STA212

Anthony Kalaydjian - Mathieu Occhipinti

2023-04-29

Exercice 1 : Modélisation probabiliste

- (a)
- (b)

Exercice 2: Classification multi-classes

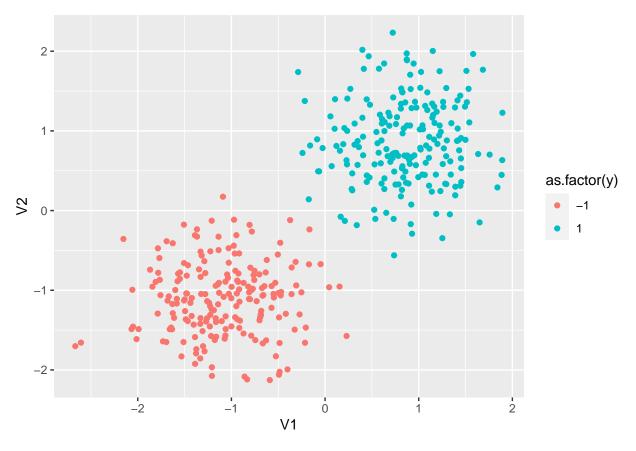
- (a)
- (b)

Exercice 3 : Implémentation d'un perceptron (origine des SVM)

- (a)
- (b)

Importation des données :

```
load(file="X_y.rda")
df <- as.data.frame(cbind(X, y))
names(df) <- c("V1", "V2", "V3", "y")
plt1 <- ggplot(data=df) + aes(x=V1, y=V2, z=y, color=as.factor(y)) + geom_point()
plt1</pre>
```



La variable V3 est une variable d'intercept.

Algorithme perceptron

```
perceptron <- function(X, y){</pre>
  theta <- c(0, 0, 0)
  n <- nrow(X)
  m \leftarrow seq(1, n)
  counter <- 0
  while (length(m) != 0){
    \#sample\ a\ random\ item\ from\ m
    index = sample(m, 1)
    #update theta
    theta <- theta + y[index]*X[index,]</pre>
    #calculate the new m
    temp <- sapply(X=seq(1, n), FUN=function(k) theta%*%X[k,])</pre>
    criterion <- y*temp</pre>
    m <- which(criterion<0)</pre>
    counter <- counter + 1</pre>
  }
  return(list(theta=theta, count=counter))
}
res <- perceptron(X, y)</pre>
```

```
theta.star <- res$theta
count.star <- res$count
theta.star</pre>
```

[1] 3.438710 4.537851 1.000000

count.star

[1] 5

L'algorithme converge en 5 itérations, et nous trouve la valeur de $\theta^* = (3.438710, 4.537851, 1.000000)^T$. **plot**

plt1 + geom_abline(intercept=-theta.star[3]/theta.star[2], slope=-theta.star[1]/theta.star[2])

