Homework 3

Due Date: April 20, 2016

April 6, 2016

1 Support Vector Machine

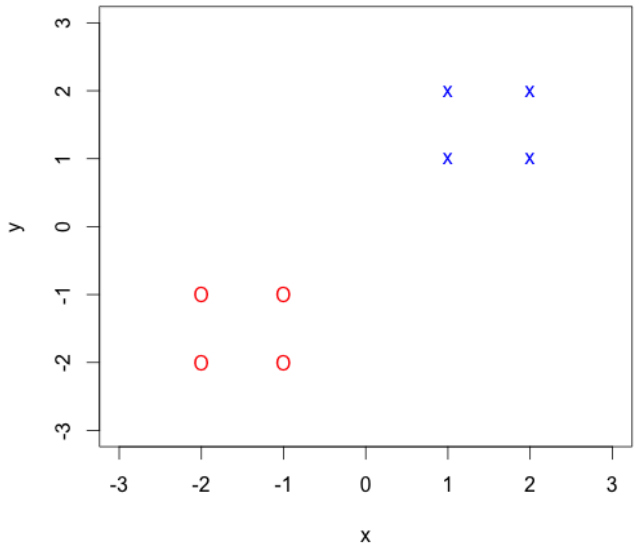
1.1 Required

1. Suppose the separating boundary of a hard-margin SVM is $w^T x + b = 0$.

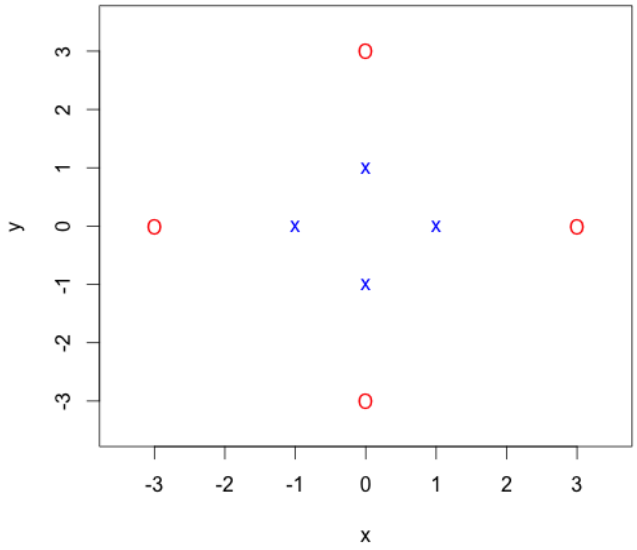
   (a) Show that w is perpendicular to the boundary.

   (b) Suppose there are two support vectors x and y, where x belongs to the positive class and y belongs to the negative class. If you know the value of w, show me how you can find the value of b.

   (c) Write down the separating boundary for the following graph. And show me the coordinates of the support vectors.



(d) Write down the separating boundary for the following graph. Show me in details how you find it. (You may have to use kernel)



2. l_2 norm soft margin SVMs
In class, we saw that if our data is not linearly separable, then we need to modify our support vector machine algorithm by introducing an error margin that must be minimized. Specifically, the formulation we have looked at is known as the l_1 norm soft margin SVM. In this problem we will consider an alternative method, known as the l_2 norm soft margin SVM. This new algorithm is given by the following optimization problem (notice that the slack penalties are now squared):

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m}\xi_i^2$$
$$s.t.\ y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \xi_i, i = 1, \ldots, m$$

(a) Notice that we have dropped the $\xi_i \geq 0$ constraint in the l_2 problem. Show that these non-negativity constraints can be removed. That is, show that the optimal value of the objective will be the same whether or not these constraints are present.

(b) What is the Lagrangian of the l_2 soft margin SVM optimization problem?

(c) Minimize the Lagrangian with respect to w, b, and ξ by taking the following gradients: $\nabla_w L$, $\frac{\partial L}{\partial b}$, and $\nabla_\xi L$,and then setting them equal to 0. Here, $\xi = [\xi_1, \xi_2, \ldots, \xi_m]^T$.

(d) What is the dual of the l_2 soft margin SVM optimization problem?

1.2 Optional

Rebuild the email classifier in hw2 with SVM using only *word* features (no feature selection required). You are encouraged to use existing libraries. (LibSVM is a good choice!)

Things to be included in your report:

1. If you have preprocessed your data, please show me in details about your procedures.
2. Use different kernels to train your model, show me which one is the best one for this problem.
3. list values of all parameters in your model
4. list test accuracy, precision, recall

Note: it may frustrate you a lot to twist the parameters. Please be patient. And if you find some good strategies, please write them down in your report. I will share your experience with your peers.