ECE368: Probabilistic Reasoning

Lab 2: Bayesian Linear Regression

Name: VARUN Sanpat	Student Number: 1003859602

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)

$z_n = [a_1 \; a_0]\begin{bmatrix} x \\ 1 \end{bmatrix} + w$ $\quad \rightarrow \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ 1 & 1 & 1 & & 1 \end{bmatrix}$

$\hat{\mu}_a = \mu_{a|x_1,z_1,\ldots x_N,z_N} = \left(X^T X + \frac{\sigma^2}{\beta^2} I_2\right)^{-1} X^T Z$

$\hat{\Sigma}_a = \Sigma_{a|x_1,z_1,\ldots x_N,z_n} = \left(X^T X + \frac{\sigma^2}{\beta^2} I\right)^{-1} \sigma^2$

$p(a|x_1,z_1,\ldots x_N,z_N) \sim N(\hat{\mu}_a, \hat{\Sigma}_a)$

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)

$z_n = [a_1 \; a_0]\begin{bmatrix} x \\ 1 \end{bmatrix} + w$, but $w \sim N(0, \sigma^2)$, so w does not affect mean of z.

$\therefore \hat{\mu}_z = E[z] = E[a_1 x] + E[a_0] + E[w] = \hat{\mu}_a^T \cdot x \rightarrow [\hat{a}_{1\,map} \; \hat{a}_{0\,map}]$.

$z \sim N(\hat{\mu}_z, x^T \underbrace{\hat{\Sigma}_a}_{} x)$

$\rightarrow \Sigma_{a|x_1,z_1,\ldots x_n,z_n}$

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)