Machine Learning, Spring 2021
Homework 8

Daniil Bulat, Jonas Josef Huwyler, Haochen Li,Giovanni Magagnin

5/6/2021

Exercise 1

Consider the Iris data set, with only the first 100 data samples (setosa and versicolor), and only the features *sepal_length* and *petal_length*.

```
## Warning: package 'e1071' was built under R version 4.0.5
```

Question 1

Run *svm* on the data (without scaling) with cost equal to 10,1,0.1, respectively, and plot the corresponding outcome. What difference do you see, and how are they related to the cost values.

```r
# set random seed
set.seed(123)
# Data Preparation
Data = iris[sample(1:100),c(1,3,5)]
# svm with cost = 0.1
fit1 = svm(Species ~., data= Data, kernel = "linear", cost = "0.1",scale = FALSE)
# svm with cost = 1
fit2 = svm(Species ~., data= Data, kernel = "linear", cost = "1",scale = FALSE)
# svm with cost = 10
fit3 = svm(Species ~., data= Data, kernel = "linear", cost = "10",scale = FALSE)
```
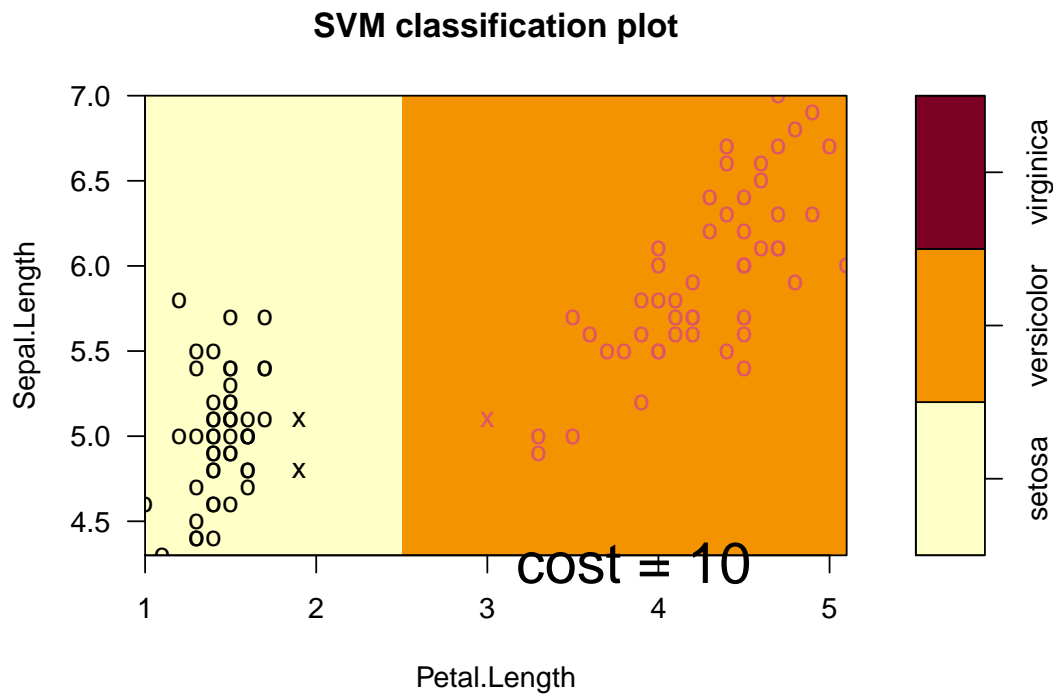
```r
# plot svm model 1
plot(fit1,data = Data)
mtext(side = 1, text = "cost = 0.1",cex = 2, outer = F)
# plot svm model 2
plot(fit2,data = Data)
mtext(side = 1, text = "cost = 1",cex = 2, outer = F)
# plot svm model 3
plot(fit3,data = Data)
mtext(side = 1, text = "cost = 10",cex = 2, outer = F)
```

SVM classification plot

SVM classification plot
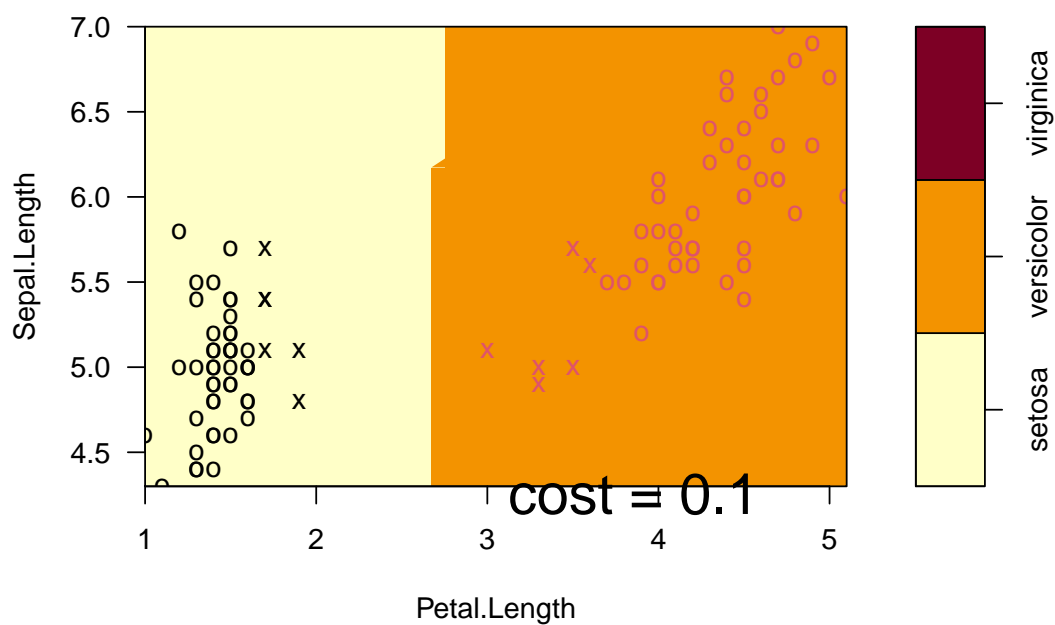
SVM classification plot
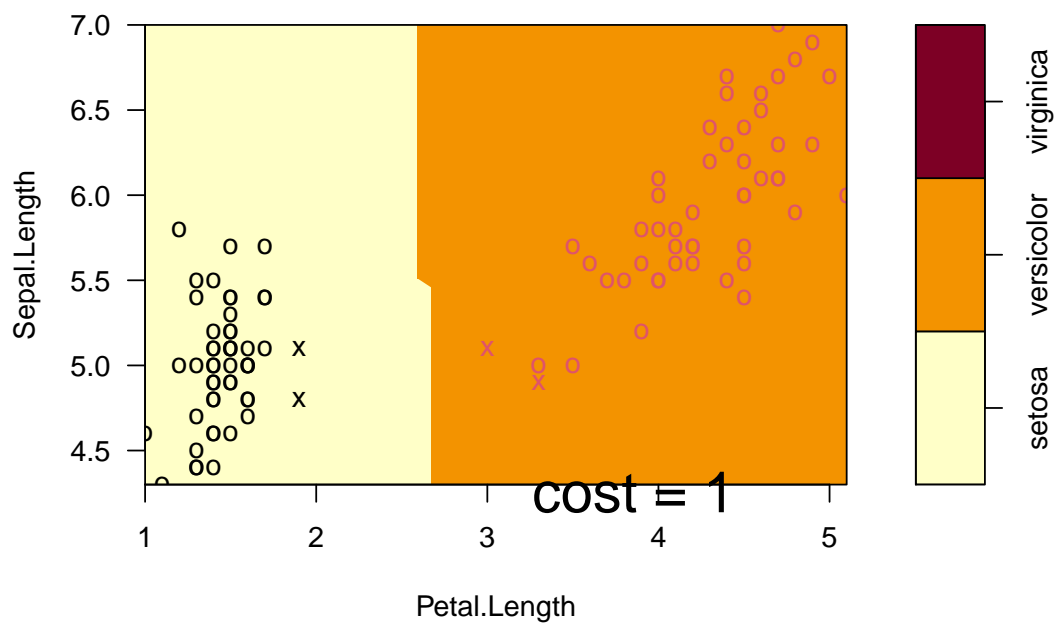
The plots of SVM results are shown above. Here we keep the legends of three classes to compare results of part 4.

With different costs, the shape of separating line and the corresponding margins (can be observed from types of points) change.As the cost increase, the points on the margin or violating the margin (denoted by "x") decrease. In other words, as the cost increase, the number of support vectors becomes smaller, so the separating line defined by the support vectors changes and the margin decreases.

Question 2

For *cost = 0.1* list the support vectors.

```
# For cost = 0.1 list the support vectors.
print(fit1$SV)
```

```
##    Sepal.Length Petal.Length
## 25          4.8          1.9
## 21          5.4          1.7
## 6           5.4          1.7
## 19          5.7          1.7
## 24          5.1          1.7
## 45          5.1          1.9
## 65          5.6          3.6
## 99          5.1          3.0
## 94          5.0          3.3
## 58          4.9          3.3
## 61          5.0          3.5
## 80          5.7          3.5
```

Question 3

For *cost = 0.1* find the equation of the separating line, as explained in Chapter 10 slide 17.

```r
# For cost = 0.1 find the equation of separating line.
cat("The beta_0 is", - fit1$rho, '\n')
cat ("The coefficients are ",t(fit1$coefs)%*%fit1$SV)
```

```
## The beta_0 is 2.442
## The coefficients are  0.02 -0.96
```

Here we compute the corresponding coefficients as above. Thus, the equation of the separating line is

$$2.442 + 0.02x_1 - 0.96x_2 = 0$$

Question 4

Now consider the data samples for versicolor and virginica, and run svm on the data (without scaling) with cost equal to 1. How many support vectors do you get and how does this compare to the result of part 1. (for the same cost)? Can you explain this behavior with respect to the different data sets (and the position of their points)?

```r
# Data Preparationg
set.seed(123)
Data2 = iris[sample(51:150),c(1,3,5)]
# SVM with cost = 1
fit4 = svm(Species ~., data= Data2, kernel = "linear", cost = "1",scale = FALSE)
# list the support vectors.
print(fit4$SV)
```

```
##      Sepal.Length Petal.Length
## 64            6.1          4.7
## 92            6.1          4.6
## 57            6.3          4.7
## 91            5.5          4.4
## 73            6.3          4.9
## 77            6.8          4.8
## 84            6.0          5.1
## 71            5.9          4.8
## 56            5.7          4.5
## 67            5.6          4.5
## 53            6.9          4.9
## 74            6.1          4.7
## 87            6.7          4.7
## 78            6.7          5.0
## 85            5.4          4.5
## 51            7.0          4.7
## 147           6.3          5.0
## 107           4.9          4.5
## 122           5.6          4.9
## 128           6.1          4.9
## 134           6.3          5.1
## 139           6.0          4.8
## 114           5.7          5.0
## 127           6.2          4.8
## 150           5.9          5.1
## 120           6.0          5.0
```
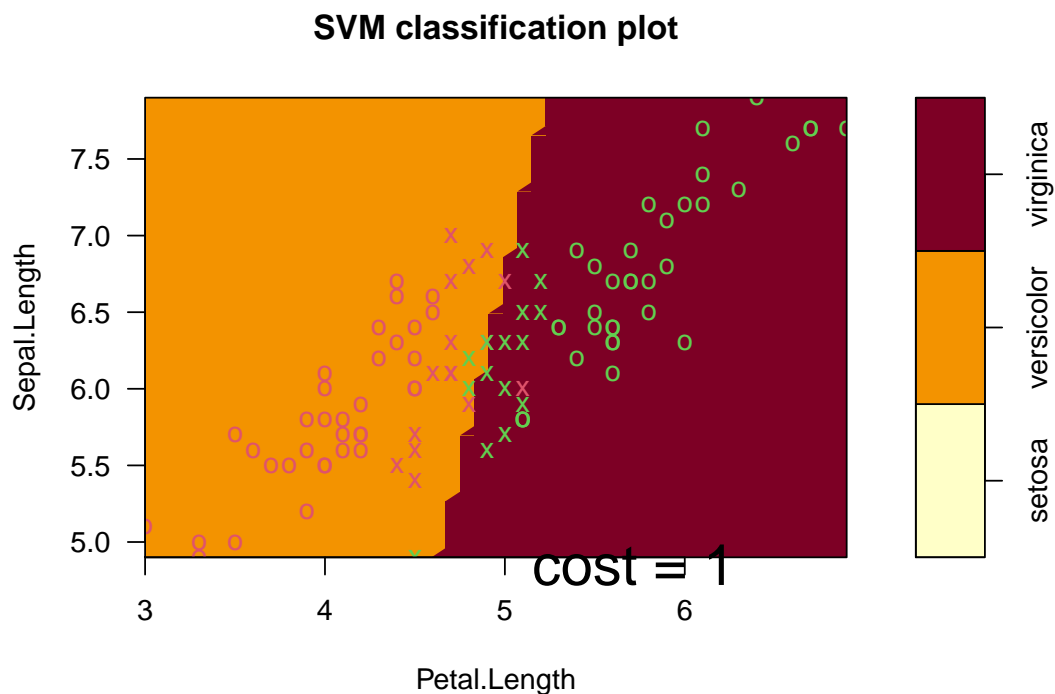
```
## 111          6.5          5.1
## 146          6.7          5.2
## 142          6.9          5.1
## 148          6.5          5.2
## 124          6.3          4.9
```

```r
cat("The number of support vectors is", nrow(fit4$SV),"\n")
```

```
## The number of support vectors is 31
```

```r
# plot svm model 4
plot(fit4,data = Data2)
mtext(side = 1, text = "cost = 1",cex = 2, outer = F)
```


For the svm on the new data set, we have 31 support vectors, which is more than the case of svm result on the old data set with same cost, 4 support vectors.

A possible explanation could be that the new data set is more concentrated distribution. That is to say, under the condition of same cost, there are more points located in the wrong side of the margin.

Exercise 2

Consider the following trainign data, where the first half is in class **-1** and the second half is in class **1**:

$$\{(-1, 2), (1, 1), (0.5, 3), (1, 4), (3, 2), (3.5, 1), (3.5, 4), (4, 2)\}$$

The maximal margin hyperplane is given by the equation $x_1 = 2$ (i.e, a vertical line intersecting the x_1-axis at the point **2**).
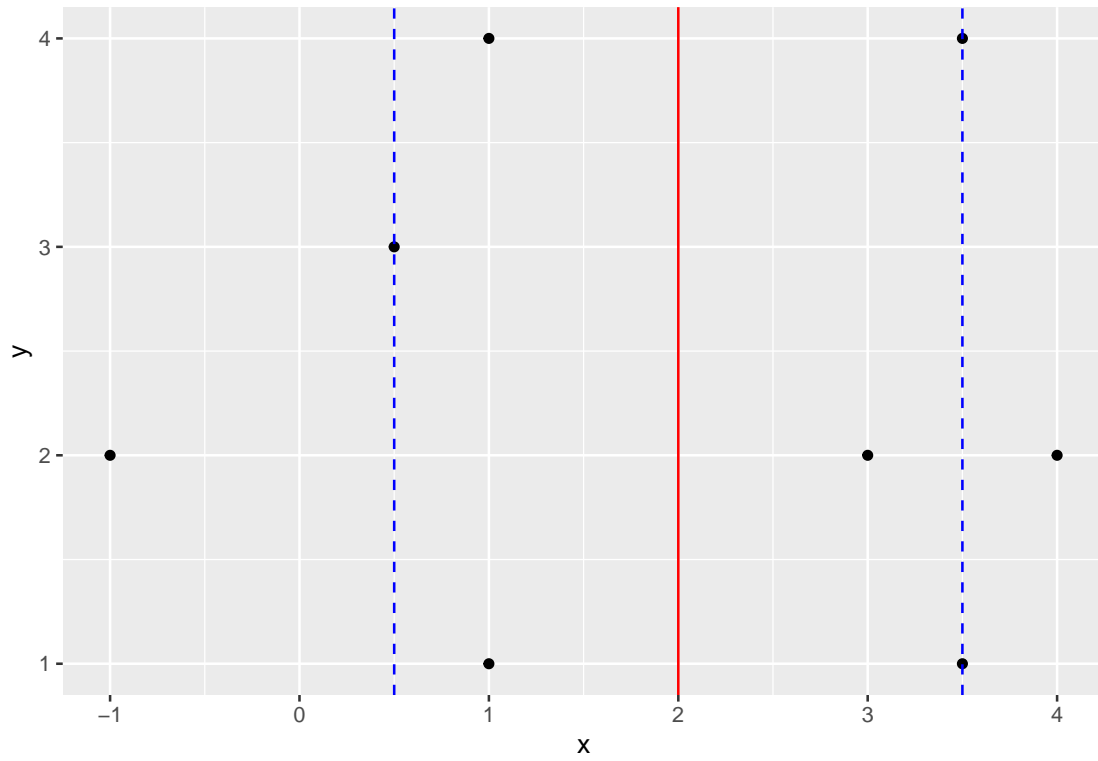
Question 1

Plot the hyperplane and the margin (on both sides). What are the support vectors and how large is the margin?
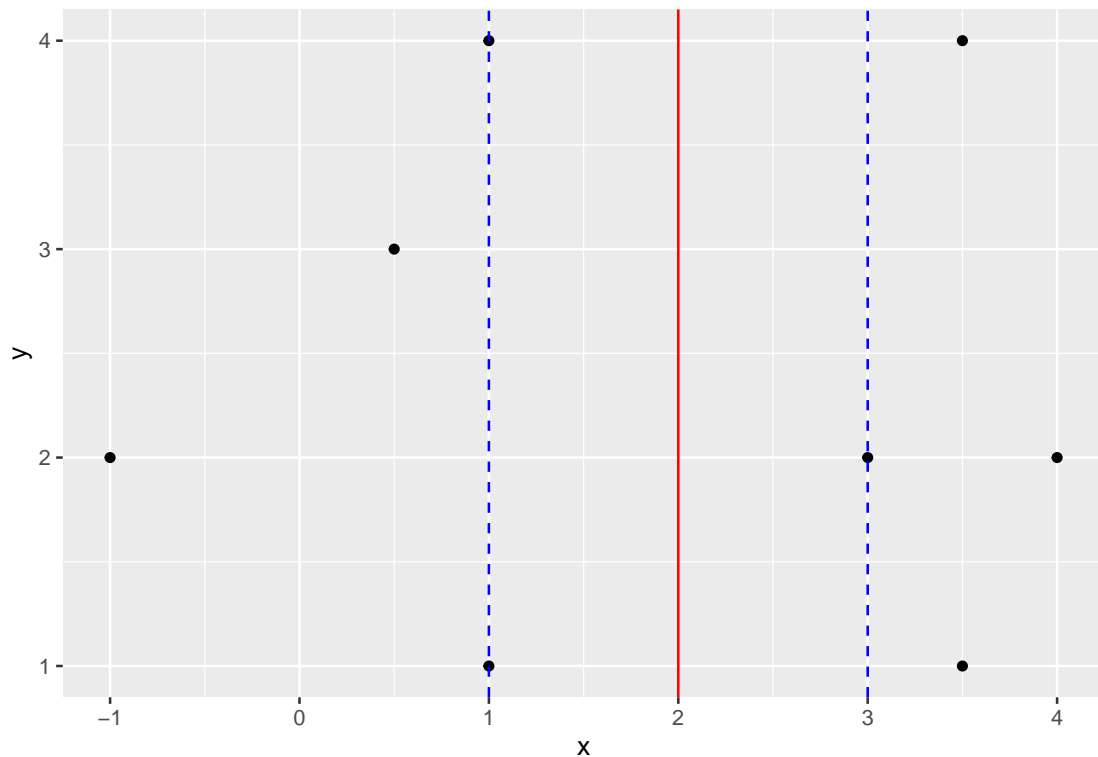
As we can find from the plot, the support vectors are $\{(1, 1), (1, 4), (3, 2)\}$. The margin is 1.

Question 2

If we leave the separating hyperplane as it is, but increase the margin to **1.5**, then what are the support vectors?
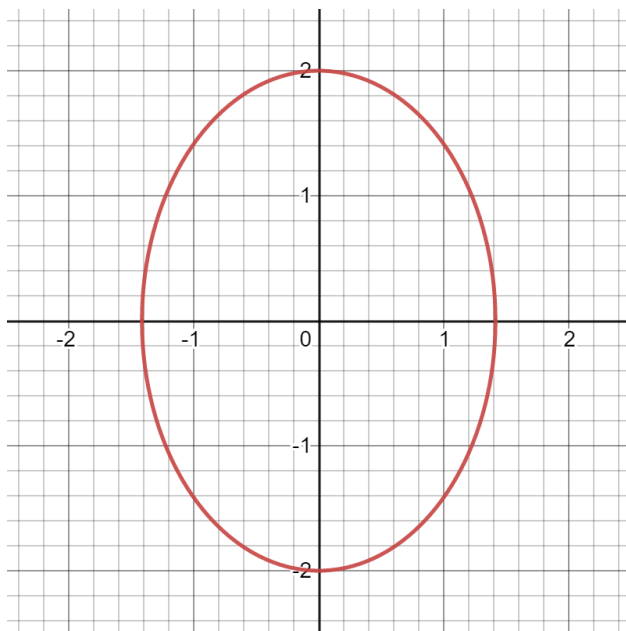
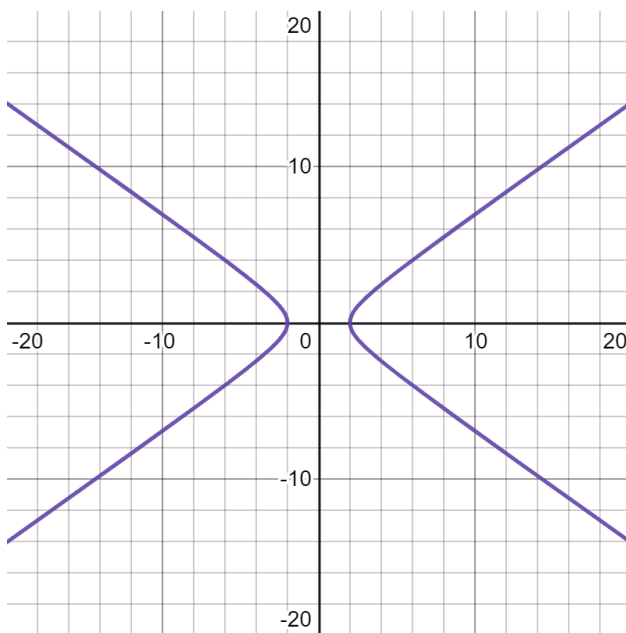As we can find from the plot, the support vectors are $\{(1, 1), (0.5, 3), (1, 4), (3, 2), (3.5, 1), (3.5, 4)\}$.

Exercise 3

Plot the graphs of the following quadratic equations:

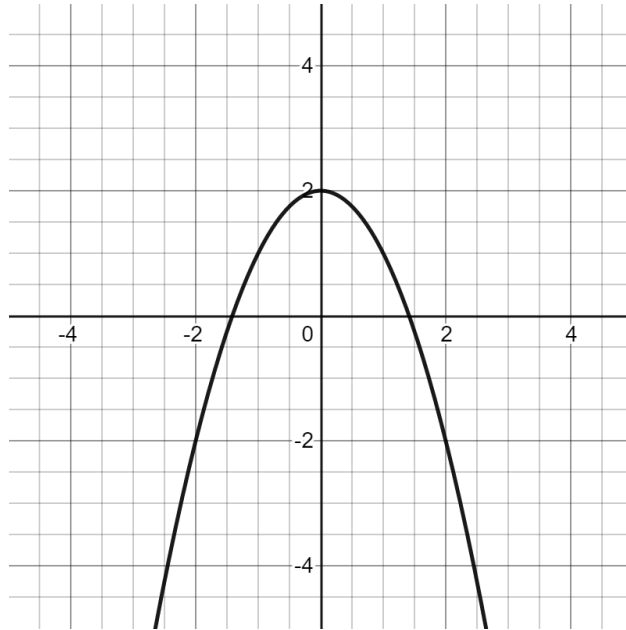1. $\frac{x_1^2}{2} + \frac{x_2^2}{4} = 1$

2. $\frac{x_1^2}{2} - \frac{x_2^2}{1} = 2$

3. $x_2 + x_1^2 = 2$

4. $x_1 - x_2 + x_1 x_2$