

ECE368: Probabilistic Reasoning

Lab 2: Bayesian Linear Regression

Name: Pratyush Menon

Student Number: 1004282661

You should hand in: 1) A scanned .pdf version of this sheet with your answers (file size should be under 2 MB); 2) four figures for Question 2 and three figures for Question 4 in the .pdf format; and 3) one Python file `regression.py` that contains your code. All these files should be uploaded to Quercus.

- Express the posterior distribution $p(\mathbf{a}|x_1, z_1, \dots, x_N, z_N)$ using $\sigma^2, \beta, x_1, z_1, x_2, z_2, \dots, x_N, z_N$. (1 pt)

$$\begin{aligned}
 X &= \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} & z &= \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix} \\
 \mu &= (X^T X + \frac{\sigma^2}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})^{-1} X^T z \\
 \Sigma &= (X^T X + \frac{\sigma^2}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})^{-1} \sigma^2 \\
 p(a|x_1, z_1, \dots, x_N, z_N) &\sim \mathcal{N}(\mu, \Sigma)
 \end{aligned}$$

- Let $\sigma^2 = 0.1$ and $\beta = 1$. Draw four contour plots corresponding to the distributions $p(\mathbf{a})$, $p(\mathbf{a}|x_1, z_1)$, $p(\mathbf{a}|x_1, z_1, \dots, x_5, z_5)$, and $p(\mathbf{a}|x_1, z_1, \dots, x_{100}, z_{100})$. In all contour plots, the x-axis represents a_0 , and the y-axis represents a_1 . Please save the figures with names **prior.pdf**, **posterior1.pdf**, **posterior5.pdf**, **posterior100.pdf**, respectively. (1.5 pt)
- Suppose that there is a new input x , for which we want to predict the corresponding target value z . Write down the distribution of the prediction z , i.e. $p(z|x, x_1, z_1, \dots, x_N, z_N)$. (1 pt)

$$\begin{aligned}
 X &= \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} & z &= \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix} & x_{\text{new}} &= \begin{bmatrix} 1 & x \end{bmatrix} \\
 \mu &= x_{\text{new}} \left((X^T X + \frac{\sigma^2}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})^{-1} (X^T z) \right) \\
 \Sigma &= (x_{\text{new}}^T (X^T X + \frac{\sigma^2}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})^{-1} \sigma^2) (x_{\text{new}}) + \sigma^2 \\
 p(z|x, x_1, z_1, \dots, x_N, z_N) &\sim \mathcal{N}(\mu, \Sigma)
 \end{aligned}$$

- Let $\sigma^2 = 0.1$ and $\beta = 1$. Given a set of new inputs $\{-4, -3.8, \dots, 3.8, 4\}$, plot three figures, whose x-axis is the input and y-axis is the prediction, corresponding to three cases:
 - The predictions are based on one training sample, i.e., based on $p(z|x, x_1, z_1)$.
 - The predictions are based on 5 training samples, i.e., based on $p(z|x, x_1, z_1, \dots, x_5, z_5)$.
 - The predictions are based on 100 training samples, i.e., based on $p(z|x, x_1, z_1, \dots, x_{100}, z_{100})$.

The range of each figure is set as $[-4, 4] \times [-4, 4]$. Each figure should contain the following three components: 1) the new inputs and the corresponding predicted targets; 2) a vertical interval at each predicted target, indicating the range within one standard deviation; 3) the training sample(s) that are used for the prediction. Use `plt.errorbar` for 1) and 2); use `plt.scatter` for 3). Please save the figures with names **predict1.pdf**, **predict5.pdf**, **predict100.pdf**, respectively. (1.5 pt)