Optmization Methods for Machine Learning - Fall 2015

Assignment # 3
Support Vector Machines
Clustering

Laura Palagi

Dipartimento di Ingegneria informatica automatica e gestionale A. Ruberti

Sapienza Università di Roma

Posted on December 4, 2015- due date December 28, 2016

In this assignment you will implement decomposition methods for Support Vector Machines (supervised learning) applied to classification problems. You can use the matlab optimization toolbox for the solution of the subproblems.

As for the **evaluation**, you may reach the max score (between 28 to 30 cum laude) by correctly answering to both Question 1 and Question 2 using the second test set (dataset from LIBSVM). Correct answers to both Question 1 and Question 2 using the two dimensional set allow to obtain only up to B+ (approximatively 27).

Question 1.

Determine the decision function

$$y(x) = sign(w^T x + b)$$

where w, b are obtained by means of nonlinear SVMs.

Answer to the following question (increasing mark).

1. (max score up to "C") Write the primal forumlation of SVM training problem and solve it by a standard QP algorithm in the optimization matlab toolbox. Check the results in terms of test error. Analyse the role of the parameter C.

2. (max score up to "B") Write a matlab code which implements an algorithm for the optimal solution of the dual quadratic problem usimg

   - a linear kernel
   $$K(x,y) = x^T y$$

- a Gaussian Kernel

$$K(x,y) = e^{-\gamma\|x-y\|^2} \qquad \gamma > 0$$

and a standard QP algorithm for its solution. Analyse the performance both from machine learning perspective (test error) and optimization perspective (e.g. computational time) Analyse the occurrence of overfitting by varying the value of the parameters (either C and the kernel one, if any)

If the problem is a two dimensional two dimensional one, produce a plot of a non-linear decision boundary.

3. (max score up to "A") Write a matlab code which implements a decomposition method with any $q \geq 2$ for the dual quadratic problem. You must define the selection rule of the working set, construct the subproblem at each iteration and uses a standard QP algorithm for its solutions. A stopping criterion on optimality is needed.

   If the problem is a two dimensional two dimensional one, produce a plot of a non-linear decision boundary.

4. (max score up to "A+") Fix $q = 2$ and implement the analytic solution of the subproblem

   As training data set you may use (attention to the maximum core obtainable)

   - (max score achievable "B" ≈ 27) The two-dimensional input samples reported in the picture 4. For the supervised learning problem you use as training set the pairs (x^i, y^i) with $x^i \in \mathbb{R}^2$ and $y^i \in \{-1, 1\}$. The coordinates x^i can be estimated from the chart.
   - (max score achievable "30 cum laude") Any of the two class problem reported on the web page "Data set" of LIBSVM (https://www.csie.ntu.edu.tw/∼cjlin/libsvmtools/datasets/) or any data set in Machine learning libraries. Choose examples with moderate number of features and size of the training set.

Question 2. Choose a multi-class problem on the web page "Data set" of LIBSVM (https://www.csie.ntu.edu.tw/∼cjlin/libsvmtools/datasets/) and implement a multiclass classificator using either one-against-all strategy or one-against-one. Use the code developed in Question 1 as a tool for the solution of the two class subproblems.

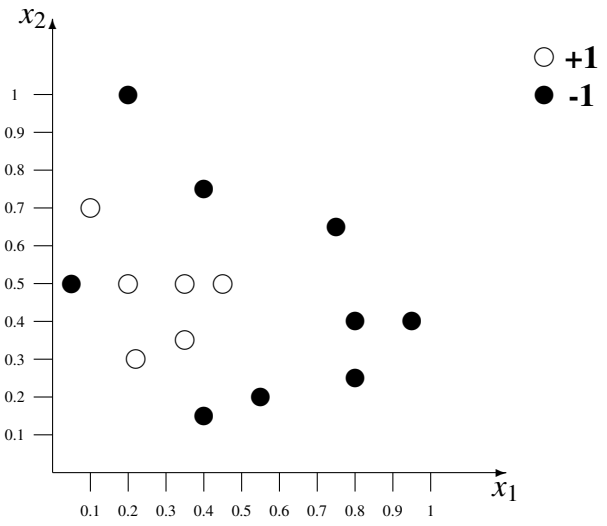

Figure 1: Sample in the two class