

## TP7 - Réduction de la dimension par analyse en composante principale et analyse factorielle discriminante - Partie 2

### Q5 - Estimation des moyennes et co-variances des classes

Le calcul de la moyenne  $\bar{\mu}$  de l'ensemble des données d'apprentissage donne les valeurs **-0.02314732** et **0.02047070**.

Le calcul de la valeur de la dispersion intra-classe  $S_w$  donne la matrice suivante :

7.039024	-3.480855
-3.480855	5.646092

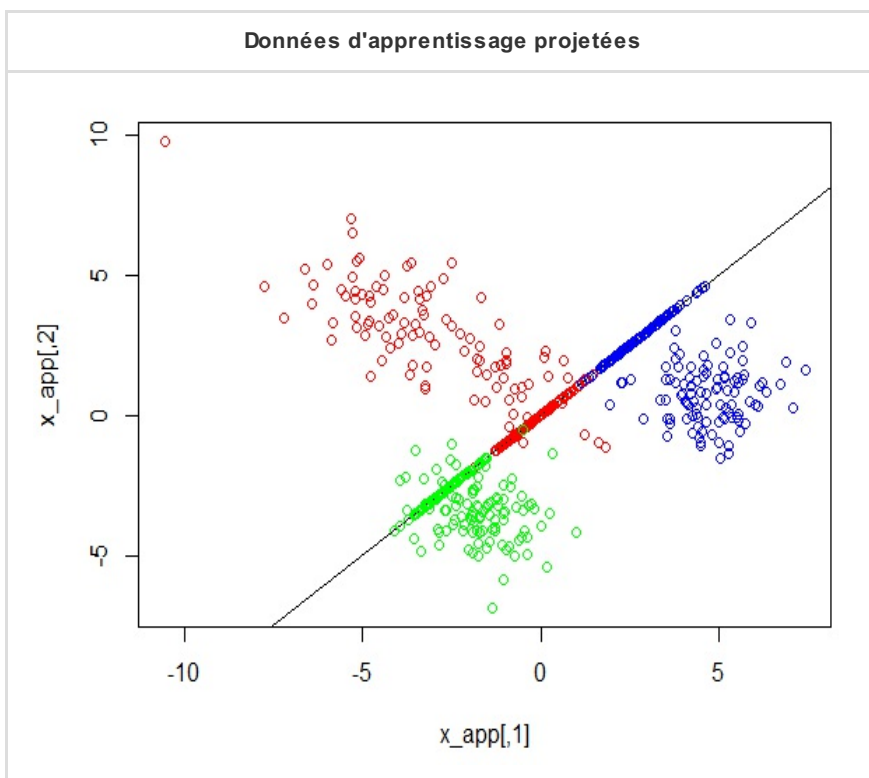
Le calcul de la valeur de la dispersion inter-classe  $S_b$  donne la matrice suivante :

34.7701687	0.5952217
0.5952217	20.9007939

### Q6 - Analyse factorielle discriminante

Sur l'image ci-dessous, on se rend bien compte qu'il est plus facile de discriminer les données d'apprentissage.

En effet, les données de la classe 1 (rouge) ne sont plus mélangées avec les données de la classe 2 (en vert).

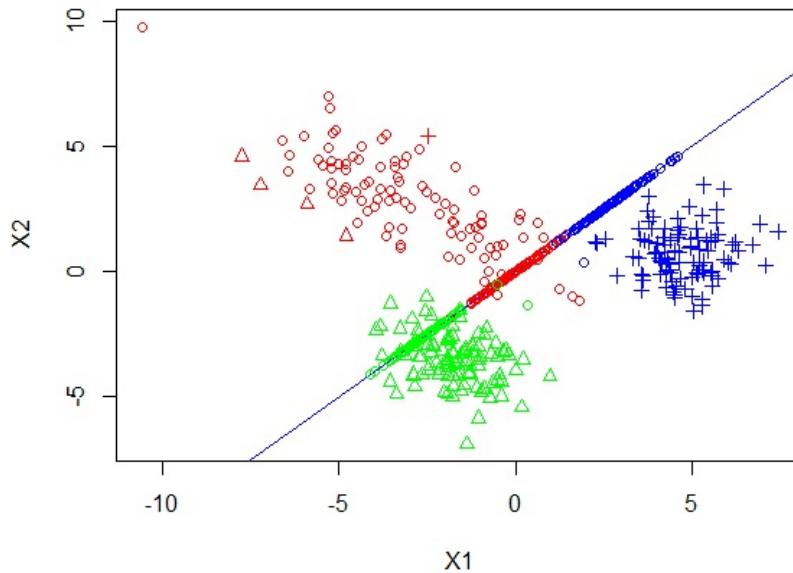


---

## Q7 - Classification par ALD des données test projetées par AFD

---

Données tests classifiées par ALD projetées sur l'axe principal déterminé par AFD



La valeur du taux de bonne classification est d'environ **98%**.

---

## Q8 - Comparaison ACP vs AFD

---

Avec l'ACP, on obtient un taux de bonne classification de **85%**.

Avec l'AFD, on obtient un taux de bonne classification de **98%**.

La raison d'une telle différence s'explique par l'axe de projection utilisé. Celui de l'AFD permet de mieux classer les données que celui de l'ACP car la droite qu'il représente permet de mieux répartir les données en évitant les chevauchements.

---

## Q9 - Comparaison classification dans l'espace d'origine bi-dimensionnel vs sur le premier axe

---