

# ECE368: Probabilistic Reasoning

## Lab 2: Bayesian Linear Regression

Name: Jiani Li

Student Number: 1003847867

**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.

1. 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)

$$\beta = \begin{bmatrix} 1 \\ \vdots \\ x_1 \\ \vdots \\ x_N \end{bmatrix} \quad \Sigma_{\mathbf{a}}^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{bmatrix} \quad \Sigma_{\mathbf{w}}^{-1} = \frac{1}{\sigma^2} \quad \mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}$$

$$p(\mathbf{a} | x_1, \dots, z_N) \sim \mathcal{N}(\mu_{\mathbf{a}} | x_1, \dots, z_N, \Sigma_{\mathbf{a}} | x_1, \dots, z_N)$$

where  $\mu_{\mathbf{a}} | x_1, \dots, z_N = (\Sigma_{\mathbf{a}}^{-1} + \beta^T \Sigma_{\mathbf{w}}^{-1} \beta)^{-1} (\beta^T \Sigma_{\mathbf{w}}^{-1} \mathbf{z})$

$$\Sigma_{\mathbf{a}} | x_1, \dots, z_N = (\Sigma_{\mathbf{a}}^{-1} + \beta^T \Sigma_{\mathbf{w}}^{-1} \beta)^{-1}$$

2. 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)
3. 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)

Correction of  $\beta$ :

$$\beta = \begin{bmatrix} 1 \\ x \end{bmatrix}$$

where  $x$  is the new input data

$$p(z|x, x_1, z_1, \dots, x_N, z_N) \sim \mathcal{N}(\mu_{z|x, x_1, \dots, z_N}, \Sigma_{z|x, x_1, \dots, z_N})$$

$$\mu_{z|x, x_1, \dots, z_N} = \beta \cdot \mu_{\mathbf{a}} | x_1, \dots, z_N \quad \text{where } \beta \text{ and } \mu_{\mathbf{a}} | x_1, \dots, z_N \text{ are from 1. } \uparrow \text{ wrong}$$

$$\Sigma_{z|x, x_1, \dots, z_N} = \Sigma_{\mathbf{w}} + \beta \cdot \Sigma_{\mathbf{a}} | x_1, \dots, z_N \beta^T \quad \text{where } \Sigma_{\mathbf{a}} | x_1, \dots, z_N \text{ from 1, } \Sigma_{\mathbf{w}} = \frac{1}{\sigma^2}$$

4. 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)