Examen : Méthodes à noyaux

Jean-Philippe Vert

A rendre avant le **14 mars 2006**

Exercice 1

Soit $\mathcal{S} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ un ensemble de points de $\mathbb{R}^d$.
1.1. Montrer que rechercher la plus petite sphère qui contient tous les points est équivalent à résoudre le problème d'optimisation:

$$\min_{c \in \mathbb{R}^d, r \in \mathbb{R}} r^2$$

sous les contraintes, pour $i = 1, \ldots, N$:

$$\| \mathbf{x}_i - c \|^2 \leq r^2$$

1.2 Ecrire le problème dual associé.
1.3. Pour $A > 0$, on considère le problème d'optimisation suivant, qui généralise le précédent en cherchant une sphère qui contient "la plupart des points":

$$\min_{c \in \mathbb{R}^d, r \in \mathbb{R}} r^2 + A \sum_{i=1}^{N} \max\left(0, \| \mathbf{x}_i - c \|^2 - r^2\right).$$

Ecrire le problème dual associé, et montrer qu'il peut être résolu quand les N point sont dans l'espace de Hilbert à noyau reproduisant associé à un noyau défini positif quelconque.
1.4. Dans le cas d'un noyau invariant par translation, montrer que ce problème est équivalent à une machine à vecteur de support que l'on décrira.

Exercice 2

On appelle variable de Rademacher une variable aléatoire σ valant 1 avec probabilité $1/2$ et -1 avec probabilité $1/2$.
2.1. Soient $(u_1, u_2, \ldots, u_N)$ des vecteurs d'un espace de Hilbert muni d'un produit scalaire $< ., . >$. Montrer que si $\sigma_1, \sigma_2, \ldots, \sigma_N$ sont N variables de Rademacher indépendantes, alors:

$$\mathbb{E} \sum_{i,j=1}^{N} \sigma_i \sigma_j < u_i, u_j > = \sum_{i=1}^{N} \| u_i \|^2.$$

2.2. Soit K un noyau défini positif sur un espace $\mathcal{X}$, $\mathcal{H}_K$ l'espace de Hilbert à noyau reproduisant associé, et $B_R = \{f \in \mathcal{H}_K, \| f \|_{\mathcal{H}_R} \leq R\}$. Pour tout ensemble de points $\mathcal{S} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ avec $\mathbf{x}_i \in \mathcal{X}$ $(i = 1, \ldots, N)$, et $\sigma_1, \sigma_2, \ldots, \sigma_N$ un ensemble de N variables de Rademacher indépendantes, montrer que

$$\mathbb{E} \sup_{f \in B_R} \left| \sum_{i=1}^{N} \sigma_i f(\mathbf{x}_i) \right| \leq R \sqrt{\sum_{i=1}^{N} K(\mathbf{x}_i, \mathbf{x}_i)}.$$

Exercice 3

Soit K un noyau défini positif sur un espace $\mathcal{X}$, $(\mathcal{H}_K, \| \,.\, \|_{\mathcal{H}_K})$ l'espace de Hilbert à noyau reproduisant associé, $\lambda > 0$, et $\phi : \mathbb{R} \to \mathbb{R}$ une fonction. On suppose que

$$\kappa = \sup_{\mathbf{x} \in \mathcal{X}} K(\mathbf{x}, \mathbf{x}) < +\infty,$$

et on note $B_R = \{f \in \mathcal{H}_K, \| f \|_{\mathcal{H}_R} \leq R\}$. On définit, pour tout $f \in \mathcal{H}$ et $\mathbf{x} \in \mathcal{X}$,

$$R_\phi(f, x) = \phi(f(\mathbf{x})) + \lambda \| f \|^2_{\mathcal{H}_K}$$

3.1. On dit the ϕ est Lipschitz si il existe $L > 0$ tel que, pour tout $x, y \in \mathbb{R}$, $| \phi(\mathbf{x}) - \phi(\mathbf{y}) | \leq L \| x - y \|$. Montrer que dans ce cas, il existe une constante C_1 qu'on déterminera telle que, pour tout $x \in \mathcal{X}$ et $f, g \in B_R$:

$$| R_\phi(f, x) - R_\phi(g, x) | \leq C_1 \| f - g \|_{\mathcal{H}_K}.$$

3.2. On dit que ϕ est convexe si pour tout $u, v \in \mathbb{R}$ et $t \in [0, 1]$, $\phi(tu + (1-t)v) \leq t\phi(u) + (1-t)\phi(v)$. On suppose que ϕ est convexe et que pour tout $x \in \mathcal{X}$, il existe $f_x \in \mathcal{H}$ qui minimise $f \mapsto R_\phi(f, x)$. montrer qu'il existe une constante $C_2 > 0$ telle que

$$\psi(f, \mathbf{x}) \triangleq R_\phi(f, \mathbf{x}) - R_\phi(f_x, \mathbf{x}) \geq C_2 \| f - f_x \|^2_{\mathcal{H}_K}.$$

3.3. Sous les hypothèses des deux question précédentes, en déduire qu'il existe une constante C qu'on déterminera telle que, si X est une variable aléatoire prenant ses valeurs dans $\mathcal{X}$:

$$\forall f \in \mathcal{H}, \quad \mathbb{E}\psi(f, X)^2 \leq C \mathbb{E}\psi(f, X)$$

Exercice 4

Les fonctions suivantes sont-elles des noyaux définis positifs:

$$\forall x, y > 0, \quad K_1(x, y) = \frac{\min(x, y)}{\max(x, y)}$$

$$\forall x, y \in \mathbb{R}, \quad K_2(x, y) = \frac{1}{2 - e^{-\| x - y \|^2}}$$

$$\forall x, y \in \mathbb{R}, \quad K_3(x, y) = \max(0, 1 - |x - y|)$$