

Compte-rendu de TP4

Cédric 16241088

Grégory 16241049

Dans ce cours, le discriminateur qu'on recherche est Bayes. Très différent que les autres discriminateurs qu'on a déjà vus, le Bayes est basé sur la minimisation d'un risque ou des probabilités à postériori. Et le noyau de cette méthode, pour nous, c'est le coût, la valeur de α_{ij} . Le choix indique les erreurs qu'on veut éviter le plus, et bien sûr change la surface discriminante.

Un autre facteur très important, c'est la distribution naturelle des classes, 'le prior' $P(\omega_i)$. On trouve bien quand le prior est loin d'équiprobable, le Bayes performe mieux que le discriminateur de μ et Γ connu. Donc, je devine que **le Bayes est plus convenable que les autres quand la distribution des classes, c'est-à-dire le prior n'est pas équilibré.**

Également, Cédric propose une idée que : **quand le changement du prior, par exemple**

$$P(\omega_1) = k P(\omega_2) = k P(\omega_3) = \dots$$

On peut le considérer comme quasi-équivalent à un changement de matrice de α avec un prior uniforme :

$$\alpha_{ij} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ k & 0 & 1 & 1 \\ k & 1 & 0 & 1 \\ k & 1 & 1 & 0 \end{bmatrix}$$

Or, Grégory n'est pas d'accord. Il est plutôt croit que **la bonne performance pour la classe qui a la plus grande probabilité, c'est grâce à l'apprentissage est plus complet dans cette classe**, n'est pas l'effet d'un équivalent de changement de α .

Ensuite, pour améliorer la performance de Bayes, on pense que puisque c'est 'à posteriori', donc **on peut la contrôler par changer les poids ou le coût α à éviter l'influence des grands coefficients de corrélation.** Par exemple, pour le problème de 10 chiffres, le chiffre 8 est facilement être décidé comme un 6 ou 9, donc on peut augmenter le α_{68} et α_{98} .

En fait, l'apprentissage de Bayes, est juste pour obtenir le μ et Γ : donc, bien sûr la quadratique est mauvais dans ce cas car il faut apprendre trop de paramètres quand la dimension est grande. Et en utilisant cette propriété on trouve que la séparation de coût et le prior avec l'apprentissage. Donc, en réalité la méthode de Bayes est très pratique et elle peut économiser le temps.

Mais pour la partie de réseau de neurones, on ne peut pas justifier théoriquement pourquoi la fonction beta est mieux que la fonction kappa. Peut-être c'est parce que la fonction beta est plus 'sigmoid' ? on ne sait pas. Pour la recherche de la solution plus rapide, on deviner que pour la matrice du coût plus rangé, la différence de ces 2 fonctions va diminuer. Et pour le prior, on constate juste quand il est uniforme la performance est mauvaise. (Ce qui montre encore une fois le Bayes est plus convenables quand le prior n'est pas équilibré.)