

# Exercices - TP5 - SVM

Thomas Laurent

3/03/2018

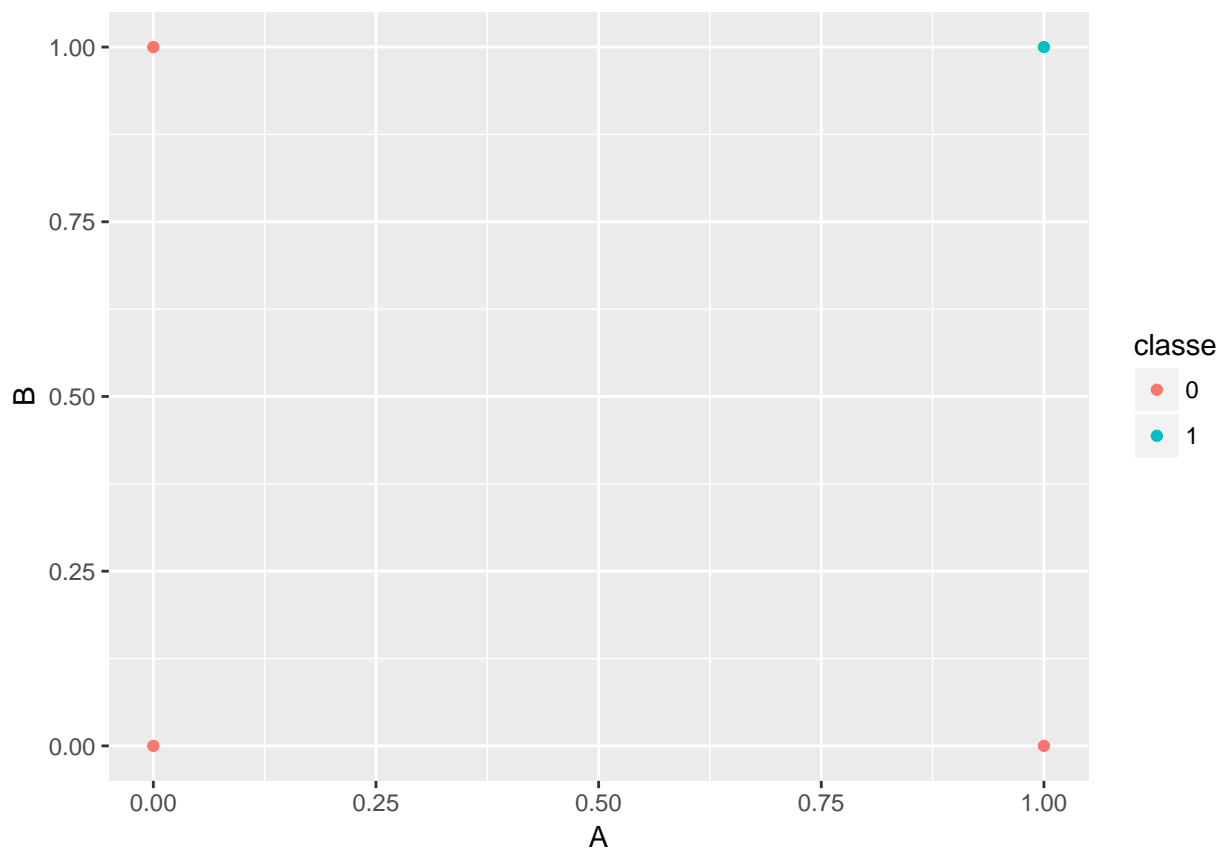
## Exercice I

### Question a)

On trace les données et le résultat est indiqué ci-dessous.

```
##Data
data=data.frame(A=c(0,0,1,1),B=c(0,1,0,1),classe=factor(c(0,0,0,1)))

library(ggplot2)
ggplot(data=data,aes(x=A,y=B,color=classe))+
  geom_point()
```

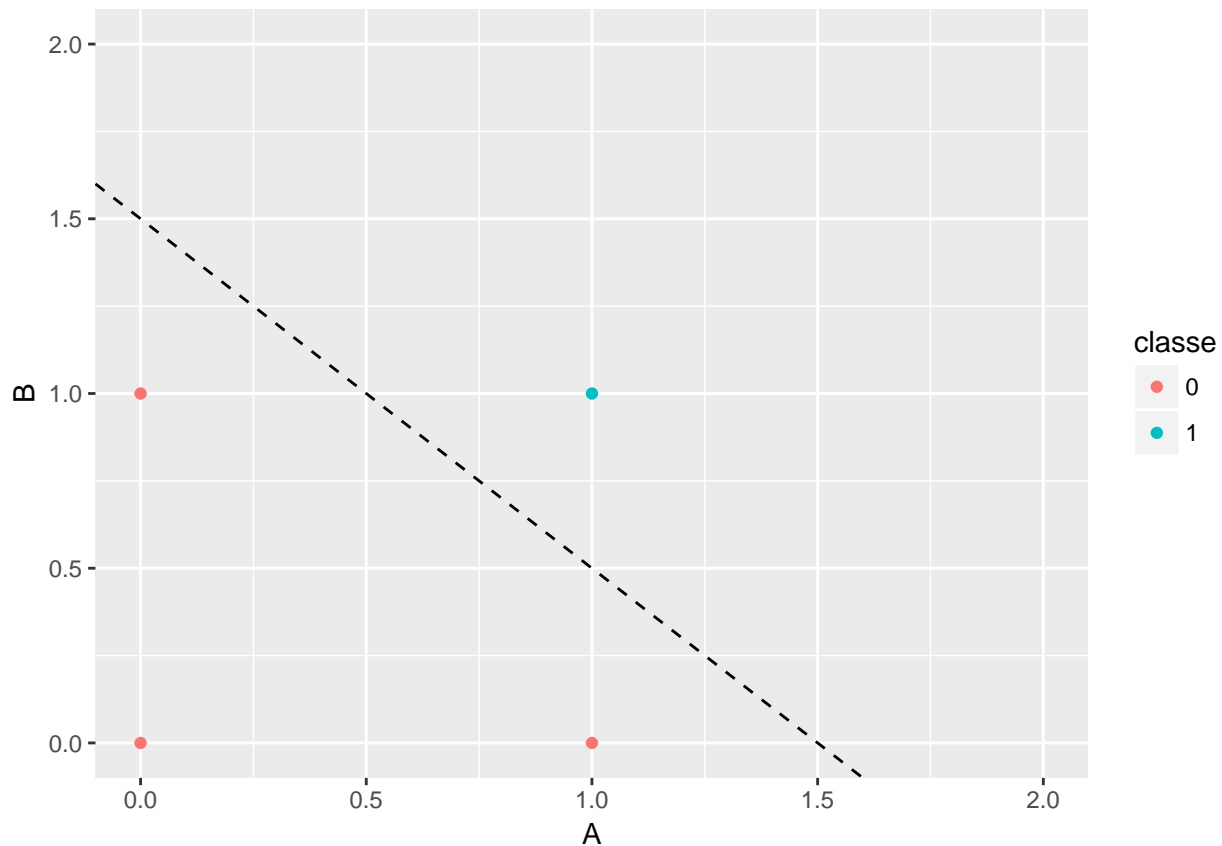


### Question b)

La ligne optimale serait la droite passant par les points (0,5;1) et (1;0,5) qui est parallèle à la droite passant par les points (0;1) et (1;0). On remarque que cette ligne optimale est équidistante pour les points (0;1), (1;0) et (1;1).

```
##Data
data=data.frame(A=c(0,0,1,1),B=c(0,1,0,1),classe=factor(c(0,0,0,1)))

library(ggplot2)
ggplot(data=data,aes(x=A,y=B,color=classe))+
  geom_point()+
  geom_abline(intercept=1.5,slope=-1,linetype=2)+
  scale_x_continuous(limits=c(0,2))+
  scale_y_continuous(limits=c(0,2))
```



### Question c)

Si la ligne passait au-dessus des trois points (0;1), (1;0) et (1;1) ou en-dessous, la marge fonctionnelle serait négative. La solution optimale se trouve donc dans la zone séparant les deux points (0;1), (1;0) et (1;1). Les points (0;1) et (1;0) sont les points les plus proches au point unique de la classe 1 (1;1) avec une distance égale à 1. Ces trois points sont donc vecteurs supports. La frontière est donc parallèle à la droite passante par les points (0;1) et (1;0) et la marge correspondante est égale à la moitié de la distance entre cette droite et le point de la classe 1. Cette droite passe par les points (1;0.5) et (0.5;1). Ainsi l'équation de cette droite est égale à  $B = -A + 1.5$ .

## Exercice 2

### Question a)

On remarque que les points (0;2) et (3;0) sont des vecteurs supports. Le vecteur  $w$  est donc orthogonal à cette droite et donc orthogonal au vecteur (-3;3). De plus, le vecteur ayant une norme égale à 1, le vecteur  $w$  (a,b):  $a^2 + b^2 = 1$  et  $-3a + 3b = 0$  donc le vecteur  $w$  est (1;1).

```
library(e1071)
data=data.frame(x1=c(3,2,1,0,0,0,-4),x2=c(-1,0,1,2,0,-4,0),classe=factor(c(1,1,1,1,0,0,0)))

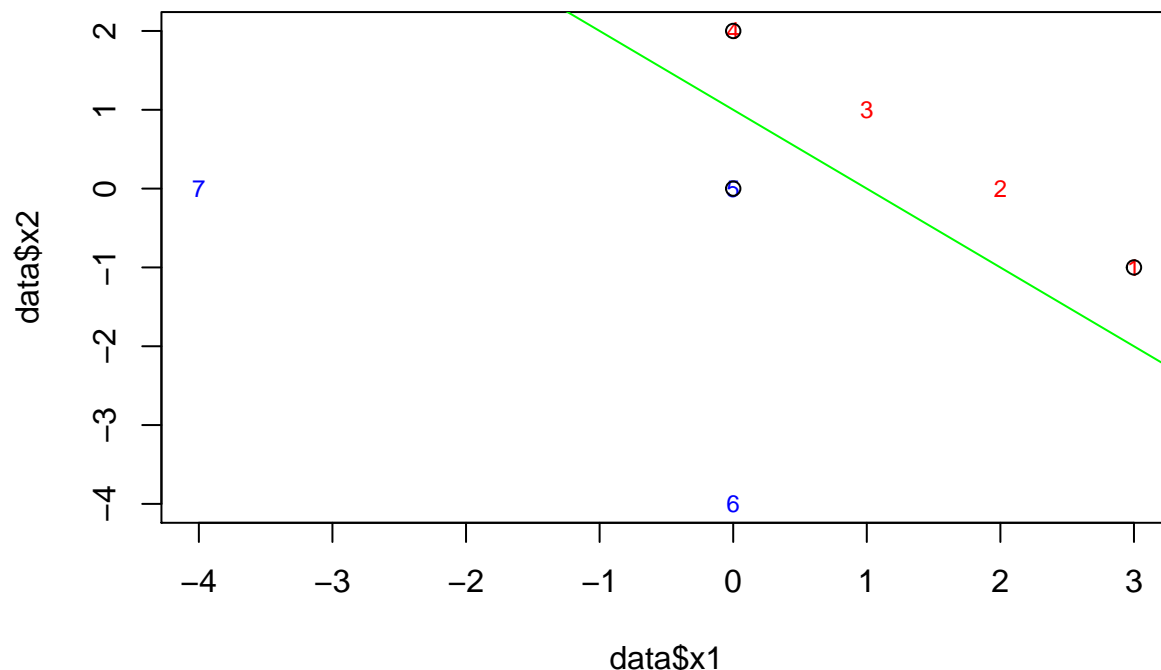
#Algorithme SVM#
Optimisation=svm(classe~ x1+x2, data=data, kernel="linear", scale=F)

#Extraction des parametres de w
beta.0=-Optimisation$rho
beta.1=sum(Optimisation$coefs*data$x1[Optimisation$index])
beta.2=sum(Optimisation$coefs*data$x2[Optimisation$index])
print(paste(beta.1,beta.2))

## [1] "0.9999 0.9997"
```

On vérifie que les paramètres de  $w$  sont égaux à 1 environ, et que a fortiori la norme est environ égale à 1. L'hyperplan séparateur est indiqué en vert dans le graphique ci-dessous.

```
plot(data$x1,data$x2,type="n")
text(data$x1,data$x2,rownames(data),col=c("blue","red")[data$classe],cex=0.75)
points(data$x1[Optimisation$index],data$x2[Optimisation$index],cex=1,col=rgb(0,0,0))
abline(-beta.0/beta.2,-beta.1/beta.2,col="green")
```



## Question b)

```
cat("Table de contigence")

## Table de contigence
round(prop.table(table(data$classe,predict(Optimisation,data))),digits=2)

##
##      0      1
## 0 0.43 0.00
## 1 0.00 0.57
```

On remarque que le pourcentage de mal classé est égal à 0%.

## Question c)

Pour effectuer une validation croisée avec  $k=7$ , on enlève une observation de l'échantillon et on garde les 6 autres pour apprendre le modèle. L'observation restante sera utilisée pour tester le modèle. On réitère ce procédé en enlevant chaque observation une fois. Cela correspond à la méthode Leave-One-Out. On teste différents modèles en faisant varier le paramètre gamma.

```
#Validation croisee#
tuned = tune.svm(classe~x1+x2, data = data, kernel="linear",
                 gamma = seq(0,2,0.1), tunecontrol=tune.control(cross=7))

summary(tuned)

##
## Parameter tuning of 'svm':
##
## - sampling method: leave-one-out
##
## - best parameters:
##   gamma
##     0
##
## - best performance: 0.1428571
##
## - Detailed performance results:
##   gamma      error dispersion
## 1    0.0 0.1428571 0.3779645
## 2    0.1 0.1428571 0.3779645
## 3    0.2 0.1428571 0.3779645
## 4    0.3 0.1428571 0.3779645
## 5    0.4 0.1428571 0.3779645
## 6    0.5 0.1428571 0.3779645
## 7    0.6 0.1428571 0.3779645
## 8    0.7 0.1428571 0.3779645
## 9    0.8 0.1428571 0.3779645
## 10   0.9 0.1428571 0.3779645
## 11   1.0 0.1428571 0.3779645
## 12   1.1 0.1428571 0.3779645
## 13   1.2 0.1428571 0.3779645
## 14   1.3 0.1428571 0.3779645
```

```
## 15  1.4 0.1428571  0.3779645
## 16  1.5 0.1428571  0.3779645
## 17  1.6 0.1428571  0.3779645
## 18  1.7 0.1428571  0.3779645
## 19  1.8 0.1428571  0.3779645
## 20  1.9 0.1428571  0.3779645
## 21  2.0 0.1428571  0.3779645
```

On remarque que l'erreur augmente avec cette méthode (0.14).

### Question 3

Si on retire un des vecteurs support de l'échantillon d'apprentissage, deux cas sont possibles:

-Il y a plusieurs plusieurs vecteurs support par groupe

Dans ce cas-là, la marge n'est pas affectée.

-Pour au moins un des groupes, il n'y a qu'un seul vecteur support avant suppression de l'observation.

Dans ce cas-là, d'autres points plus éloignés de la frontière seront sélectionnés comme vecteurs support. Ainsi, la marge augmente.