Compte rendu

Dans ce TP, on a comparé les performances de discrimination de trois discriminateurs : (1)μ、Γ connus; (2)linéaire; (3)quadratique, pour choisir une meilleur discriminateur dans le cas μ、Γ inconnus.

Dans la première partie, on est en cas (1) μ、Γ connus. On a changé les identités de μ_1 et μ_2, et de Γ_1 et Γ_2 pour simuler les différentes densités de probabilité de x. Et après, on observe la frontière dans les différents cas. Dans le cas $\Gamma_1=\Gamma_2$ et $\mu_1\neq\mu_2$, on trouve que la frontière est linéaire. Dans le cas $\Gamma_1\neq\Gamma_2$ et $\mu_1=\mu_2$ et le cas $\Gamma_1\neq\Gamma_2$ et $\mu_1\neq\mu_2$, on trouve que la frontière est conique. La surface discriminante obtenue est arg $\max_{1,2}$ des densités de probabilités. Du coup, on pense que c'est possible de choisir les discriminateurs linéaires ou quadratiques pour discriminer les vecteurs x. Mais quels sont meilleurs, c'est ce que on discute dans la deuxième partie.

Dans la deuxième partie, on compare les performances des discriminateurs linéaires et quadratiques avec le cas μ、Γ connus. D'après TP1, Il est possible d'estimer la barre d'erreur de l'estimation τ_g. Dans l'exemple 1, en général, il y a une perte de performance significative avec le discriminateur linéaire, mais pas significative avec le discriminateur quadratique. Mais dans l'exemple 2, il n'y a pas de perte de performance significative avec le discriminateur linéaire et quadratique. La moyenne de discriminateurs linéaires est toujours plus petite que les discriminateurs quadratiques, et leur variance est toujours plus grand. Du coup, le discriminateur quadratique est mieux dans toutes les situations. Géométriquement, il existe des sections qui ressemblent à "linéaire" dans le courbe conique, du coup, le discriminateur quadratique peut discriminer mieux que linéaire. Mais mathématiquement, on ne sait pas comment expliquer. D'après (2.6), on trouve que quand P_{app} petit ou grand, le discriminateur quadratique performe assez bien comme μ、Γ connus, mais quand P_{app} approche 78, sa performance est très mauvais, pire que linéaire. Et quand P_{app} approche 38, la performance de linéaire est pire que quadratique. Pour améliorer, peut-être on peut définir max(lin, para) pour prendre toujours le meilleur résultat.