Homework 1 of Machine learning

February 12, 2018

Zhen Yang

Department of Civil and Environmental Engineering

University of Michian Ann Arbor, Ann Arobor, MI, 48109

zhenyang@umich.edu

Collaborators: Hoang Duc Bui, Junding Wang, Bochang Wang

1 Problem 1

(a)The predicted labels of all test instances is shown in Figure 1. The blue dots are the predicted labels and the orange dots are the real value. The fluctuation of the real value are noises. The test error is 0.609.

(b) Figure 2 shows the predicted label of test set by locally weighted linear regression. When the

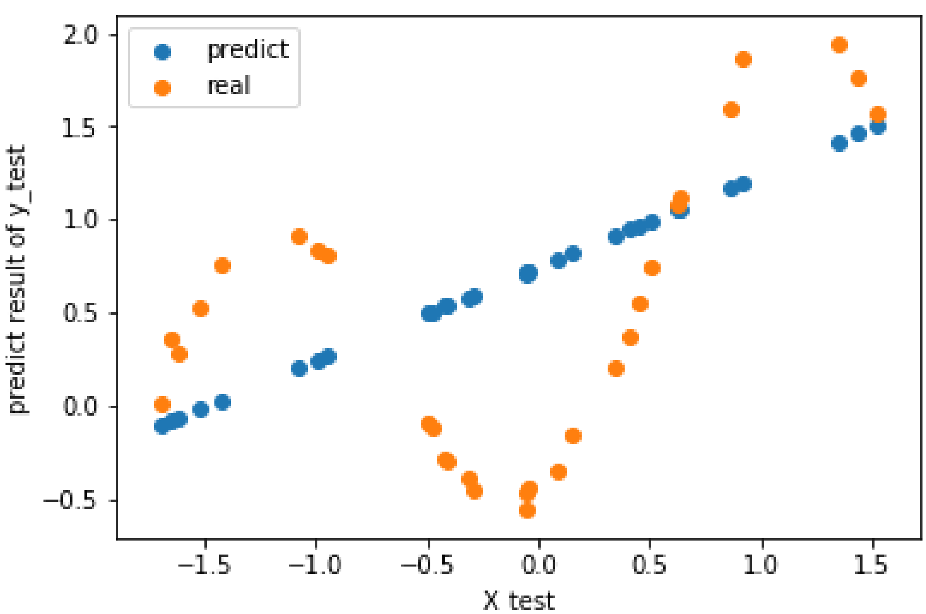
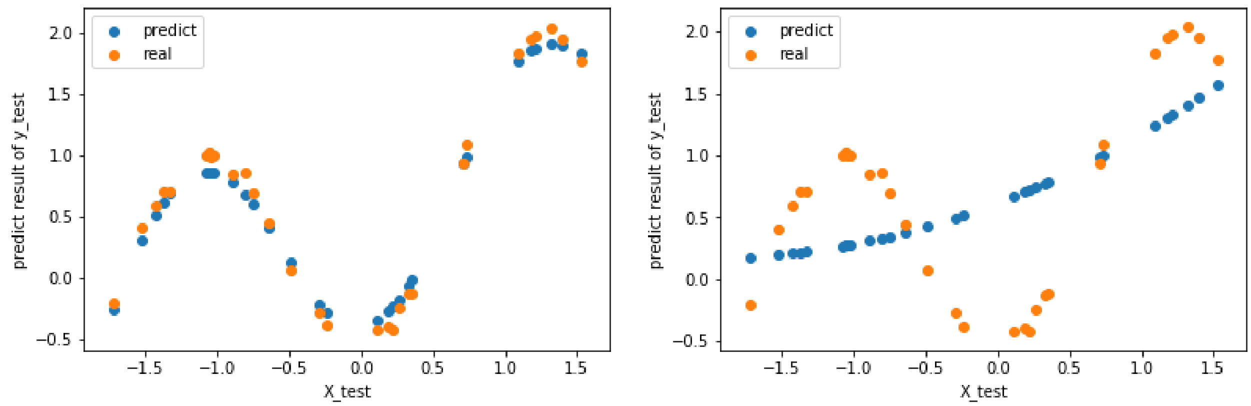


Figure 1: predict label of test set by Linear Regression

kernel width is 0.2, the test error is 0.01, and when the kernel width is 2, the test error is 0.44.



(a) predict label of test set with kernel width 0.2

(b) predict label of test set with kernel width 2

Figure 2: predict label of test set by locally weighted Linear Regression

2 Problem 2

(b)X_1 $N(0, 1.0)$. covariance is 1.0 and mean is 0. Figure 3 shows the distribution of $P(X_1)$.

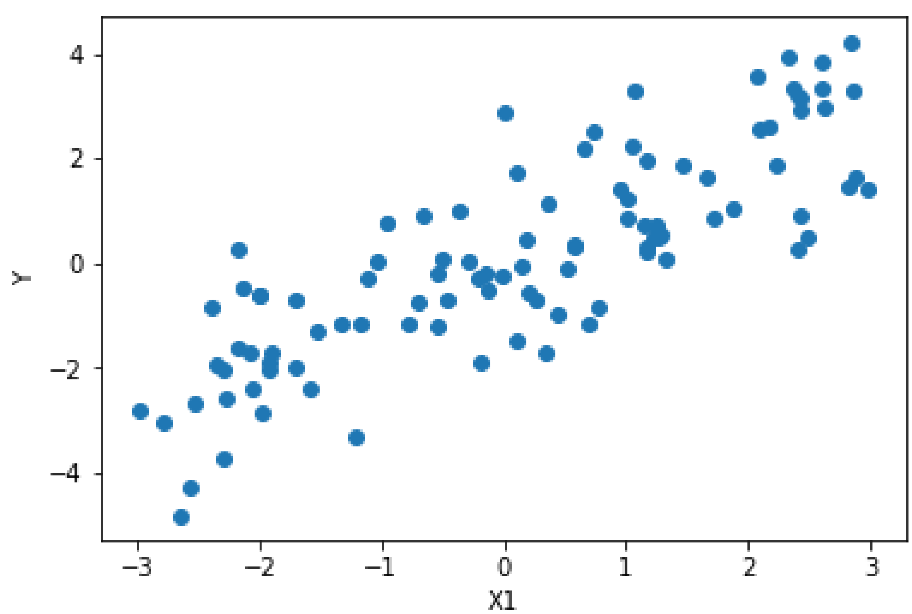


Figure 3: marginal distribution

(d)Mean vector is [0.550.15] and covariance matrix is

$$\begin{bmatrix} 0.75 & -0.75 \\ -0.75 & 1.75 \end{bmatrix}$$

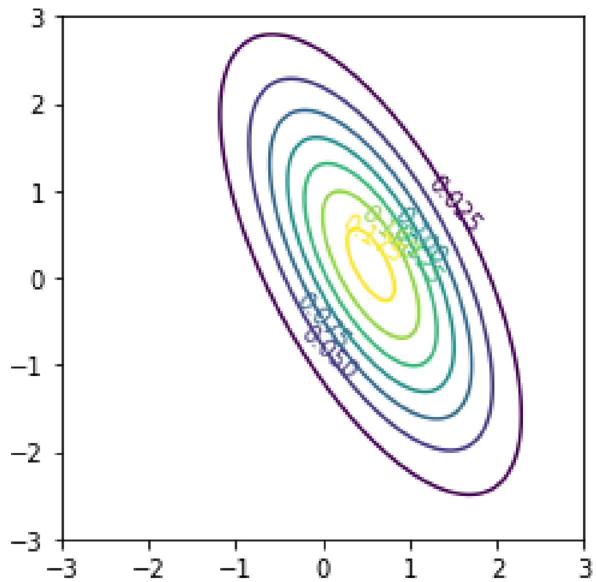


Figure 4: conditional distribution

3 Problem 3

The Figure 6 shows the distribution of w after different times of iterations. After initialization, mean vector is $[0, 0]$, and the covariance matrix:

$$\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

After 1 step of iteration, mean vector is $[0.0266, -0.0288]$, and the covariance matrix:

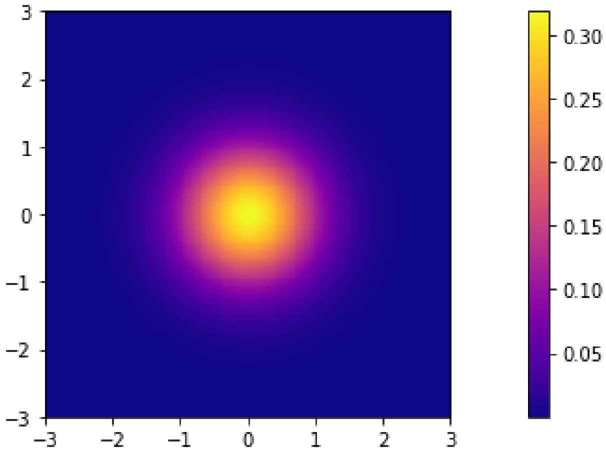
$$\begin{bmatrix} 0.389 & 0.120 \\ 0.120 & 0.370 \end{bmatrix}$$

After 10 steps of iteration, mean vector is $[0.554, -0.158]$, and the covariance matrix:

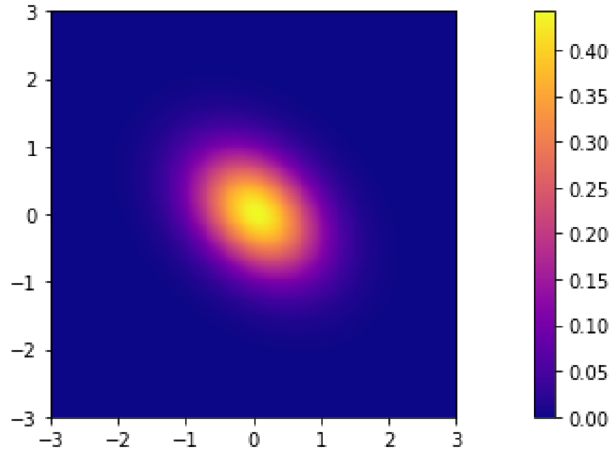
$$\begin{bmatrix} 0.029 & 0.008 \\ 0.008 & 0.085 \end{bmatrix}$$

After 20 steps of iteration, mean vector is $[0.452, -0.207]$, and the covariance matrix:

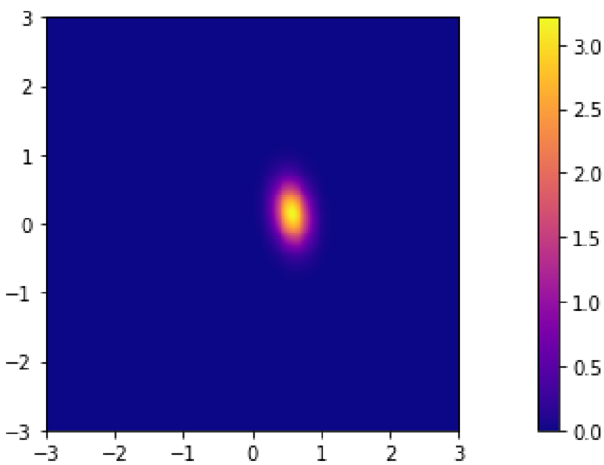
$$\begin{bmatrix} 0.008 & 0.004 \\ 0.004 & 0.021 \end{bmatrix}$$



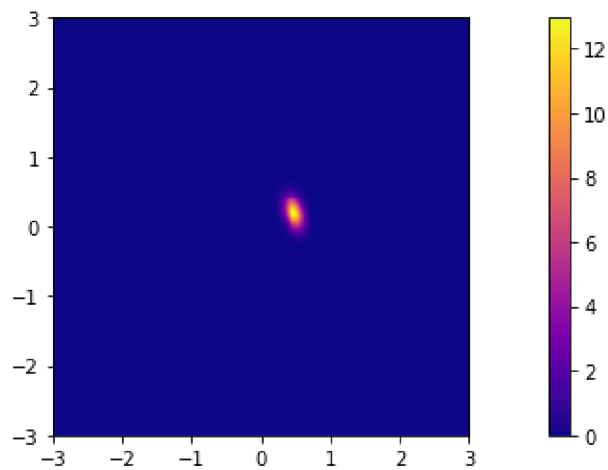
(a) initial distribution



(b) 1-step distribution



(a) 10-step distribution



(b) 20-step distribution

Figure 6: distribution of w after different times of iterations

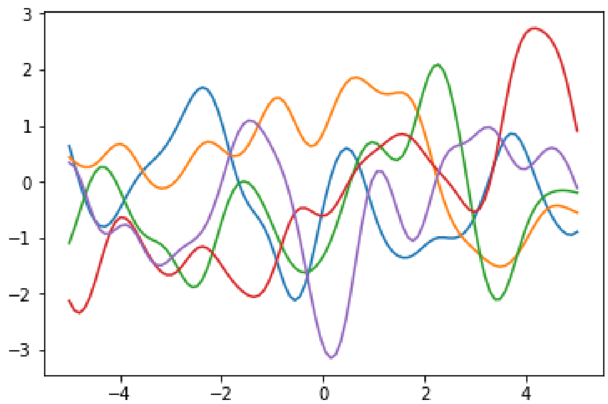
4 Problem 4

(a)i.

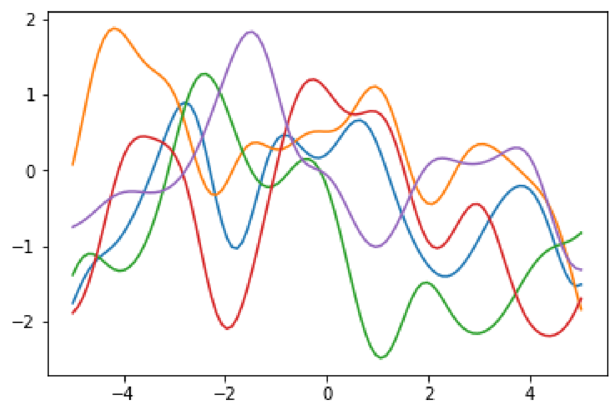
$$\begin{bmatrix} y(x_1) \\ y(x_2) \\ ... \\ y(x_n) \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \\ ... \\ 0 \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{bmatrix}) \tag{1}$$

$$k(x,x') = exp(-\frac{||x-x'||^2}{2\sigma}) \tag{2}$$

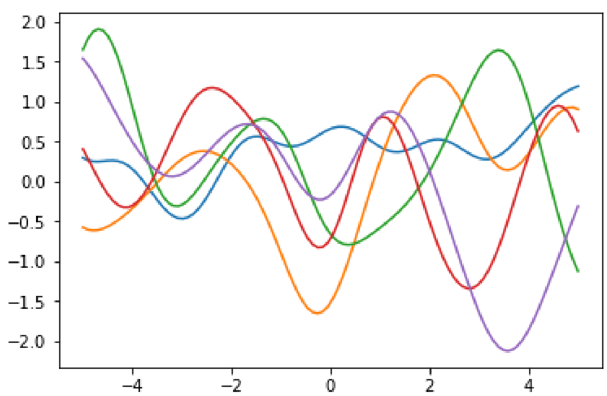
ii Figure 8 shows the sample function with different kernel parameter σ



(a) $\sigma = 0.3$



(b) $\sigma = 0.5$



(c) $\sigma = 1.0$

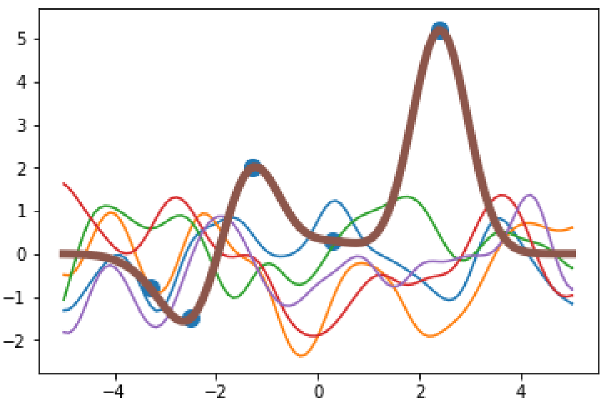
Figure 7: 5 sample functions of Gaussian Process

(b)i

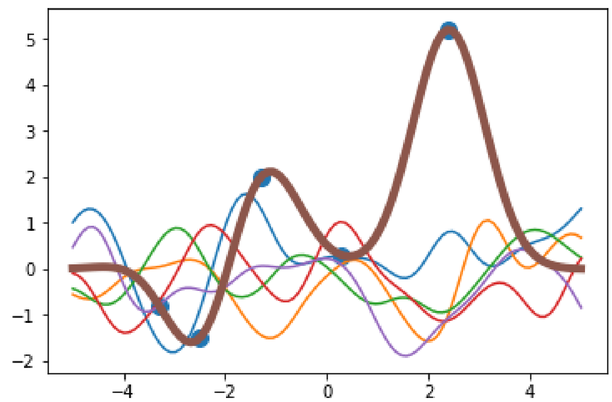
$$\begin{bmatrix} y^D \\ y \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{K}(x^D,x^D) & \mathbf{K}(x^D,x) \\ \mathbf{K}(x,x^D) & \mathbf{K}(x,x) \end{bmatrix}) \tag{3}$$

$$P(y|y^D) \sim N(\mathbf{K}(x,x^D)\mathbf{K}(x^D,x^D)^{-1}y^D, \mathbf{K}(x,x) - \mathbf{K}(x,x^D)\mathbf{K}(x^D,x^D)^{-1}\mathbf{K}(x^D,x)) \tag{4}$$

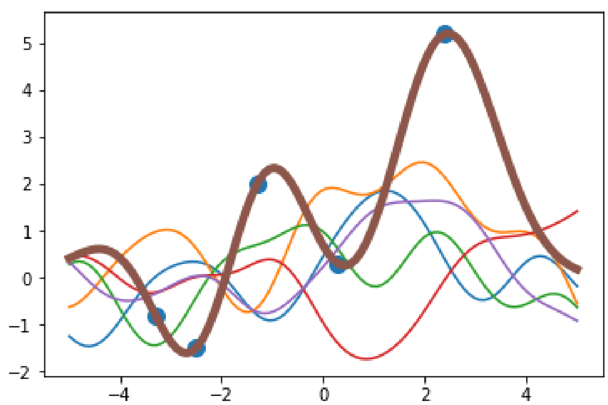
ii



(a) $\sigma = 0.3$



(b) $\sigma = 0.5$



(c) $\sigma = 1.0$

Figure 8: 5 sample functions of Conditional Gaussian Process