Lilien (利廉) - 76246003
Magalie (周诗曈) – 16241018

Compte rendu de TP (méthode probabiliste)

1) Le vecteur moyenne μ définit le centre de la répartition des vecteurs x appartenant à une même classe d tandis que la covariance Γ définit la distribution des vecteurs x autour du centre. On peut ensuite effectuer une séparation des classes en traçant la surface discriminante : le facteur qui est alors optimisé est la minimisation de la probabilité d'erreur (on cherche à minimiser la probabilité qu'un vecteur x se situe du mauvais côté de la surface discriminante). La surface discriminante correspond à la ligne reliant les points maximum où les densités de probabilités se recouvrent : il s'agit donc de la ligne de niveau des équiprobabilités des deux classes. L'équation suivie par la surface discriminante est une fonction quadratique, ce qui est logique avec nos observations expérimentales sur Matlab.

2) En exécutant le programme du discriminateur avec μ et Γ connus, on obtient au premier tirage $\tau g = 0{,}90$, on obtient donc un écart-type de 0,03 (avec la formule de calcul). Pour un intervalle de confiance à 95%, on peut considérer une barre d'erreur de $2x\sigma$ et on peut ainsi générer les barres d'erreurs sur le graphique. Cela est cohérent lorsqu'on relance le programme plusieurs fois.

Le discriminateur avec μ et Γ connus a une moyenne d'environ 0,90 pour τg tandis que le discriminateur linéaire a une moyenne d'environ 0,83 pour τg. Cette baisse en linéaire est due au fait que l'on prend des estimateurs de μ et Γ pour les calculs. De plus, on choisi un Γ linéaire qui est un Γ commun à toutes les classes, alors que pour le discriminateur avec μ et Γ connus, chaque classe possède son propre Γ.

On remarque, en effectuant plusieurs simulations sur Matlab, que τg(quadratique) est supérieur à τg(linéaire) : cela est du au fait qu'on prend un Γ propre à chaque classe en quadratique alors qu'on prend un Γ commun à toutes les classes en linéaire. Par ailleurs, on remarque que τg(μ et Γ connus) est supérieur à τg(quadratique). Cela est du au fait qu'on prend un estimateur de μ et un estimateur de Γ pour les calculs en quadratique. Graphiquement, on remarque aussi des différences significatives de performances entre les trois discriminateurs.

Pour le nouvel exemple, on remarque que les performances sont quasi-identiques pour les trois discriminants car les classes sont choisies de façon à être bien séparées en trois domaines « linéairement séparés ».

Le quadratique est plus performant que le linéaire lorsque les classes sont réparties aléatoirement et le quadratique est au moins équivalent au linéaire lorsque les classes sont séparées « linéairement », on peut donc effectivement dire que « Qui peut le plus, peut le moins ».

Lorsque la dimension de l'espace est N=38, on remarque que τg(μ et Γ connus) est constant à 1. Par contre, τg(linéaire) diminue lorsque Papp=38 (la longueur de la base d'apprentissage est égale à la dimension de l'espace) et τg(quadratique) diminue lorsque Papp=76 (la longueur de la base d'apprentissage est égale à 2 fois la dimension de la base de l'espace). Il s'agit d'observations et de conjectures, à quoi est dû ce phénomène ?

Quoi qu'il en soit, il est préférable de choisir, si possible, la méthode avec μ et Γ connus pour obtenir un meilleur taux de généralisation.