$n \times 2 \quad 2 \times 1$
$n \times 1$

ECE368: Probabilistic Reasoning

Lab 2: Bayesian Linear Regression

Name: Raymond Hong	**Student Number:** 1003942629

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

$z = a_1 x + a_0 + w$

$\underline{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$

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

$\underline{z} = X\underline{a} + \underline{w}$

$(n \times 2) \; (2 \times 1)$

$(n \times 1)$

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)

$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \quad \Sigma_a^{-1} = \begin{bmatrix} 1/\beta & 0 \\ 0 & 1/\beta \end{bmatrix} \quad \Sigma_w^{-1} = 1/\sigma^2 \quad z = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}$

$\mu_{\underline{a}|x_1, z_1, \ldots x_N, z_N} = (\Sigma_a^{-1} + X^T \Sigma_w^{-1} X)^{-1} (X^T \Sigma_w^{-1} z)$

$\Sigma_{\underline{a}|x_1 z_1 \ldots x_N, z_N} = (\Sigma_a^{-1} + X^T \Sigma_w^{-1} X)^{-1}$

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

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)

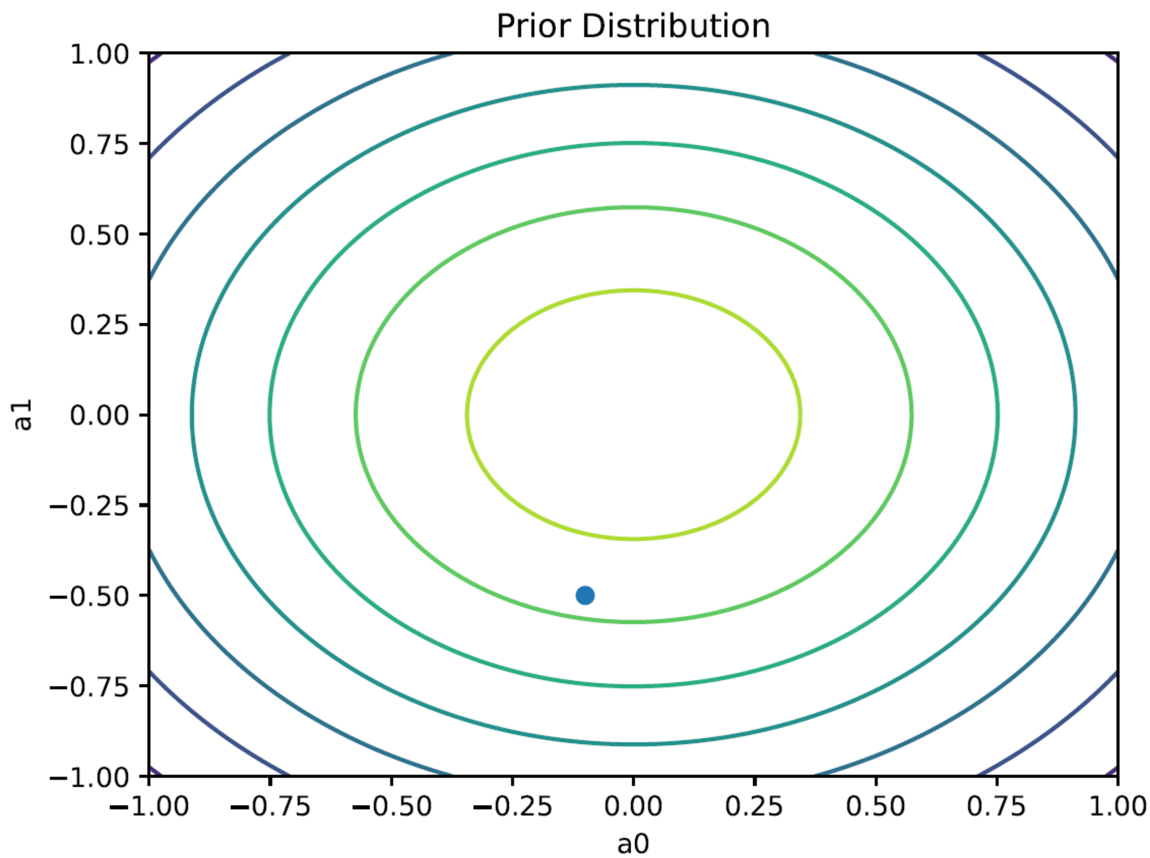
$\mu_{z|x, x_1, z_1, \ldots x_N, z_N} = X \cdot \mu_{\underline{a}|x_1, z_1 \ldots x_N, z_N} \quad X = [1 \;\; x] \quad \mu_{\underline{a}|x_1, z_1 \ldots x_N, z_N}$ from part (1)

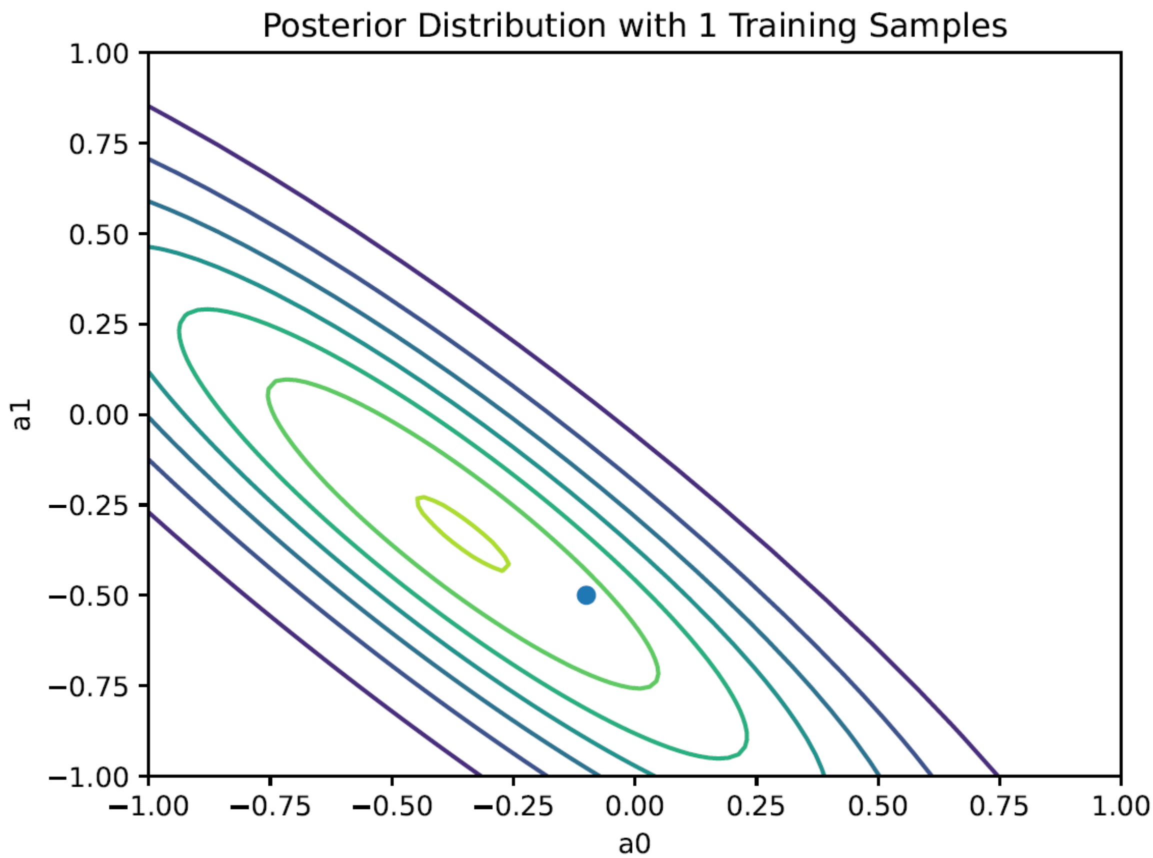
$\Sigma_{z|x, x_1, z_1, \ldots x_N, z_N} = \Sigma_w + X \Sigma_{\underline{a}|x_1, z_1, \ldots x_N, z_N} X^T \quad \Sigma_w = \sigma^2 \quad \Sigma_{\underline{a}|x_1, z_1 \ldots x_N, z_N}$ from part (1)

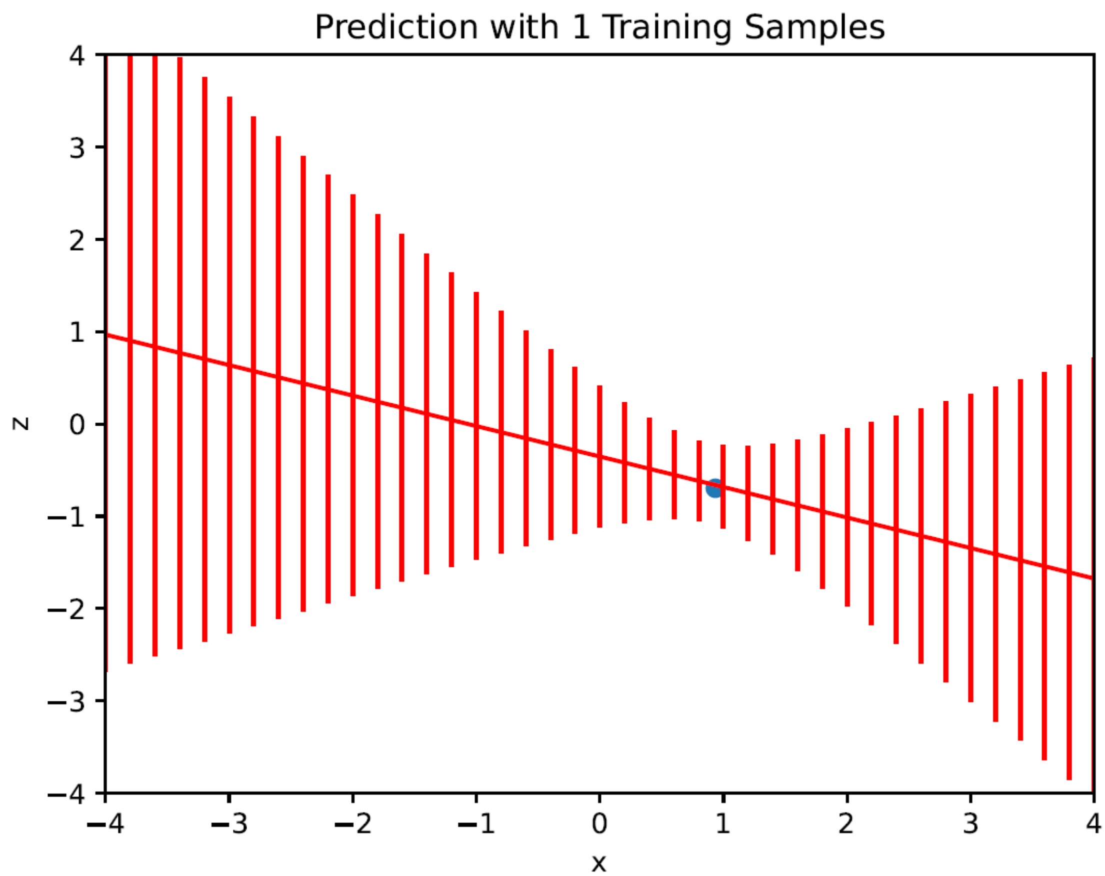
$p(z|x, x_1, z_1 \ldots x_N, z_N) \sim N(\mu_{z|x, x_1 \ldots z_N}, \Sigma_{z|x, x_1, z_1 \ldots x_N, z_N})$

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 $p(z|x, x_1, z_1, \ldots, x_{100}, z_{100})$.

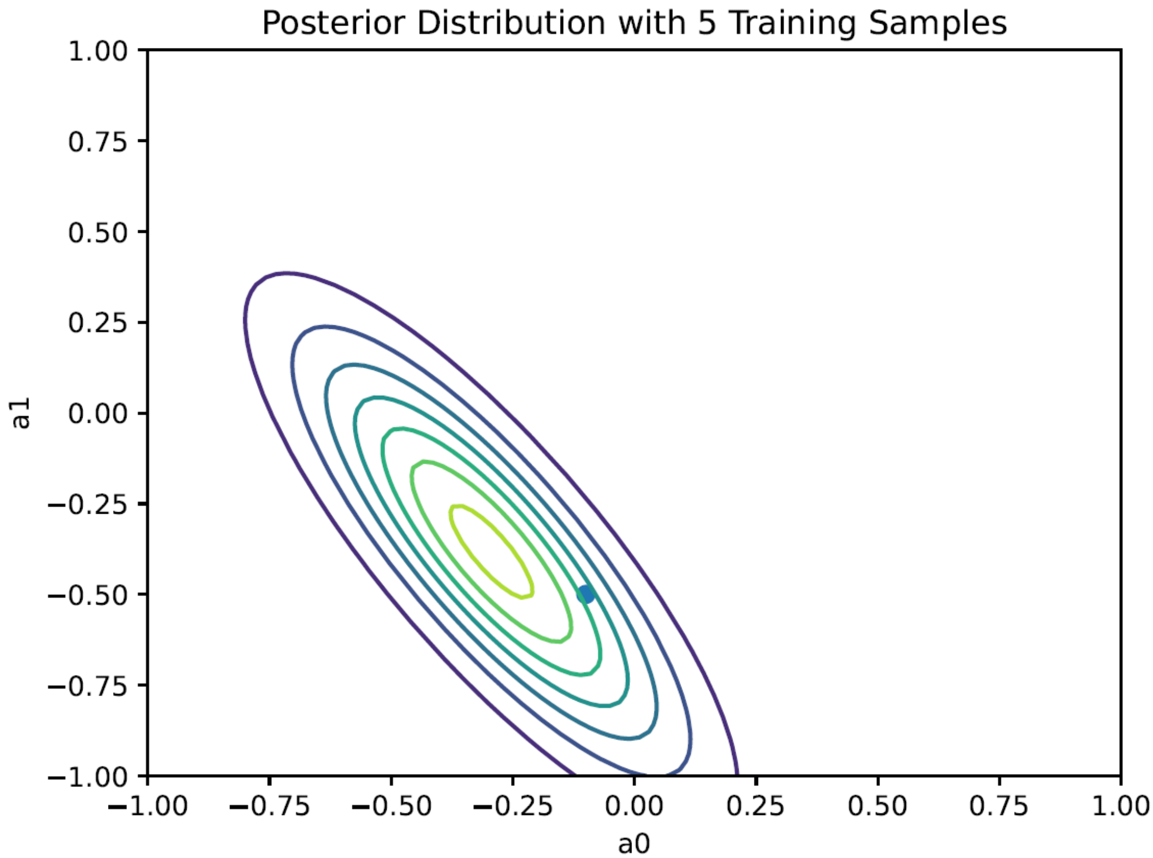
   The range of each figure is set as $[-4, 4] \times [-4, 4]$. Each figure should contain the following three components: 1) the new inputs and the corresponding predicted targets; 2) a vertical interval at each predicted target, indicating the range within one standard deviation; 3) the training sample(s) 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)


Prior Distribution
a1
a0
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00

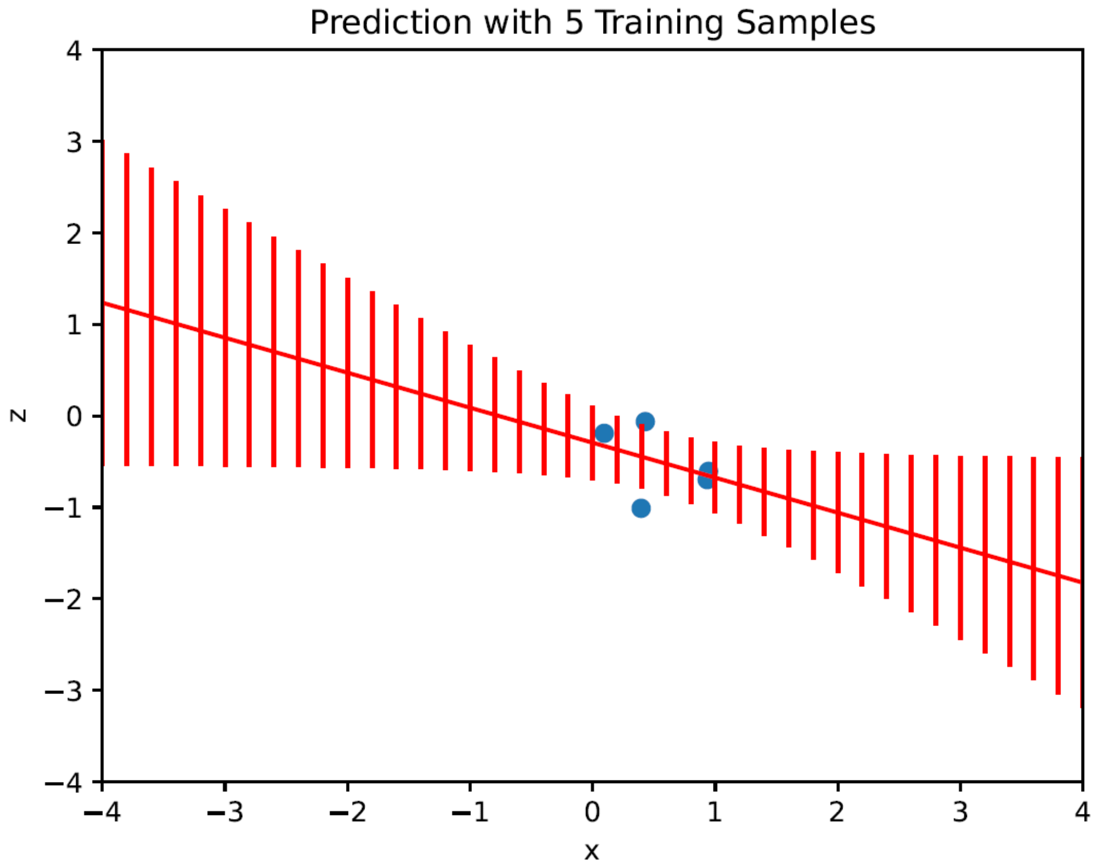

Posterior Distribution with 1 Training Samples
a1
a0
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00



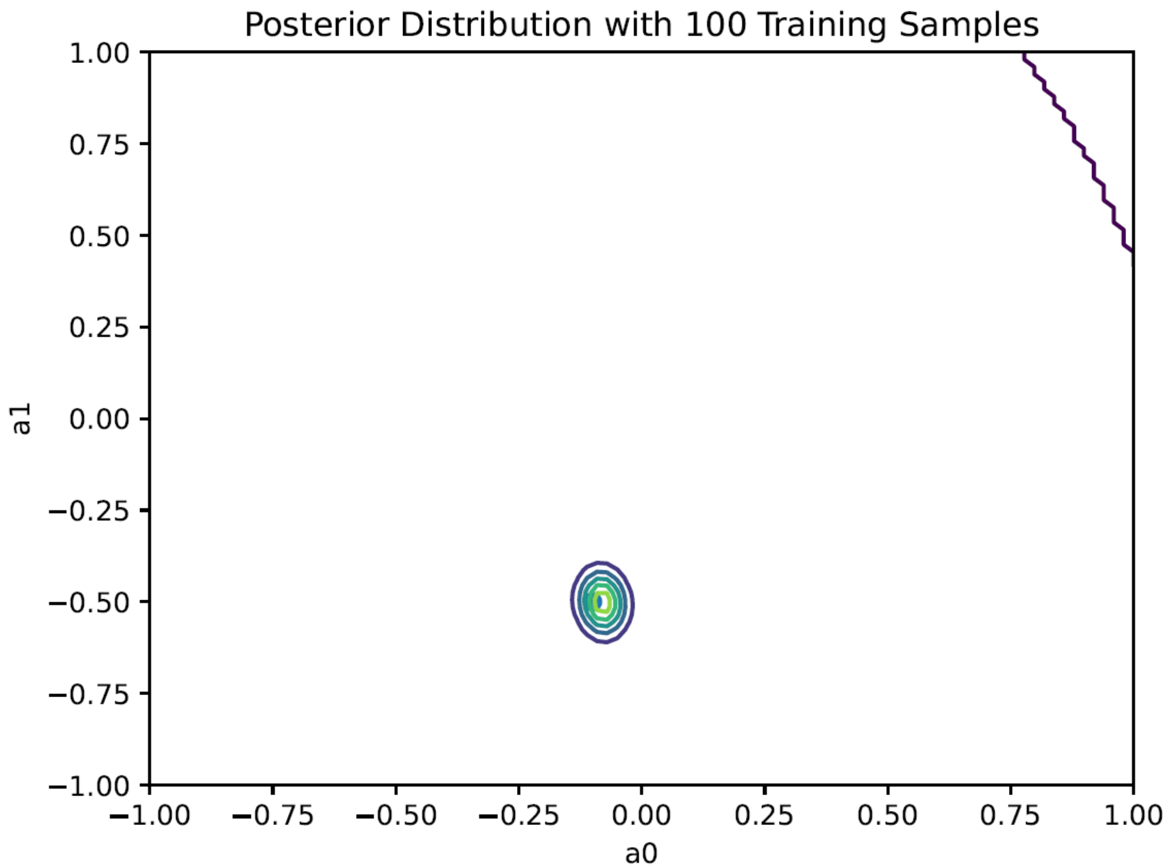
Prediction with 1 Training Samples
z
x
4
3
2
1
0
−1
−2
−3
−4
−4
−3
−2
−1
0
1
2
3
4

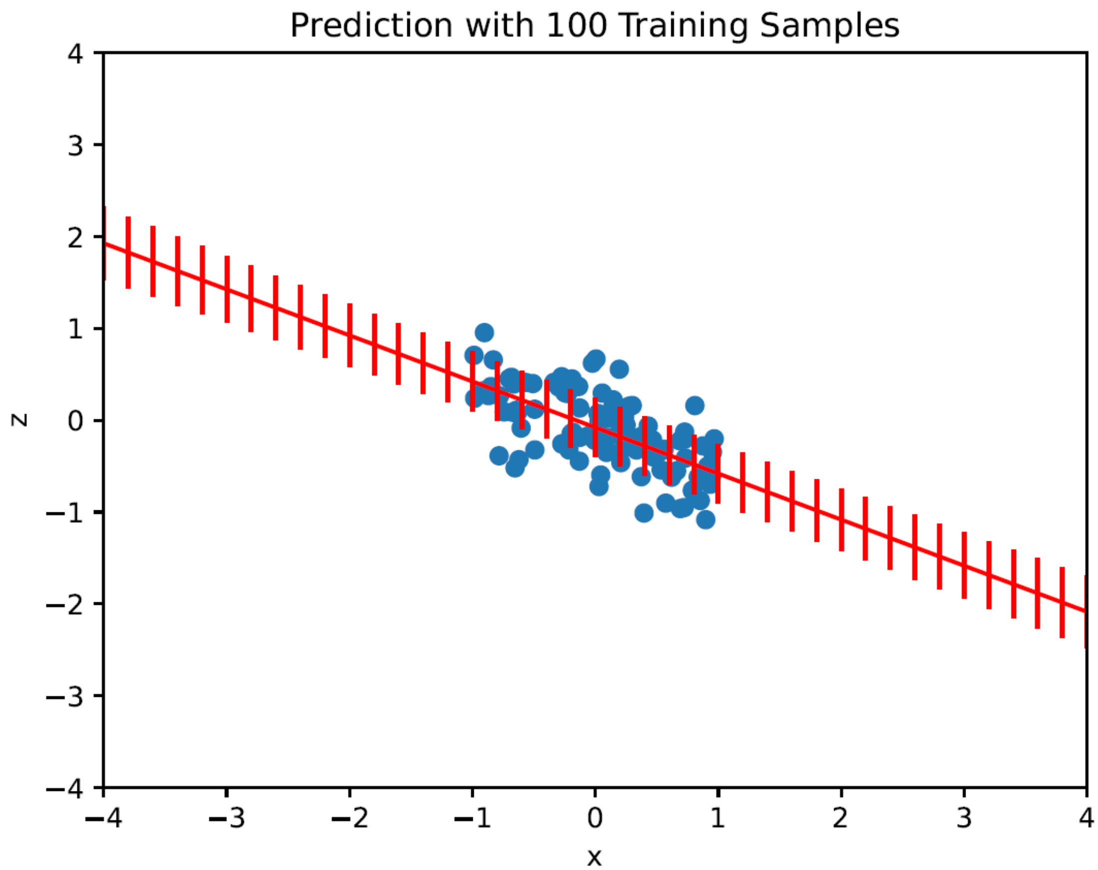


Posterior Distribution with 5 Training Samples
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
a1
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
a0



Prediction with 5 Training Samples
4
3
2
1
0
−1
−2
−3
−4
z
−4
−3
−2
−1
0
1
2
3
4
x


Posterior Distribution with 100 Training Samples
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
a1
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
a0



Prediction with 100 Training Samples
4
3
2
1
0
−1
−2
−3
−4
z
−4
−3
−2
−1
0
1
2
3
4
x