Lab2 Answer

Tuesday, March 22, 2022 9:08 AM

ECE368: Probabilistic Reasoning
Lab 2: Bayesian Linear Regression

Name: RUI ZENG	Student Number: 1003979091

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**)

$Z = [1 \quad x] \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} + W$

$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} + W$

Gaussian system :

$y = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \quad A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \quad x = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \quad z = w$

$x \sim N(\mu_x, \Sigma_x) \quad \mu_x = \mu_a = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Sigma_x = \Sigma_a = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}$

$z \sim N(0, \Sigma_z) \quad \Sigma_z = \Sigma_w = \delta^2$

$P(a | x_1 z_1, \cdots x_N z_N) \sim N(\mu_{x|y}, \Sigma_{x|y})$

$\mu_{x|y} = (\Sigma_x^{-1} + A^T \Sigma_z^{-1} A)^{-1} \{A^T \Sigma_z^{-1} y\}$

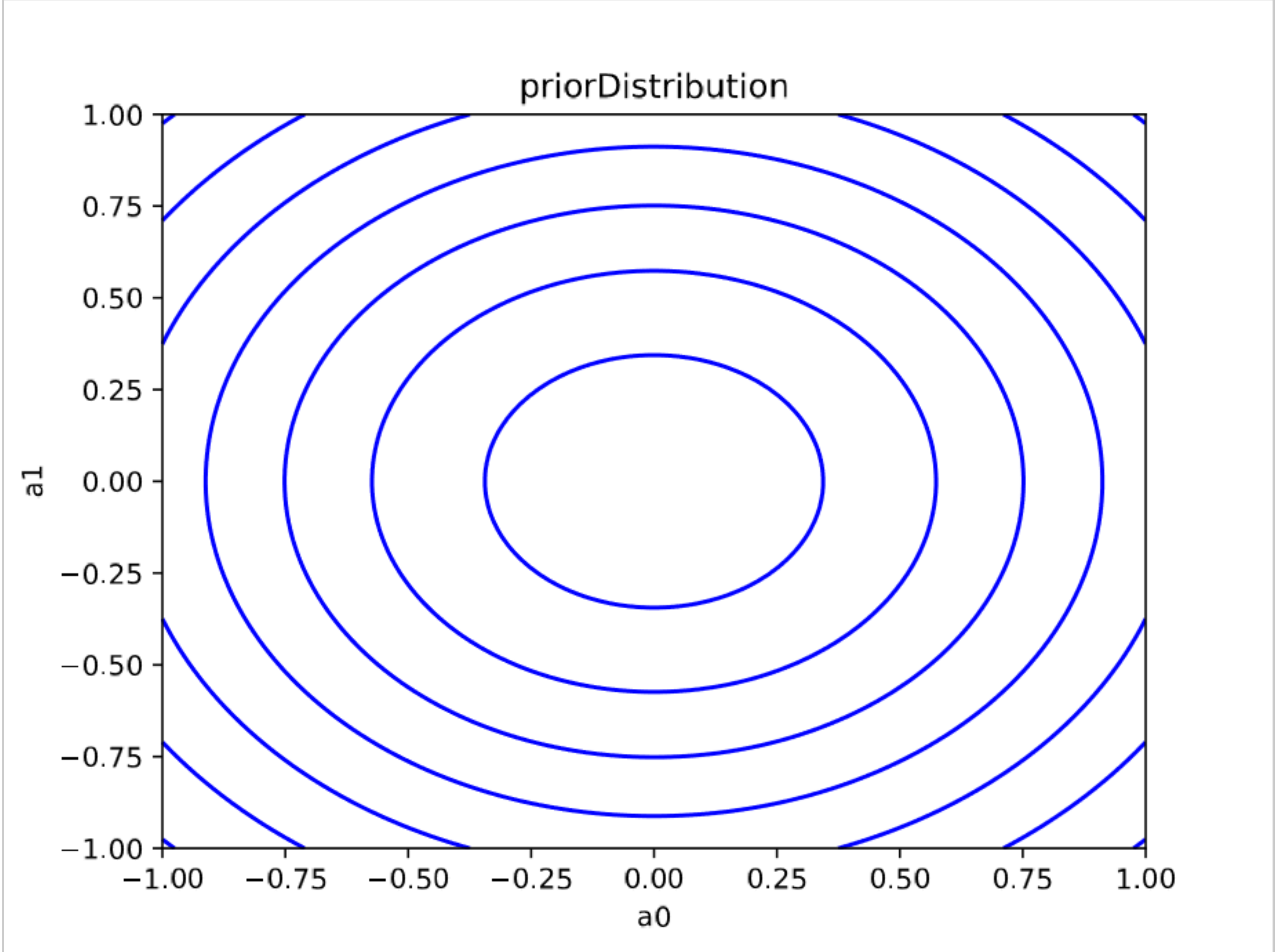
$\Sigma_{x|y} = (\Sigma_x^{-1} + A^T \Sigma_z^{-1} A)^{-1}$

$\Sigma_x^{-1} = \begin{bmatrix} \frac{1}{\beta} & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \quad \Sigma_z^{-1} = \frac{1}{\delta^2}$

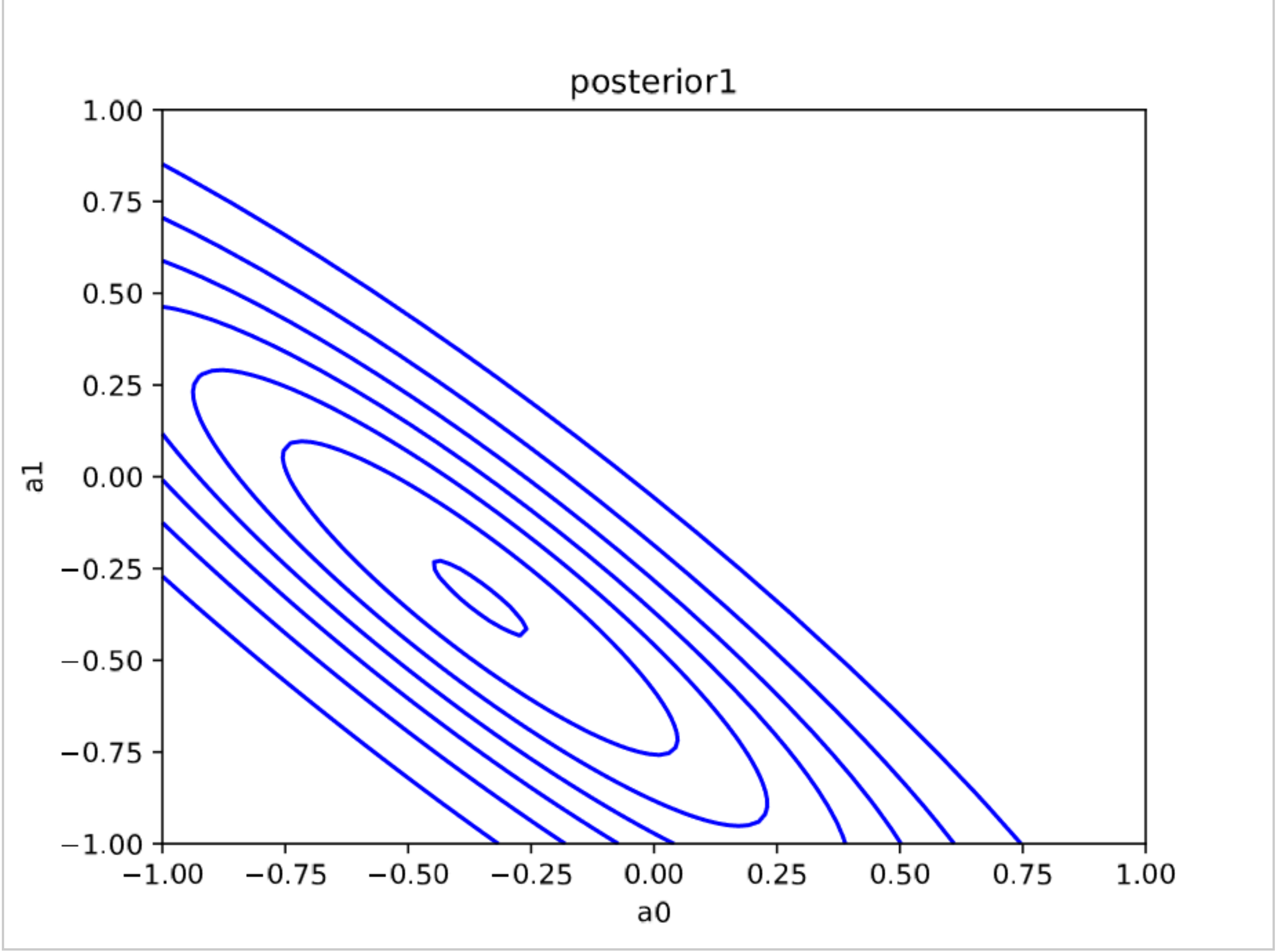
2. Let $\sigma^2 = 0.1$ and $\beta = 1$. Draw four contour plots corresponding to the distributions $p(\mathbf{a})$, $p(\mathbf{a}|x_1, z_1)$, $p(\mathbf{a}|x_1, z_1, \ldots, x_5, z_5)$, and $p(\mathbf{a}|x_1, z_1, \ldots, x_{100}, z_{100})$. In all contour plots, the x-axis represents a_0, and the y-axis represents a_1. Please save the figures with names **prior.pdf, posterior1.pdf, posterior5.pdf, posterior100.pdf**, respectively. (1.5 **pt**)

$p(a) = x \sim N(\vec{0}, \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix})$

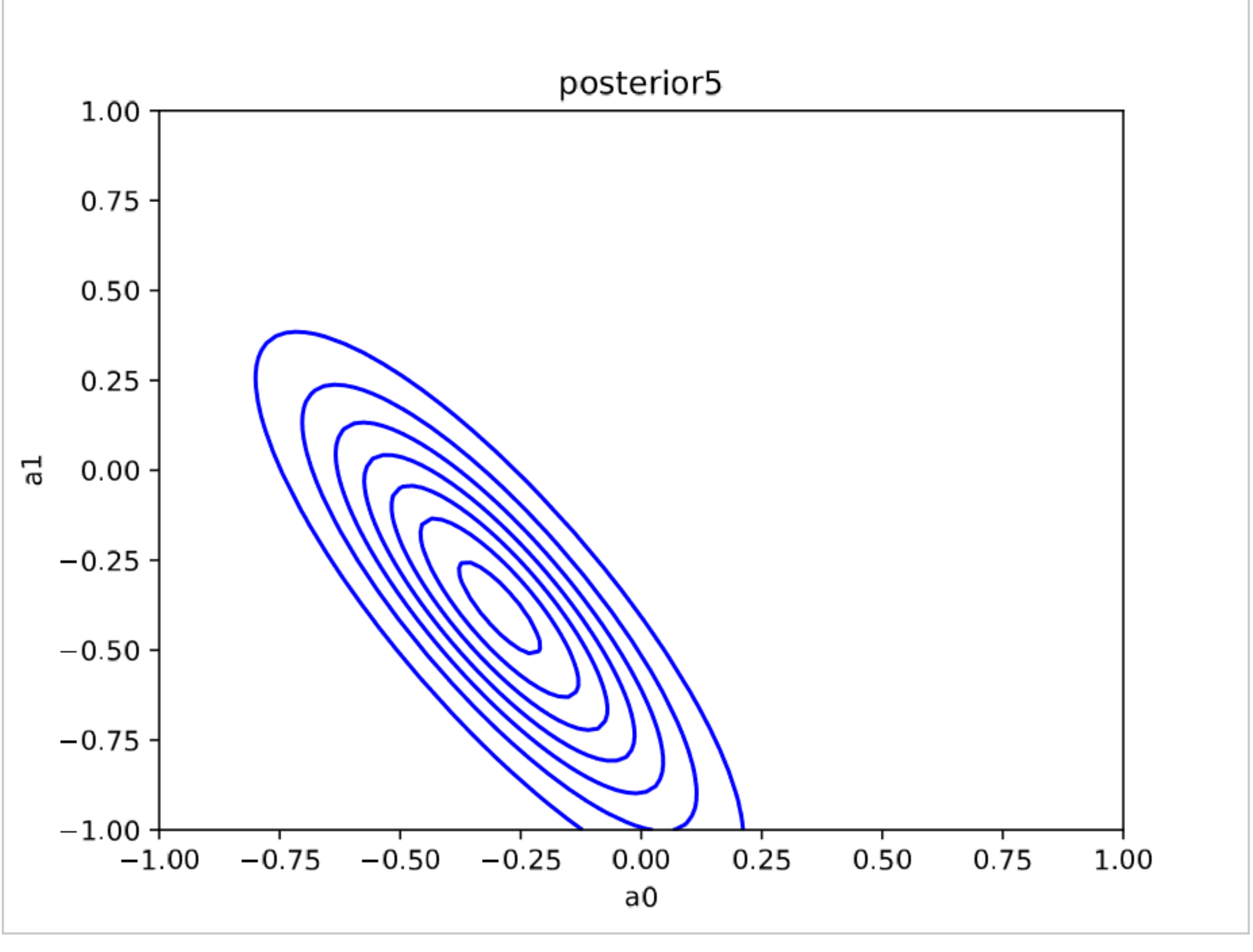
Prior.pdf



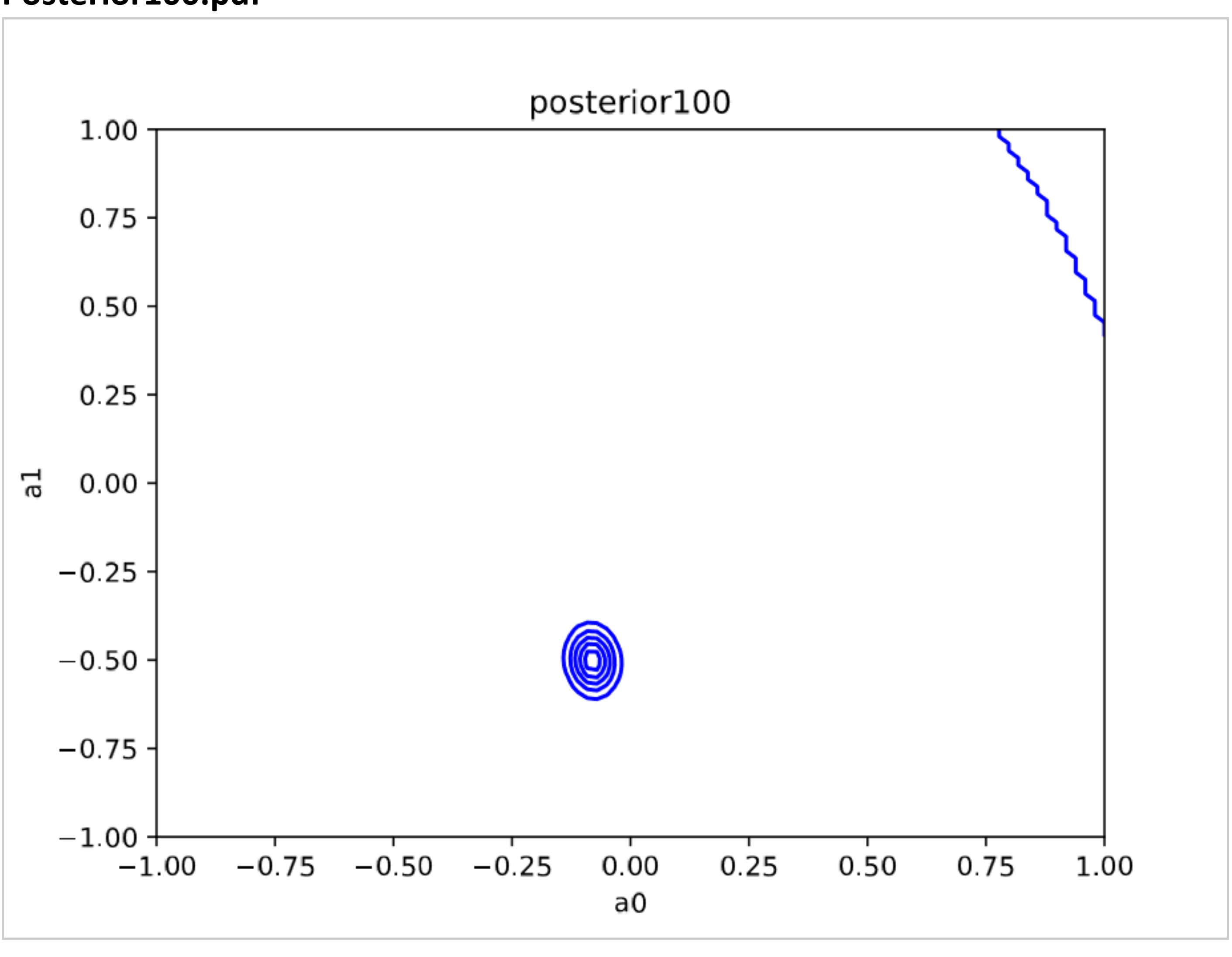
Posterior1.pdf



Posterior5.pdf



Posterior100.pdf



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**)

from above $P(a | x_1 z_1 \cdots x_N z_N) \sim N(\mu_{x|y}, \Sigma_{x|y})$

① $\mu_{x|y} = (\Sigma_x^{-1} + A^T \Sigma_z^{-1} A)^{-1} \{A^T \Sigma_z^{-1} y\}$

② $\Sigma_{x|y} = (\Sigma_x^{-1} + A^T \Sigma_z^{-1} A)^{-1}$

$\Sigma_x^{-1} = \begin{bmatrix} \frac{1}{\beta} & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \quad A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$

$\Sigma_z^{-1} = \frac{1}{\delta^2} \quad y = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}$

New Gaussian System

$a \sim N(\mu_{x|y}, \Sigma_{x|y})$

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

$z = [1 \quad x] \vec{a} + w$

$z \sim N(\mu_z, \Sigma_z)$

$\mu_z = [1 \quad x] \mu_a$

$\mu_a = \mu_{x|y}$ from ①

$\Sigma_z = [1 \quad x] \Sigma_a \begin{bmatrix} 1 \\ x \end{bmatrix} + \Sigma_w$

$\Sigma_a = \mu_{x|y}$ from ②

$\Sigma_w = \delta^2$

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:

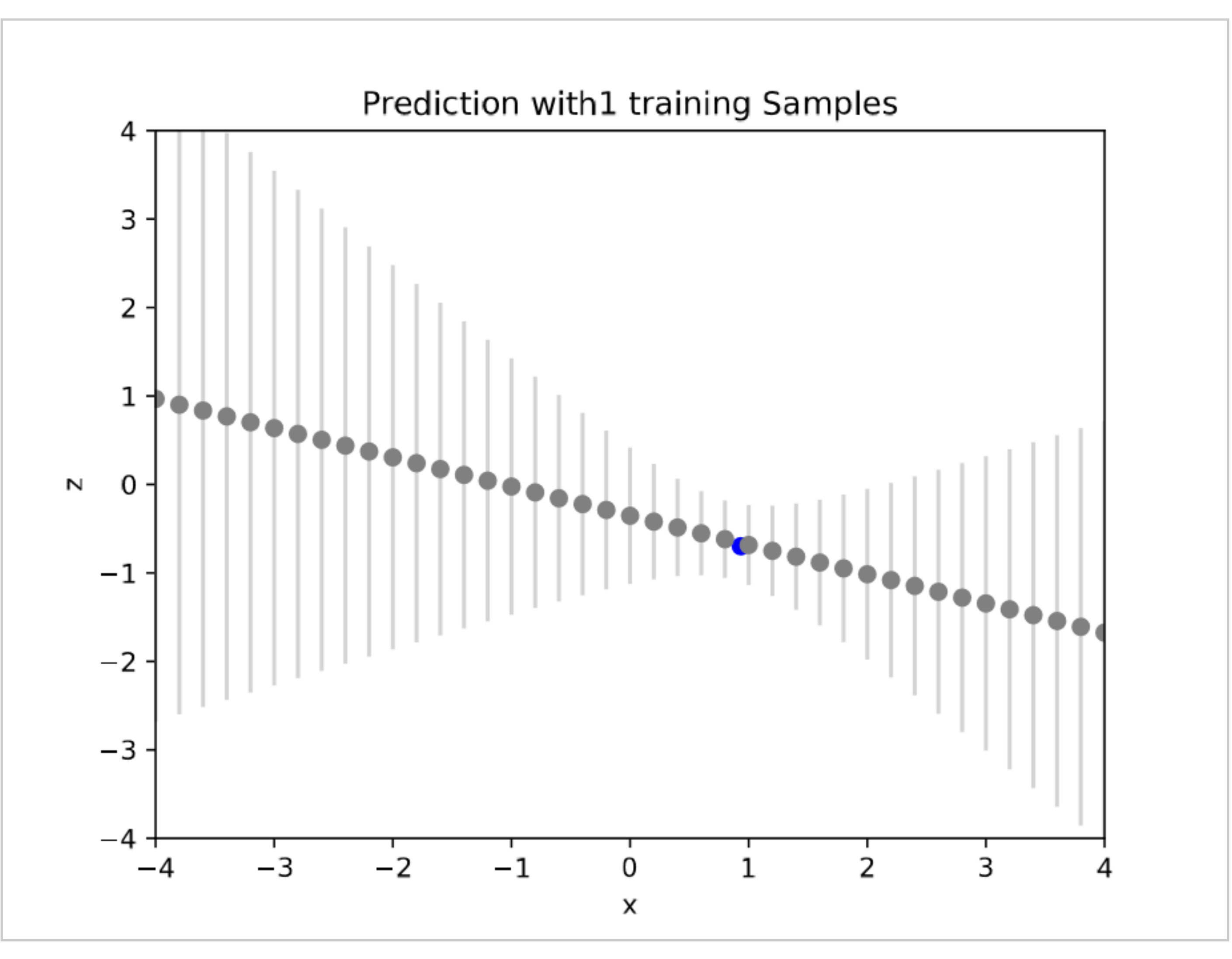
   (a) The predictions are based on one training sample, i.e., based on $p(z|x, x_1, z_1)$.

   (b) The predictions are based on 5 training samples, i.e., based on $p(z|x, x_1, z_1, \ldots, x_5, z_5)$.

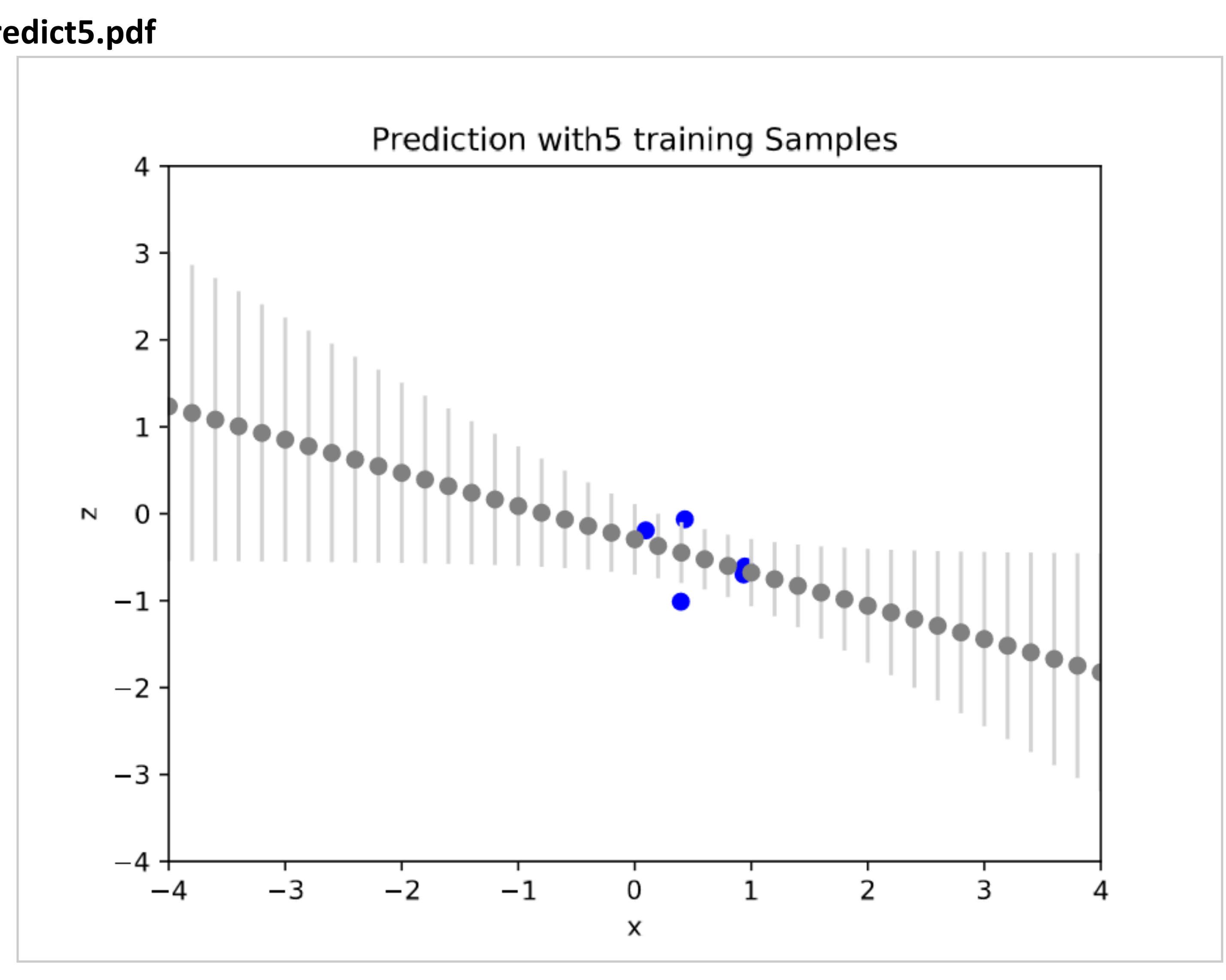
   (c) The predictions are based on 100 training samples, i.e., based on $p(z|x, x_1, z_1, \ldots, x_{100}, z_{100})$.

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**)

Predict1.pdf



Predict5.pdf



Predict100.pdf