TD 8 : MÉTHODES À NOYAUX

COURS D'APPRENTISSAGE, ECOLE NORMALE SUPÉRIEURE, 13 NOVEMBRE 2015

Jean-Baptiste Alayrac
jean-baptiste.alayrac@inria.fr

Résumé. Ce TD a pour but de faire manipuler la notion de noyaux en passant tout d'abord en revue divers exemples. Ensuite on manipulera la notion de feature space engendrée par le noyau.

1. Exemples de noyaux définis positifs

Dans cet exercice, $\mathcal{X}$ est l'ensemble sur lequel sont définis nos noyaux. On rappelle qu'un noyau $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ est défini positif si :

$$\forall(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n, \forall(x_1, \dots, x_n) \in \mathcal{X}^n, \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \geq 0.$$

1) **Opération sur les noyaux :** Soient K et L deux noyaux définis positifs sur $\mathcal{X}$.

(1) Montrer que $K + L$ est aussi un noyau défini positif.

(2) Montrer que KL est aussi un noyau défini positif.

Prouvez que les fonctions K suivantes définies sur $\mathcal{X} \times \mathcal{X}$ dans $\mathbb{R}$ sont des noyaux définis positifs.

2) **Minimum :** $\mathcal{X} = \mathbb{R}^+,\ K(x, y) = \min(x, y)$

3) **Chi-2 :** $\mathcal{X} = \mathbb{R}^+_*,\ K(x, y) = 2\frac{xy}{x+y}$

4) **Sur des ensembles :** $\mathcal{X} = \mathcal{P}(\mathcal{A})$ avec $\mathcal{A}$ un ensemble de cardinal fini. $K(A, B) = \frac{|A\cap B|}{|A\cup B|}$.

5) ***Bonus :*** $\mathcal{X} = \mathbb{N},\ K(n, m) = \mathrm{PGCD}(n, m)$.

2. Manipulation de la distance dans le feature space

6) a) Soient $(x, y) \in \mathcal{X}$. Soit K un noyau défini positif sur $\mathcal{X}$. On rappelle qu'il existe un espace de Hilbert $\mathcal{F}$ pour lequel on a $K(x, y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{F}}$. Exprimer la distance $\|\phi(x) - \phi(y)\|^2_{\mathcal{F}}$ uniquement en fonction de K.

b) Exprimer la distance dans le cas du noyau Chi 2 vue dans l'exercice précédent. Commenter.

7) **Distance à la moyenne dans le feature space**. On considère ici des points $(x_1, \dots, x_n) \in \mathcal{X}^n$ et des réponses binaires associées $(y_1, \dots, y_n) \in \{-1, 1\}^n$. Soit un noyau K défini positif sur $\mathcal{X}$. On se propose d'étudier une règle de classification très simple qui va simplement décider en fonction des distances aux centroïdes respectifs de chaque classe.

a) Soit $x \in \mathcal{X}$. Donnez la distance entre $\phi(x)$ et $\frac{1}{n_+}\sum_{i,y_i=1}\phi(x_i)$ uniquement en fonction de K (où n_+ est le nombre de y_i tels que $y_i = 1$).

b) Proposez une règle de classification simple pour le vecteur x en fonction des données $(x_i)_{i=1,\dots,n}$ et $(y_i)_{i=1,\dots,n}$ et du noyau K.

c) Supposez maintenant que $\frac{1}{n}\sum_i \phi(x_i) = 0$ (donnée centrée dans $\mathcal{F}$) ainsi que $\sum_i y_i = 0$ (autant de points positifs que négatifs). Simplifiez alors la règle déduite précédemment. Pour la suite on supposera cette propriété vraie.

d) **Application au noyau gaussien.** On prend maintenant $\mathcal{X} = \mathbb{R}^d$. On considère alors le noyau gaussien : $K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$. Montrez que la règle de classification déduite précedemment est équivalente à celle du moyennage local vu au cours 2.