ECE368: Probabilistic Reasoning

Lab 2: Bayesian Linear Regression

Name: Jia Ming Huang **Student Number:** 1003893245.

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**)

$p(\vec{a}|x_1, z_1, \ldots, x_N, z_N) = N(\mu_z, \Sigma_z)$.

$A = [\vec{1}\ \vec{x}] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$ $\Sigma_{ww} = \sigma^2$ $\vec{b} = \vec{0}$

$\Sigma_z = \Sigma_{\vec{a}|x_1, z_1, \ldots, x_N, z_N} = (\Sigma_{\vec{a}\vec{a}}^{-1} + A^T \Sigma_{ww}^{-1} A)^{-1}$

$\Sigma_{\vec{a}\vec{a}} = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}$ $\vec{z} = [z_1\ z_2 \cdots z_N]^T$.

$\mu_z = \mu_{\vec{a}|x_1, z_1, \ldots, x_N, z_N} = (\Sigma_{\vec{a}\vec{a}}^{-1} + A^T \cdot \Sigma_{ww}^{-1} A)^{-1} \cdot (A^T \Sigma_{ww}^{-1} (\vec{z} - \vec{b}) + \Sigma_{\vec{a}\vec{a}}^{-1} \vec{\mu}_{\vec{a}})$ $\vec{\mu}_a = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

So, $p(\vec{a}|x_1, z_1, \ldots, x_N, z_N)$ $= \Sigma_{\vec{a}|x_1, z_1, \ldots, x_N, z_N} \cdot (A^T \Sigma_{ww}^{-1} \cdot \vec{z})$.

$= N\left(\left(\begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}^{-1} + [\vec{1}\ \vec{x}]^T \cdot \frac{1}{\sigma^2} \cdot [\vec{1}\ \vec{x}]\right)^{-1} \cdot \left([\vec{1}\ \vec{x}]^T \cdot \frac{1}{\sigma^2} \cdot [z_1\ z_2 \cdots z_N]^T\right), \left(\begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}^{-1} + [\vec{1}\ \vec{x}]^T \cdot \frac{1}{\sigma^2} \cdot [\vec{1}\ \vec{x}]\right)^{-1}\right)$.

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**)

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

$\mu_z' = A' \cdot \mu_z$

$\Sigma_z' = \Sigma_{ww} + A' \cdot \Sigma_z \cdot A'^T$. $\Rightarrow A' = [1\ x]$

Recall that $\Sigma_z = (\Sigma_{\vec{a}\vec{a}}^{-1} + A^T \Sigma_{ww}^{-1} A)^{-1}$, $\Sigma_{ww} = \sigma^2$.

and, $\mu_z = \Sigma_z \cdot (A^T \Sigma_{ww}^{-1} (\vec{z} - \vec{b}) + \Sigma_{\vec{a}\vec{a}}^{-1} \vec{\mu}_{\vec{a}})$.

So, $p(\vec{z}|x, x_1, z_1, \ldots, x_N, z_N)$

$= N([1\ x] \cdot \mu_z,$

$\sigma^2 + [1\ x] \cdot \Sigma_z \cdot [1\ x]^T)$.

Note that $\vec{z}|x, x_1, z_1, \ldots, x_N, z_N = \mu_z'$.

and $\mu_z' = A' \cdot \mu_{\vec{a}|x_1, z_1, \ldots, x_N, z_N} = A' \cdot \mu_z$.

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**)