ECE368: Probabilistic Reasoning

Lab 2 – Part I: Bayesian Linear Regression

Name: Calvin Kwei Hoi Tan	**Student Number:**

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 $p(\mathbf{a}|x_1, z_1, \ldots, x_N, z_N)$ using $\sigma^2, \beta, x_1, z_1, x_2, z_2, \ldots, x_N, z_N$. (1 pt)

Given $[1\ x]\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} + w$

A $\underline{a}$

$\underline{a} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}\right)$

$\mu_{\underline{a}}$ $\Sigma_{\underline{a}\underline{a}}$

$w \sim (0, \sigma^2)$

μ_w Σ_{ww}

$$p(\underline{a}|x_1, z_1, \ldots, x_N, z_N) = N(\mu_{\underline{a}|x_1,z_1,\ldots,x_N,z_N}, \Sigma_{\underline{a}|x_1,\ldots,z_N})$$

where...

$$\mu_{\underline{a}|\ldots} = (\Sigma_{\underline{a}\underline{a}}^{-1} + A^T\Sigma_{ww}^{-1}A)^{-1}(A^T\Sigma_{ww}^{-1}(\underline{z}-\underline{b}) + \Sigma_{\underline{a}\underline{a}}^{-1}\mu_{\underline{a}})$$

$$= (\Sigma_{\underline{a}\underline{a}}^{-1} + A^T\Sigma_{ww}^{-1}A)^{-1}(A^T\Sigma_{ww}^{-1}\underline{z})$$

$A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}, \mu_{\underline{a}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mu_w = 0$

and

$$\Sigma_{\underline{a}|\ldots} = (\Sigma_{\underline{a}\underline{a}}^{-1} + A^T\Sigma_{ww}^{-1}A)^{-1}$$

$\underline{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}, b = \underline{0}, \Sigma_{\underline{a}\underline{a}} = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}, \Sigma_{ww} = \sigma^2$

2. Let $\sigma^2 = 0.1$ and $\beta = 1$. Draw four contour plots corresponding to the distributions $p(\mathbf{a})$, $p(\mathbf{a}|x_1, z_1)$, $p(\mathbf{a}|x_1, z_1, \ldots, x_5, z_5)$, and $p(\mathbf{a}|x_1, z_1, \ldots, x_{100}, z_{100})$. 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, $p(z|x, x_1, z_1, \ldots, x_N, z_N)$. (1 pt)

$$p(z|x, x_1, z_1, \ldots, x_N, z_N) = N(\underline{z}|\mu_z, \Sigma_z)$$

x new data

new data

where...

$$\mu_z = A\mu_{\underline{a}|x_1,z_1,\ldots,x_N,z_N}$$

$$\Sigma_z = \Sigma_w + A\Sigma_{\underline{a}|x_1,z_1,\ldots,x_N,z_N}A^T$$

$A = [1\ x]$, $\mu_{\underline{a}|\ldots}$ from above

$\Sigma_w = \sigma^2$, $\Sigma_{\underline{a}|\ldots}$ from above

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

   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)