Projet : Distribution Gaussienne multivariée – Fonctions discriminantes

Rendu du projet : Un fichier pdf nommé « NOM_PRENOM » que vous envoyez par mail à : lazhar.labiod@u-paris.fr

L'objectif de ce projet est d'explorer la distribution gaussienne multivariée (simulation et visualisation), et ces liens forts avec les classifieurs linéaires et quadratiques.

Partie 1 : Prise en main des packages **mvtnorm et MixSim**

1. **mvtnorm** : Simulation et visualisation d'une gaussienne multivariée
   a. Simuler et visualiser les contours d'une gaussienne multivariée selon différentes configurations des paramètres Sigma et mu.

```
## Exemple
library(mvtnorm)
x.points <- seq(-3,3,length.out=100)
y.points <- x.points
z <- matrix(0,nrow=100,ncol=100)
mu <- c(1,1)
sigma <- matrix(c(1,1,1,5),nrow=2)
for (i in 1:100) {
  for (j in 1:100) {
    z[i,j] <- dmvnorm(c(x.points[i],y.points[j]),
                 mean=mu,sigma=sigma)
  }
}
contour(x.points,y.points,z)
```

2. **MixSim** : Simulation d'un mélange de gaussiennes multivariées
   a. Prise en main du package **MixSim** : utiliser le papier et le code R illustrant les différents exemples traités dans le papier. L'objectif n'est pas de revoir les développements mathématiques mais de comprendre l'idée d'un mélange gaussien et de s'intéresser en particulier à:
      - la simulation de données à partir d'un mélange gaussien
      - le degré de mélange
      - les proportions des classes
      - la structure de la matrice de variance covariance
   b. Simulation de jeux donnés à partir de mélanges gaussiens selon différents paramétrages
      - Jeu1 : une matrice de taille 500x2 avec deux classes sphériques et bien séparées
      - Jeu2 : une matrice de taille 500x2 avec 3 classes sphériques et un degré de mélange différent de zéro.
      - Jeu3 : une matrice de taille 500x2 avec 3 classes non- sphériques et un degré de mélange différent de zéro.

Partie 2 : Classifieur linéaire et quadratique

1. Implémenter sous R la règle de classification $g_i(X)$
2. Appliquer cette règle aux 3 jeux donnés simulés plus haut.
3. Faites une conclusion de vos expérimentations

$$\boxed{g_i(x) = -\frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu) - \frac{1}{2}\ln|\Sigma_i| + \ln P(c_i)}$$