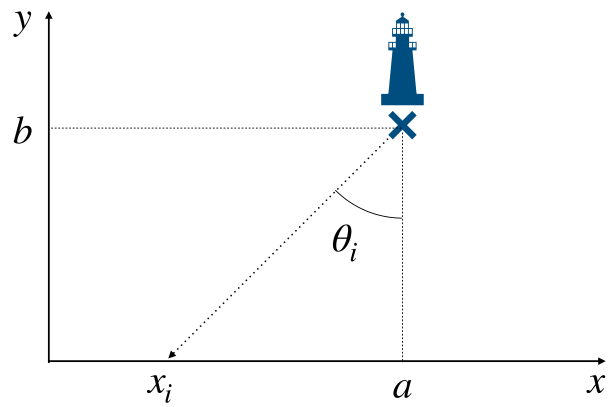
b) Express the posterior distribution of the parameters. Assume uniform priors with large enough bounds for a and $\sigma > 0$. Approximate the posterior distribution of the parameters with the Laplace approximation.

c) Use the Metropolis-Hasting algorithm to sample the posterior distribution of the parameters. Compare the obtained density with the laplace approximation on a plot (show on a (a, σ) space a thousand samples, the contours of the log-posterior and the contours of the log-Laplace approximation).

d) Use the samples of the posterior obtained in the previous subquestions to plot the predictions of the model together with the data. The plot should include the median prediction, the 50% and 90% confidence intervals.

Question 2: Where is the beacon?

A beacon positioned at unknown location (a, b) emits light in random directions $\theta \sim \mathcal{U}(-\pi/2, \pi/2)$. The light is detected by sensors located on the shore $(y = 0)$. The light detections are recorded at coordinates x_i, $i = 1, 2, \ldots, N$. Given these coordinates, where is the beacon?



a) Write x_i as a function of θ_i, the angle at which the light was emitted to produce x_i. Derive the likelihood of the data x_i, $i = 1, 2, \ldots, N$ given the location of the data. Hint: x_i is the transformation of the random variable θ_i.

b) Write down the posterior distribution of (a, b). Choose appropriate priors for the parameters.

c) The file `data.csv` contains the list of impacts coordinates x_i. Use Korali to sample the posterior distribution of (a, b) given this data. Report the mean of the parameters and show the posterior density of the beacon's position.