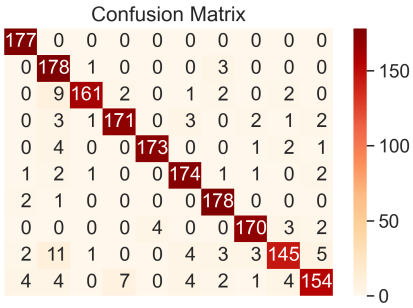
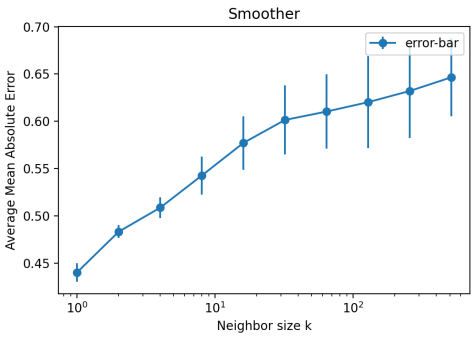
		Q1	Q2	Q3	Q4	Q5	Total
Grade	Max	1	1	1	1	1	5
	Expected						



(a) Non-parametric Classification



(b) Gaussian Smoother

Figure 1. Example figures.

Question 1

Implement the multivariate nonparametric classification algorithm with Gaussian kernel (**8.4**). Use Mahalanobis distance $d(\vec{x}, \vec{y})^2 = (\vec{x} - \vec{y})S^{-1}(\vec{x} - \vec{y})$ to normalize the variances. S denotes the covariance matrix of the training set.

Question 2

Run your classifier on Optdigits data set. Use optdigits.tra as your training set and optdigits.tes as the test set. Plot the confusion matrix similar to the 1.a and report the accuracy of your classifier. You can use external libraries for the confusion matrix.

Question 3

Implement k-NN Smoother (**8.8**) using Gaussian kernel and a fixed number of neighbors (k). Use Mahalanobis distance as you did in the previous questions.

Question 4

You will be applying K-fold cross-validation to tune the closest neighbors parameter k of your regressor. Write a function that randomly splits the given data into N many groups. Download the winequality-red.csv file from Wine Quality Data Set. Run your regressor for K=5 and k=2 on the wine quality data set. Report the mean absolute error for each fold.

Question 5

In order to tune the closest neighbors parameter k, run K-fold cross-validation for K=5 and for each $k \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512\}$. Repeat the experiment with 20 different seeds and plot an error-bar plot similar to the one shown in 1.b and report your average mean absolute errors for each value of k. Note that, in the error-bar figure 1.b, the length of a bar is determined by the standard deviation of the mean absolute errors.