Assignment for Lecture 9

BAYESIAN CURVE FITTING AND VARIANCE REDUCTION

Lecture Date: 4/29/2026

"C" denotes for "computational" problems, language suggestion: Python/Julia

please include codes and results with analyses for computational problems

please write in pdf format and submit to `bjcai@fudan.edu.cn` ***before the lecture of 5/6/2026***

1. Assume that when $\theta = 1$, the random variable x follows the distribution $\mathcal{N}(1, \sigma^2)$, and when $\theta = 2$, it follows $\mathcal{N}(2, \sigma^2)$.

   (a) For $\sigma = 2$, write down the pdf of x and plot it. Under the same condition $\sigma = 2$, what is the value of $p(\theta = 1|x = 1)$?

   (b) As the parameter σ varies, how does the posterior distribution for the parameter θ, i.e., $p(\theta = 1|x = 1)$, change?

2. [C] Program the Bayesian linear regression example discussed in the lecture. The ground-truth physical model is $f_{\text{phys}}(x) = w_0' + w_1' x$ with $w_0' = -0.2$ and $w_1' = 0.5$. Generate $m = 10$ data samples with $x^{(i)} \sim \text{Unif}[-1, 1]$, $y^{(i)} = f_{\text{phys}}(x^{(i)}) + a_\delta^{(i)}$ and $a_\delta^{(i)} \sim \mathcal{N}(0, 0.5^2)$. Use the learning model $f_{w_0,w_1}(x) = w_0 + w_1 x$, with likelihood precision $\beta = (1/0.5)^2 = 4$ and prior precision $\alpha = 2.0$. *Write a program to infer the posterior distribution of* $\mathbf{w} = (w_0, w_1)^\top$. *Plot the data, the ground-truth line, the Bayesian mean prediction, and the posterior distribution in the* (w_0, w_1) *plane. Repeat the calculation for several values of* m *and explain how the posterior becomes more concentrated around the true parameters as the number of data samples increases.*

3. [C] Let $\mathbf{x}^{(i)} = (x_1^{(i)}, \cdots, x_d^{(i)})^\top$ be the vector $\mathbf{x}$ at the ith iteration. At the $(i+1)$th iteration, we draw $x_j^{(i+1)}$ from the conditional distribution,

$$p(x_j | x_1^{(i+1)}, \cdots, x_{j-1}^{(i+1)}, x_{j+1}^{(i)}, \cdots, x_d^{(i)}), \tag{1-1}$$

   for $j = 1 \sim d$. For the 3-dimensional Gaussian distribution,

$$\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix}\right), \tag{1-2}$$

   where ρ_{ij} is the correlation between x_i and x_j, sample $\mathbf{x}$ from the initial point $\mathbf{x}^{(0)} = (1, 1, 1)^\top$ with $\rho_{12} = 0.2, \rho_{13} = 0.8$ and $\rho_{23} = 0.5$.