ECE368: Probabilistic Reasoning
Lab 2: Bayesian Linear Regression

Name: Michael Vu	**Student Number:** 1002473272

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 σ^2, β, $x_1, z_1, x_2, z_2, \ldots, x_N, z_N$. (1 **pt**)

$$P(a) = \frac{1}{2\pi\beta}\exp(-(\frac{a_0^2 + a_1^2}{2\beta})) = \frac{1}{2\pi\beta}\exp(-\frac{||a||^2}{2\beta})$$

$$P(x_i, z_i|a) = P(z_i|x_i, a) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(z_i - a_1 x_i - a_0)^2}{2\sigma^2})$$

$$P(x_i, z_i) = P(z_i|x_i) = \int P(z_i|x_i, a)P(a)da = \Gamma_i$$

$$P(a|x, z) = \frac{\frac{1}{2\pi\beta}(\frac{1}{\sqrt{2\pi\sigma^2}})^N \exp(-\frac{||a||^2}{2\beta} - \frac{1}{2\sigma^2}\sum_{i=1}^N (z_i - a_1 x_i - a_0)^2)}{\prod_{i=1}^N \Gamma_i}$$

$$= C\exp\Big(-\frac{||a||^2}{2\beta} - \frac{1}{2\sigma^2}\sum_{i=1}^N (z_i - a_1 x_i - a_0)^2\Big)$$

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**)

$$\text{let } D = \{x_1, z_1, \ldots, x_N, z_N\}$$

$$\text{let } \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$$

$$\hat{\mathbf{a}} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = (\mathbf{X}^T\mathbf{X} + \frac{\sigma^2}{\beta}I)^{-1}\mathbf{X}^T\mathbf{z}$$

$$P(z|x, D) = P(a_1 x + a_0 + w|x, D) \sim \mathcal{N}(a_1 x + a_0, \sigma^2)$$

$$P(z|x, D) = \frac{1}{\sqrt{2\pi}\sigma}\exp\Big(-\frac{1}{2\sigma^2}(z - a_1 x - a_0)^2\Big)$$

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**)