ECE368: Probabilistic Reasoning
Lab 2: Bayesian Linear Regression

Name: Armaan Lalani **Student Number:** 1005023225

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 posterior distribution $p(\mathbf{a}|z_1, \ldots, z_N; x_1, \ldots, x_N)$ using σ^2, β, $x_1, z_1, x_2, z_2, \ldots, x_N, z_N$. (1 **pt**)

Let $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix}$, $\Sigma_a^{-1} = \begin{bmatrix} 1/\beta & 0 \\ 0 & 1/\beta \end{bmatrix}$, $\Sigma_w^{-1} = \frac{1}{\sigma^2}$, $Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$

$p(a \mid z_1, \ldots z_n; x_1 \ldots x_n) \sim N(\mu_{a|x_1, \ldots z_n}, \Sigma_{a|x_1, \ldots z_n})$

$\mu_{a|x_1, \ldots z_n} = (\Sigma_a^{-1} + A^T \Sigma_w^{-1} A)^{-1} (A^T \Sigma_w^{-1} Z)$, $\Sigma_{a|x_1, \ldots z_n} = (\Sigma_a^{-1} + A^T \Sigma_w^{-1} A)^{-1}$

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

Let $A = [1 \;\; x]$, $\Sigma_w = \sigma^2$, and all other variables hold their same value as in question 1.

$p(z \mid x, x_1, \ldots z_N) \sim N(\mu_{z|x, x_1, \ldots z_n}, \Sigma_{z|x, x_1, \ldots z_n})$

$\mu_{z|x, x_1, \ldots z_n} = A \cdot \mu_{a|x_1, \ldots z_N}$, $\Sigma_{z|x, x_1, \ldots z_n} = \Sigma_w + A \Sigma_{a|x_1, \ldots z_n} A^T$

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 $p(z|z_1; x, x_1,)$.
   (b) The predictions are based on 5 training samples, i.e., based on $p(z|z_1, \ldots, z_5; x, x_1, \ldots, x_5)$.
   (c) The predictions are based on 100 training samples, i.e., based on $p(z|z_1, \ldots, z_{100}; x, x_1, \ldots, x_{100})$.

   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**)