

## 2.7 Risque de Bayes

On considère un problème à  $C$  classes  $\{\omega_1, \dots, \omega_C\}$  pour lequel les vecteurs de caractéristiques  $\mathbf{x} \in \mathbb{R}^N$  sont distribués suivant  $P(\mathbf{x}|\omega_c)$ .

Soit  $(R_i)_{i=1, \dots, C}$  une partition de l'espace, le risque de Bayes est défini par :

$$\mathcal{R} = \sum_{i=1}^C \sum_{j=1}^C \int_{R_i} \alpha_{ij} P(\omega_j, \mathbf{x}) d\mathbf{x}$$

où  $\alpha_{ij}$  est le coût si on décide que la classe de  $\mathbf{x}$  est  $\omega_i$  alors que la vraie classe est  $\omega_j$ .

**Pour quelles valeurs de  $\alpha_{ij}$  le risque de Bayes est égal à la probabilité d'erreur ?**

**Attention.** Le choix des coûts  $\alpha_{ij}$  est un choix de valeur et ce n'est pas un choix scientifique (en général).

**Exemple ?**

On peut montrer que :

$$\mathcal{R} = \sum_{i=1}^C \int_{R_i} \rho_i(\mathbf{x}) P(\mathbf{x}) d\mathbf{x}$$

où  $P(\mathbf{x}) = \sum_j P(\mathbf{x}|\omega_j)P(\omega_j)$  est la densité de  $\mathbf{x}$  et

$$\rho_i(\mathbf{x}) = \sum_{j=1}^C \alpha_{ij} P(\omega_j|\mathbf{x})$$

Le discriminateur qui minimise le risque de Bayes peut alors être défini par

$$\hat{d}_{Bayes}(\mathbf{x}) = \arg \min_{i=1, \dots, C} \rho_i(\mathbf{x})$$

L'une des questions qui sera abordée en TP est : **comment appliquer ce concept à un réseau de neurones ?**

On étudiera 2 solutions empiriques :

La première solution consiste à estimer  $P(\omega_i|\mathbf{x})$  avec  $\frac{\beta_i}{\sum_{i=1}^C \beta_i}$  où  $\beta_i = \frac{1}{1+e^{-2\eta_i}}$

La deuxième solution consiste à estimer  $P(\omega_i|\mathbf{x})$  avec  $\frac{\kappa_i}{\sum_{i=1}^C \kappa_i}$  où  $\kappa_i = e^{2\eta_i}$ .

Où  $\eta_i$  est l'entrée du neurones  $i$  de sortie.

### Synthèse

L'objectif de cet enseignement est de vous présenter les concepts de bases du traitement de données, et de vous former à l'analyse des performance de ces techniques.

On a insisté sur 3 concepts :

- attention à l'inversion des matrices en grande dimension
- "qui peut le plus peut le moins" n'est pas vrai en apprentissage
- attention à la barre d'erreur lors de l'analyse des performances.

Ce sont des éléments de base qui vous pourrez retrouver dans toute les techniques d'apprentissage.