SY09 – TP3 – Théorie de la décision

# Exercice 1 :

Cet exercice a pour but de nous faire étudier les performances du classifieur euclidien sur des échantillons issus de 2 classes distinctes.

1. *Génération des différents échantillons*

La fonction à implémenter doit générer une matrice (n,3) renvoyant les échantillons des deux classes différentes w1 et w2. Ces échantillons suivant des lois normales, on va utiliser la fonction R mvrnorm afin de générer aléatoirement la valeur des ces échantillons.

echantillon = function(n,mu1,mu2,a1,a2)

{

library(MASS)

sigma1 <- a1 \* matrix(c(1,0,0,1), 2,2)

sigma2 <- a2 \*matrix(c(1,0,0,1), 2,2)

ech1 <- mvrnorm(n/2, mu1, sigma1);

ech2 <- mvrnorm(n/2, mu2, sigma2);

ech <- rbind(ech1,ech2)

classe <- matrix(1: n)

classe[ 1: n/2 ] <- 1

classe[ n/2 + 1: 600 ] <- 2

res <- cbind(ech, classe)

plot( res[,1], res[,2], asp=1, col = c("orange", "purple")[ res[,3] ] )

return(res )

}

*On obtient les graphiques suivants : avec n = 600, u1 = (0,0)’, u2=(10,0)’.*

**

*a1=1 et a2=1 a1=1 et a2=6*

**

*a1=1 et a2=9 a1=5 et a2=5*

**

*a1=10 et a2=10*

1. *Estimation de la probabilité d’erreur*

* La fonction Theta renvoie un vecteur contenant les estimateurs des paramètres mu1 et mu2 des deux échantillons passés en paramètres.
* La fonction regleEuclidienne calcule la classe retenue par le classificateur pour l’observation x.
* La fonction erreurEstimee calcule la probabilité d’erreur estimée sur l’échantillon D.

theta <- function(D1App,D2App)

{

rbind(apply(D1App,2,mean),apply(D2App,2,mean))

}

regleEuclidienne = function(x,Theta)

{

d1<-(x[1]-Theta[1,1])^2+(x[2]-Theta[1,2])^2

d2<-(x[1]-Theta[2,1])^2+(x[2]-Theta[2,2])^2

res=(d1>d2)+1

return(res)

}

erreurEstimee = function ( D1Test, D2Test, Regle, Theta )

{

n = dim(D1Test)[1] + dim(D2Test)[1]

classement1=apply(D1Test,1,Regle,Theta=Theta)

classement2=apply(D2Test,1,Regle,Theta=Theta)

probaderreur=sum((classement1==2),(classement2==1))/n

return(probaderreur)

}

1. Nous utilisons la fonction suivante pour répéter 10 fois la génération d’échantillons et l’estimation de la probabilité d’erreur associée au classifieur euclidien

erreur <- function(n, mu1, mu2, a1, a2)

{

res = matrix ( 1, 10 )

for ( i in 1:10 )

{

ech <- echantillon(n,mu1,m2,a1,a2)

theta<-Theta(ech[1:n/2,1:2], ech[n/2+1:n,1:2])

res[i] <- erreurEstimee(ech[1:n/2,1:2], ech[n/2 +1:n,1:2], regleEuclidienne, theta)

}

return( res )

}

*On obtient les résultats suivants :*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Situation** | **a1=1 a2=1** | **a1=1 a2=6** | **a1=1 a2=9** | **a1=5 a2=5** | **a1=10 a2=10** |
| Erreur Estimée moyenne | 0 | 0,023333375 | 0,02183333 | 0,01266667 | 0,054166667 |
| Variance | 0 | 9,16667E-16 | 1,8796E-05 | 1,3086E-05 | 7,42284E-05 |

La probabilité d’erreur reste assez faible, même dans le pire des cas (a1=a2=10), elle ne vaut que 5.4%.

## Exercice 2 :

On considère un problème de détection de cibles dans lequel la classe ω1 correspond aux missiles, et la classe ω2 correspond aux avions. Chaque cible est décrite par deux variables X1 et X2 issues de deux capteurs différents. Chaque variable suit, dans chaque classe, une loi normale avec les paramètres suivants :

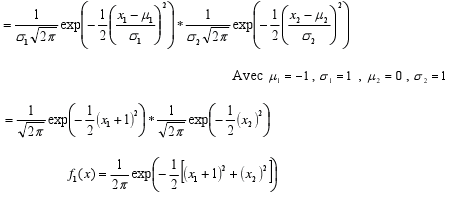
f11(x1) ∼ N(−1, 1) f21(x1) ∼ N(1, 1) f12(x2) = f22(x2) ∼ N(0, 1)

On suppose l'indépendance conditionnelle de X1 et X2. Les densités conditionnelles du vecteur

X= (X1,X2)’ sont donc : f1(x) = f11(x1)f12(x2) dans la classe ω1 f2(x ) = f21(x1)f22(x2) dans la classe ω2.

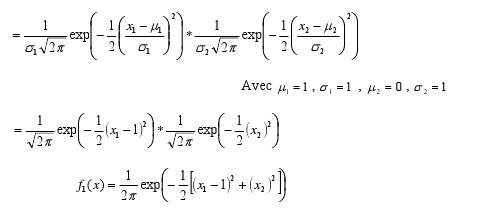
1. Nous cherchons à prouver que les distributions f1 et f2 suivent une loi normale bidimensionnelle dont les paramètres sont à déterminer. Nous allons pour cela calculer le produit des distributions monodimensionnelles constituant f1 et f2 pour identifier les différents paramètres en jeu.

f1 ( x ) = f11(x1) ∗ f12( x 2)



Donc f1(x) ∼ N avec µ=(-1,0)’ et ∑=I

f2 ( x ) = f21(x1) ∗ f22( x 2)



Donc f2(x) ∼ N avec µ=(1,0)’ et ∑=I

1. *Simulation des échantillons*

*Grâce à la fonction mvrnorm de R, nous générons les différents échantillons en utilisant les paramètres déterminés à la question précédente :*

* mu1 = (mu11, mu12) et sigma1 = (sigma11, 0, 0, sigma12 )
* mu2 = (mu21, mu22) et sigma1 = (sigma21, 0, 0, sigma22 )

mu1 <- matrix(c(-1,0), 2, 1)

sigma1 <- matrix(c(1,0,0,1), 2,2)

vect1 <- mvrnorm(n, mu1, sigma1)

mu2 <- matrix(c(1,0), 2, 1)

sigma2 <- matrix(c(1,0,0,1), 2,2)

vect2 <- mvrnorm(n, mu2, sigma2)

On estime les paramètres de cet échantillon grâce aux fonctions colMeans et var des échantillons générés.

Pour la classe ω1 :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Taille de l’échantillon | 10 | 100 | 1000 | 10 000 | 100 000 |
| Mean X1 | -0.5946 | -1.00335 | -1.03772 | -0.998243 | -0.9988623 |
| Mean X2 | 0.26514 | 0.014959 | 0.0063340 | 0.0152704 | 0.00113909 |
| Var X1 | 1.3699 | 1.02254 | 1.03777 | 0.984495 | 1.004036 |
| Var X2 | 0.2812 | 0.80886 | 1.00639 | 0.979752 | 1.001740 |

Et pour la classe ω2 :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Taille de l’échantillon | 10 | 100 | 1000 | 10 000 | 100 000 |
| Mean X1 | 1.14333 | 0.95758 | 0.972161 | 0.9972412 | 1.000985 |
| Mean X2 | -0.03626 | -0.0097831 | 0.0481819 | 0.01261037 | -0.000247 |
| Var X1 | 0.93537 | 0.92436 | 0.993347 | 0.986927 | 1.002532 |
| Var X2 | 1.1693 | 1.116906 | 0.99587 | 0.9943532 | 1.001366 |

On retrouve le principe de la loi des grands nombres : plus l’échantillon est important et plus les estimateurs (moyenne et variance empiriques) tendent vers les valeurs théoriques (espérance et variance).

1. *Détermination des expressions f1(x) et f2(x)*

Densité d’une loi Normale *N* (µ,σ) dans R :

Les courbes sont des cercles d’iso-densité si les équations suivantes sont vérifiées :

f1(x) = K1

f2(x) = K2

K1 et K2 étant des constantes

Soit :

L’équation d’un cercle étant de type (avec r le rayon et a et b les coordonnées du centre du cercle), nous pouvons remarquer que les équations obtenues plus haut sont de même type. f1(x) est représentée par un cercle de rayon ) centré en (-1,0) et f2(x) par un cercle de rayon centré en (1,0).

1. *Utilisation de la règle de Neyman-Pearson*
2. Règle de Neyman-Pearson:

Il faut réduire le problème à une seule variable.

Avec , on a :

Après simplification:

Il ne nous reste donc qu’une seule variable, x1.

1. La première équation est

Or,

On obtient

1. On détermine la fonction frontière avec l’équation précédente

frontiere = function(alpha) {

exp(-2\*qnorm(1-alpha)+2)

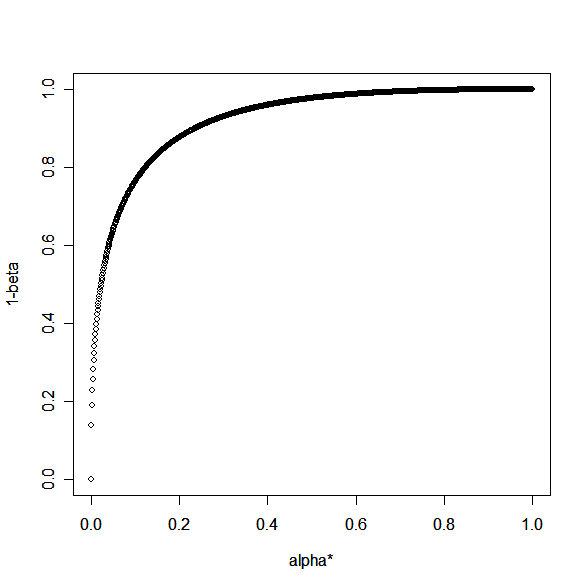
}

On trace les 2 frontières avec la fonction R abline. La première représentation est obtenue avec le plus petit α\* (0.05). On remarque qu’avec la plus petite valeur les points bleu se retrouvent plus nombreux à droite de la frontière (et la frontière est donc plus à gauche). Avec α la probabilité d’erreur (ici, une erreur serait de considérer un élément de la classe 2 comme un élément de la classe 1) on a α ≤ α\*, on peut en déduire que plus α\* sera faible et plus la probabilité d’erreur sera limitée.

1. La valeur de α est estimée en calculant la part des vecteurs vect1X1 qui dépassent la frontière et, pour β, la part des vecteurs vect2X1 qui ne dépassent pas la frontière

|  |  |  |
| --- | --- | --- |
| α\* | 0,05 | 0,1 |
| α | 0.1166667 | 0.06666667 |
| β | 0.8833333 | 0.9333333 |

1. Pour obtenir l’équation de la courbeCOR 1-β = g(α\*), on utilise l’expression :



Or, on a : donc

Alors :

Donc

1. Etude du problème de Bayles

a. Expression de Bayles

La règle de Bayes permet de minimiser le risque (π) en prenant en compte les coûts (c).

b. Frontière de décision dans les différents cas :

On calcule l’équation de la frontière avec

On remarque tout d’abord en calculant les équations des frontières que les cas 2 et 3 sont équivalents (frontière rouge). C’est peu étonnant, augmenter le coût force à minimiser le nombre de points bleus à gauche de la frontière, et augmenter la probabilité de trouver des points bleus revient également à déplacer la frontière sur la gauche. Pour le cas 1, la frontière est en 0 puisque les probabilités et les coûts sont identiques.



c. α est la probabilité de considérer un élément de la classe ω1 comme appartenant à la classe ω2, inversement pour β.

|  |  |  |
| --- | --- | --- |
|  | Cas 1 | Cas 2 et 3 |
| α | 0.138 | 0.563 |
| β | 0.157 | 0.003 |

Pour le cas 1, les probabilités et coûts étant identiques, il est logique de retrouver des valeurs de α et β assez proches.

Pour les cas 2 et 3, α est bien plus élevé que β puisque, dans un cas, le coût d’une erreur concernant un élément de 2 est très élevé, et dans l’autre, la probabilité de trouver un élément de 2 est également nettement supérieure à celle de trouver un élément de 1.