Assignment 05

Machine Learning, Summer term 2018

Norman Hendrich, Marc Bestmann, Philipp Ruppel

May 07, 2018

Solutions due by May 14

Assignment 05.1 (Multi-class classification, 2+2+2 points)

The Iris flower data-set consists of 50 samples from each of three species (Iris setosa, Iris virginica, Iris versicolor). Four features were measured from each sample: the length and the width of the sepals and petals, in centimeters. Your task is to predict the species from these features.

A simple approach to multi-class classification is to divide the problem into binary subproblems. Each subproblem deals with the classification for one of the classes c (class c vs. not c, also called one-vs-the-rest). The corresponding target vector is $y_c \in \{0, 1\}^n$ (for n examples), where $y_{c,i} = 1$ is x_i belongs to c and $y_{c,i} = 0$ otherwise.

In this exercise, we approximate each binary classification problem with regression (in other words, we regress on only two values, namely 0 and 1). For each class c, we calculate the least-squares regression function f_c from X and y_c. Based on all $\{f_c\}$, we predict $y_{pred} = \mathrm{argmax_c f_c(x)}$; this is the class of the regression function with the largest value for the given input x. For example, for $f_1(x) = 0.31$, $f_2(x) = 0.39$ and $f_3(x) = 0.18$, we precit label 2.

a. Load `iris_multiclass.mat`. You have 3 classes in your dataset. Note that you have to create appropriate training and test data yourself (`indices_train` and `indices_test`).

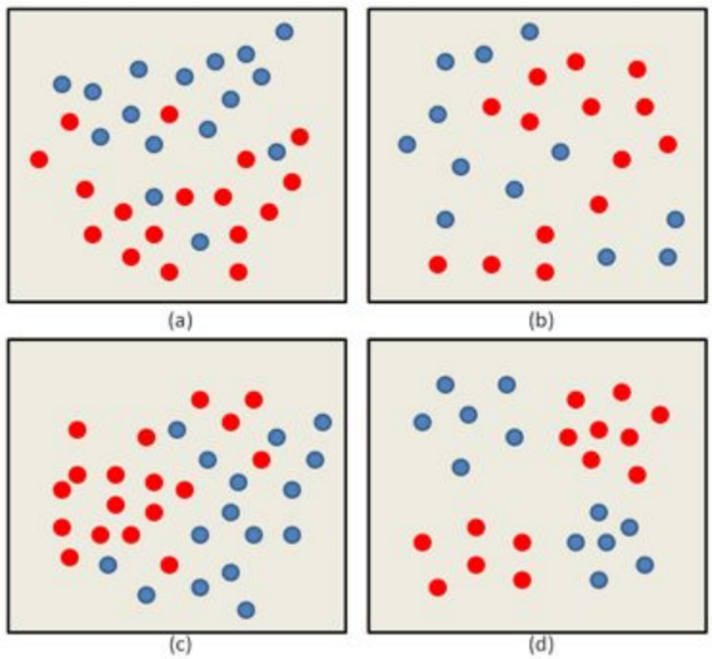
   Estimate w_{setosa}, $w_{versicolor}$ and $w_{virginica}$ based on the training data.

b. Predict the classes for the test data. Report the misclassification error (0/1 loss).

c. What are potential problems of using least-squares regressions for multi-class classification?

Assignment 05.2 (Decision boundaries, 1+1+1 points)

Consider the following four samples, where colors indiciate class labels. Each sample is typical of an underlying distribution.



Plot the decision boundaries for the four samples when solving this classification problem with:

a. a 1NN classifier;

b. linear least-squares regression (linear basis functions);

c. linear least-squares regression with quadratic basis functions.

Assignment 05.3 (Play with SVM, 1+1+1 points)

In this exercise, you play with applets to get a better understanding of SVMs. For example use the libsvm applet at `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`.

Try to set training points as depicted in Figure 2-a.

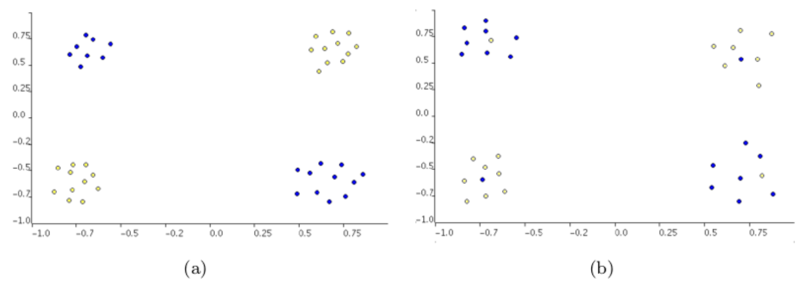


Figure 2: (a) Training data. (b) Training data with outliers.

Now train the SVM with the following settings. Capture the output screen for your report.

a. Linear kernel (simple dot product) with $C = 100$.

b. Polynomial kernel of degree 2, and degree 8. Choose a proper C.

c. Gaussian kernel (radial basis fuction): Try different values for σ. Note that sometimes a different parametrization is used (e.g. $\beta = 1/(2\sigma^2)$. Try different values for β in this case. choose a proper C.

d. Now add noise to your training data as depicted in Figure 2 b. Try a Gaussian kernel with $\beta = 10$ and $C = 0, 10, 1000$. Based on your observation, describe the effect of the parameter C.

Assignment 05.4 (Play with SVM II, 1+1+1 points)

Use again the applets for SVMs as in the previous exercise.

a. Create a data set which for fixed $C = 0.0001$ is perfectly separable with a polynomial kernel of degree 3, but not with a polynomial kernel of degree 2.

b. Create a data set which for fixed $C = 0.0001$ is perfeclty separable with a radial basis function kernel with small σ (e.g., $\sigma = 0.1/\beta = 50$), but not with large σ (e.g. $\sigma = 10$, / $\beta = 1/200$).

c. Create a data set which for fixed $C = 0.0001$ is perfeclty separable with a radial basis function kernel, but not with a polynomial kernel of degree 3.