- **Fabrice ROSSI (INRIA, Paris). Discrimination de fonctions par machines à vecteurs de support.** Travail en collaboration avec **Nathalie VILLA**

Une machine à vecteurs de support (MVS, [6]) est un outil de discrimination basé sur la maximisation de la marge d'un séparateur affine : les données à classer sont envoyées dans un espace de Hilbert à noyau reproduisant (RKHS) dans lequel on choisit un séparateur affine en minimisant un compromis entre les erreurs de classement du séparateur et la norme du vecteur normal qui le définit. Le passage de l'espace de départ au Hilbert est réalisé de manière implicite en s'appuyant sur son noyau, dont la donnée est suffisante pour construire la MVS.

Dans le cas de données fonctionnelles, le passage par un espace à noyau reproduisant peut sembler superflu, deux classes de fonctions étant généralement linéairement séparable dans un espace de dimension infinie. Cependant, une MVS ainsi construite s'appuie sur une régularisation de type *ridge* dont l'efficacité dans un cadre fonctionnel est assez limitée ([4]).

Il est donc naturel d'étudier les noyaux adaptés aux données fonctionnelles. On étudie deux types de noyaux. Le premier consiste en la combinaison d'un noyau classique de MVS (par exemple le noyau Gaussien) avec un opérateur de projection sur une base tronquée : en s'inspirant de [2], on suppose que les fonctions observées sont éléments d'un Hilbert dont on se donne une base $\{\Psi_j\}_{j\geq 1}$. En utilisant une méthode de validation, on construit ainsi une MVS classique sur les d premières coordonnées sur la base des fonctions observées. Tous les paramètres de la MVS (d inclus) sont choisis automatiquement à partir des données, de façon consistante (l'erreur de la MVS converge avec le nombre de fonctions observées vers l'erreur bayésienne optimale, voir [5]).

Le deuxième type de noyau est adapté aux données fonctionnelles régulières et discrétisées. On considère un opérateur différentiel $L = D^m + \sum_{j=0}^{m-1} a_j D^j$ définit sur l'espace de Sobolev $\mathcal{H}^m([0,1])$, qui conduit à un RKHS $\mathcal{H}_1$ (sous-espace de $\mathcal{H}^m$) dans lequel le produit scalaire est obtenu à partir de L ($\langle u, v\rangle_1 = \langle Lu, Lv\rangle_{L^2}$) (voir [1]). On suppose alors les fonctions observées à valeurs dans $\mathcal{H}_1$ et discrétisées en $t_1, \ldots, t_d$. À chaque fonction x, on associe son interpolation L-spline dans $\mathcal{H}_1$, qui coïncide exactement avec x en $t_1, \ldots, t_d$ et qui est de norme minimale. L'intérêt d'une MVS construite dans $\mathcal{H}_1$ de cette manière est que le noyau utilisé s'appuie sur l'opérateur L et permet donc de considérer les dérivées des fonctions observées, plutôt que les fonctions elles-mêmes, ce qui est fructueux en pratique (voir [5]). Or, on peut construire directement le noyau à partir des valeurs des fonctions observées en $t_1, \ldots, t_d$, sans passer par le calcul des fonctions d'interpolation. On montre de plus qu'une MVS obtenue de cette façon est consistante (voir [7]).

Références

[1] Besse, P. and Ramsay, J.O. (1986). Principal component analysis of sampled curves. *Psychometrica*, **51**, 285-311.

[2] Biau, G., Bunea, F. and Wegkamp, M. (2005). Functional Classification in Hilbert Spaces. *IEEE Transactions on Information Theory*, **51**, 2163-2172.

[3] Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines*, Cambridge University Press, UK.

[4] Hastie, T. and Buja, A. and Tibshirani, R. (1995). Penalized Discriminant Analysis. *Annals of Statistics*, **23**, 73-102.

[5] Rossi, F. and Villa, N. (2006). Support Vector Machine For Functional Data Classification. *Neurocomputing*, **69**, 730-742.

[6] Vapnik, V. (1995). *The Nature of Statistical Learning Theory*,Springer Verlag New York.

[7] Villa, N. and Rossi, F. (2006). Un résultat de consistance pour des SVM fonctionnels par interpolation spline. *Comptes Rendus Mathématiques*, **343**(8), 555-560.