Stat 627/427 (Statistical Machine Learning)

Homework # 11: Support Vector Machines (Chap. 9)

Due April 24 by noon on Blackboard. Quiz #9 is on April 25.

1. (**Chap. 9, # 1a, p. 368**) This problem involves a hyperplane in two dimensions. Sketch the hyperplane $1 + 3X_1 - X_2 = 0$. Indicate the set of points for which $1 + 3X_1 - X_2 > 0$, as well as the set of points for which $1 + 3X_1 - X_2 < 0$.

2. (**Chap. 9, # 2, p. 368**) We have seen that in p = 2 dimensions, a linear decision boundary takes the form $\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$. We now investigate a **non-linear** decision boundary.

   (a) Sketch the curve $(1 + X_1)^2 + (2 - X_2)^2 = 4$.

   (b) On your sketch, indicate the set of points for which $(1 + X_1)^2 + (2 - X_2)^2 > 4$, as well as the set of points for which $(1 + X_1)^2 + (2 - X_2)^2 \leq 4$.

   (c) Suppose that a classifier assigns an observation to the blue class if $(1+X_1)^2+(2-X_2)^2 > 4$, and to the red class otherwise. To what class is the observation $(0, 0)$ classified? $(-1, 1)$? $(2, 2)$? $(3, 8)$?

   (d) Argue that while the decision boundary in (c) is not linear in terms of X_1 and X_2, it is linear in terms of X_1, X_1^2, X_2, and X_2^2.

3. (**Chap. 9, # 3, p. 368**) Here we explore the maximal margin classifier on a toy data set.

   (a) We are given $n = 7$ observations in $p = 2$ dimensions. For each observation, there is an associated class label.

Obs.	X_1	X_2	Y
1	3	4	Red
2	2	2	Red
3	4	4	Red
4	1	4	Red
5	2	1	Blue
6	4	3	Blue
7	4	1	Blue

   Sketch the observations.

   (b) Sketch the optimal separating hyperplane, and provide the equation for this hyperplane such as in exercise #1.

   (c) Describe the classification rule for the maximal margin classifier. It should be something along the lines of "Classify to Red if $\beta_0 + \beta_1 X_1 + \beta_2 X_2 > 0$, and classify to Blue otherwise". Provide the values for β_0, β_1, and β_2.

   (d) On your sketch, indicate the margin for the maximal margin hyperplane.

   (e) Indicate the support vectors for the maximal margin classifier.

   (f) Argue that a slight movement of the seventh observation would not affect the maximal margin hyperplane.

   (g) Sketch a hyperplane that is not the optimal separating hyper- plane, and provide the equation for this hyperplane.

   (h) Draw an additional observation on the plot so that the two classes are no longer separable by a hyperplane.

4. (**Chap. 9, # 7, p. 371)** In this problem, you will use support vector approaches in order to predict whether a given car gets high or low gas mileage based on the Auto data set (from `ISLR`).

   (a) Create a binary variable that takes on a 1 for cars with gas mileage above the median, and a 0 for cars with gas mileage below the median.

   (b) Fit a support vector classifier to the data with various values of cost, in order to predict whether a car gets high or low gas mileage. Report the cross-validation errors associated with different values of this parameter. Comment on your results.

   (c) Now repeat (b), this time using SVMs with radial and polynomial basis kernels. Comment on your results.

   (d) Make some plots to back up your assertions in (b) and (c). When $p > 2$, you can use the `plot()` function to create plots displaying pairs of variables at a time. For example, to plot horsepower and year, the syntax is `plot(svm.result, data=Auto, horsepower~year)`.