Publications du Groupe de travail STAPH: Statistique Fonctionnelle et Opératorielle

STAPH-2010-01 *Recueil de résumés de l'année 2009-2010*

Alain Boudou, Frédéric Ferraty, Yves Romain, Pascal Sarda, Philippe Vieu et Sylvie Viguier-Pla

Institut de Mathématiques, Université Paul Sabatier, Toulouse, France

.

STAPH: Groupe de travail en Statistique Fonctionnelle et Opératorielle. Présentation des activités 2009-2010.

Alain BOUDOU, Frédéric FERRATY, Yves ROMAIN, Pascal SARDA, Philippe VIEU et Sylvie VIGUIER-PLA

Adresse pour correspondance:
Institut de Mathématiques, Université Paul sabatier
Toulouse

Dans ce document, sont regroupés les résumés des exposés donnés lors des séances du groupe de travail en Statistique Fonctionnelle et Opératorielle STAPH au cours de l'année 2009-2010. Nous tenons tout d'abord à remercier chaleureusement tous les orateurs pour la qualité de leurs contributions qui ont donné lieu à des échanges fructueux. Nous saluons tout particulièrement les nouveaux venus à notre groupe de travail (la moitié des orateurs) et avec lesquels des membres de STAPH ont noués des liens scientifiques ou amorcés des collaborations. De manière générale, nous avons accueilli des doctorants, des jeunes chercheurs ou encore des chercheurs confirmés, suivant en cela un objectif de pluralité qui nous tient à coeur.

On pourra se rendre compte à la lecture de ces résumés que les exposés qui y sont présentés témoignent du dynamisme et de la richesse de la recherche en Statistique Fonctionnelle. Ils portent d'abord sur la diversité et le renouvellement des thèmes abordés qui touchent aux différents aspects de la Statistique Fonctionnelle : modèles nonparamétriques, statistique opératorielle, modèles pour variables fonctionnelles ... Les différentes approches mèlent recherche théorique et applications (économétrie, océanologie, chimiométrie, ...) : c'est un aspect auquel nous tenons particulièrement.

Cette introduction est également l'occasion pour nous de faire un bref bilan de l'activité de notre groupe au terme de sa onzième année d'existence. Tout d'abord, nous félicitons tout particulièrement ici Emmanuel Cabral qui a soutenu brillamment sa thèse sous la direction de A. Boudou et Y. Romain sur le thème de l'étude spectrale des processus stationnaires multidimmensionnels et l'ACP dans le domaine des fréquences. De manière générale, le groupe de travail STAPH a consolidé son implication dans la recherche en statistique qui lui vaut une reconnaissance à l'échelle internationale. Notre groupe a accueilli cette année sept professeurs invités au sein de l'équipe Statistique et Probabilités de l'IMT. Plusieurs d'entre nous ont été invités dans des Universités étrangères. A cela s'ajoute l'activité éditoriale et d'expertise pour des revues de statistique internationales de haut niveau. Signalons à ce sujet le numéro spécial de Journal of Multivariate Analysis paru cette année et coordonné par F. Ferraty. Au plan des conférences internationales, F. Ferraty et P. Vieu ont organisé avec W. Gonzalez-Manteiga une session à la seconde édition du Workshop du groupe ERCIM à Limassol, Chypre, en novembre 2009 (cf. site http://www.cfe-csda.org/ercim09/). Celle-ci portait sur le thème de la "sparsité" et des données fonctionnelles, qui est à la pointe de la recherche dans le domaine. Les rencontres de Statistique Fonctionnelle et Opératorielle organisées chaque année par notre groupe à l'échelle nationale ont cette année été "déplacées" dans le cadre des journées de statistique au mois de mai dernier à Marseille. Nous avons organisé en effet deux sessions lors de ces journées (cf. http://jds2010.univmed.fr/) regroupant jeunes chercheurs et chercheurs confirmés. Ces sessions ont eu un écho très positif. Nous terminons enfin par les événements marquants à venir. Pour la troisième année consécutive nous serons présents au Workshop de l'ERCIM qui se tiendra au mois de décembre à l'Université de Londres (http://www.cfe-csda.org/ercim10/). Le second Workshop international de statistique fonctionnelle et opératorielle se tiendra au mois de juin 2011 à Santander. Le comité scientifique est présidé par R. Cao, F. Ferraty et W. Gonzalez-Manteiga. La première édition qui a eu lieu à Toulouse en juin 2008 a eu un très fort impact et la première liste de conférenciers invités permet de penser qu'il en sera également de même pour cette seconde édition. Notons enfin que la monographie co-édité par F. Ferraty et Y. Romain et regroupant des chapitres d'ordre méthodologique et bibliographique sur les différents aspects de la Statistique Fontionnelle paraitra à l'automne 2010.

Simultaneous confidence bands for nonparametric regression with repeated measurements data

DEGRAS David

Adresse pour correspondance:
University of Chicago, Department of Statistics,
5734 S. University Avenue, Chicago, IL 60637, USA
e-mail: degras@galton.uchicago.edu

Résumé

We look into nonparametric regression with repeated measurements collected on a fine grid. An asymptotic normality result is obtained in a function space equipped with the supremum norm. This result can be used to build simultaneous confidence bands (SCB) for various tasks in statistical exploration, estimation and inference. Two applications are proposed: one is a SCB procedure for the regression function and the other is a goodness-of-fit test for linear regression models. The first one improves upon other available methods in terms of accuracy while the second can detect local departures from a parametric shape, as opposed to the usual goodness-of-fit tests which only track global departures. A numerical study is also provided.

Références

Adler, R. J. (1990). An introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes. IMS, Hayward, CA.

Azzalini, A. and Bowman, A. (1993). On the use of nonparametric regression for checking linear relationships. *J. Roy. Statist. Soc. B* **55**, 549–557.

Benhenni, K. and M. Rachdi, M. (2007). Nonparametric estimation of the average growth curve with a general nonstationary error process. *Commun. Stat., Theory and Methods 36*, 1173–1186.

Degras, D. (2008). Asymptotics for the nonparametric estimation of the mean function of a random process. *Statist. Probab. Lett.* **78**, 2976–2980.

Degras, D. (2009). Nonparametric estimation of a trend based upon sampled continuous processes. *C. R. Acad. Sci. Paris, Ser. I* **347**, 191–194.
Fan, J. and Gijbels, I. (1996). Local Polynomial Modeling and Its Applications. Chapman and Hall, London.
Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis. Springer, New York.
Hall, P. and Titterington, D. M. (1988). On confidence bands in nonparametric density estimation and regression. *J. Multivariate Anal.* **27**, 228–254.
Härdle, W., and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21**, 1926–1947.
Hart, J. D. and Wehrly, T. E. (1986). Kernel regression estimation using repeated measurements data. *J. Amer. Statist. Assoc.* **81**, 1080–1088.
Hart, J. D. and Wehrly, T. E. (1993). Consistency of cross-validation when the data are curves. *Stoch. Proces. Applic.* **45**, 351–361.
Pollard, D. (1990). Empirical processes: theory and applications. *Region. Conf. Ser. Probab. Statist., vol. 2.* Institute of Mathematical Statistics, Hayward, CA.
Robinson, P. M. (1997). Large-sample inference for nonparametric regression with dependent errors. *Ann. Statist.* **25**, 2054–2083.
Sun, J. and Loader, C. (1994). Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.* **22**, 1328–1347.
Yao, F. (2007). Asymptotic distributions of nonparametric regression estimators for longitudinal or functional data. *J. Multivariate Anal.* **98**, 40–56.
Zhao, Z. and Wu, W. B. (2008). Confidence bands in nonparametric time series regression. *Ann. Statist. 36* , 1854–1878.

Concentration de mesures spectrales

VIGUIER-PLA Sylvie* et BOUDOU Alain

* Adresse pour correspondance:
Équipe de Statistique et Probabilités
IMT, Université Paul Sabatier
118 Route de Narbonne
31062 Toulouse Cedex 09
e-mail: viguier@cict.fr

Résumé

Il est bien connu qu'il y a, dans le cas des fonctions stationnaires, dualité entre le temporel et le fréquentiel. Nous nous proposons d'examiner comment se traduit dans le temporel la concentration de la mesure spectrale en un point du spectre.

Nous allons examiner le cas des séries, mais nos résultats s'étendent également au cas des fonctions.

Dans cet exposé, un rapide rappel mathématique présentera les notions de mesure aléatoire, mesure spectrale et opérateur unitaire.

Désignant par μ_Z la mesure spectrale d'une série donnée, nous envisageons ensuite différents types de concentration autour d'un point particulier λ du spectre $\Pi = [-\pi, \pi[$, qui peuvent se résumer de la façon suivante :

- $\mu_Z(\Pi - \{\lambda\}) = 0$ (μ_Z concentrée en λ),
- $\mu_Z(\Pi - \{\lambda\}) \leq \alpha$ (μ_Z "essentiellement" concentrée en λ),
- et $\mu_Z(\Pi - [\lambda - \alpha, \lambda + \alpha]) = 0$ (μ_Z concentrée dans un voisinage de λ).

Enfin, c'est le dernier type de concentration que nous développerons. Pour cela, nous étudierons comment la proximité entre deux opérateurs unitaires se traduit au niveau des mesures spectrales associées.

Nous illustrerons notre propos par des exemples simulés.

Références

Azencott, R. et Dacunha-Castelle, D. (1984). *Séries d'observations irrégulières.* Masson, Paris

Bosq, D. (2000). *Linear Processes in Function Spaces. Theory and Applications.* Lecture notes in statistics, **149**, Springer, Berlin.

Boudou, A. (2007). Groupe d'opérateurs unitaires déduit d'une mesure spectrale - une application. *C. R. Acad. Sci. Paris, Ser. I* **344** 791-794

Boudou, A. and Romain, Y. (2002). On spectral and random measures associated to continuous and discrete time processes. *Stat. Proba. Letters* **59** 145-157.

Brillinger, D. (2001). *Time Series. Data Analysis and Theory.* SIAM, Classics in Applied Mathematics, San Francisco.

Dacunha-Castelle, D. and Duflo, M. (1982). *Probabilités et Statistiques.* Masson.

Riesz, F. and Nagy, B. (1991). *Functional Analysis.* Dover Publications.

Forecasting next-day electricity demand and price using nonparametric functional methods

ANEIROS* Germán, CAO Ricardo, VILAR Juan Manuel
space5mm * Adresse pour correspondance:
Universidade da Coruña, Facultad de Informática,
Campus de Elviña s/n, 15011 A Coruña, Spain
e-mail: ganeiros@udc.es

Résumé

One-day-ahead forecasting of demand and prices electricity are important issues in the competitive electric power markets. These problems have been studied in previous published works, most of which have used ARIMA models, dynamic regression or neural nets.

This talk proposes to obtain demand and prices electricity forecasts based on nonparametric functional (NF) or semi-functional partial lineal (SFPL) regression models. For this, we need to estimate the regression function in each model under dependence conditions. For the NF model we use the estimator proposed in Ferraty and Vieu (2004), while for the SFPL model we consider the estimators proposed in Aneiros-Pérez and Vieu (2008).

First, we give some asymptotic properties showing the good asymptotic behavior of the estimators for the SFPL models, and these asymptotic properties are compared with those of the estimator for the NF model. Then, forecasts based on these functional models are compared with those obtained from a naive method and with ARIMA forecasting. The real data used to perform these comparisons come from the electricity market of mainland Spain in years 2008-09.

All the theoretical results presented in this talk can be seen in Ferraty and Vieu (2004) and Aneiros-Pérez and Vieu (2008).

Références

Aneiros-Pérez G, Cao R, Vilar-Fernández JM (2009) Functional methods for time series prediction: a nonparametric approach. To appear in Journal of Forecasting

Aneiros-Pérez G, Vieu P (2008) Nonparametric time series prediction: A semi-functional partial linear modeling. J Multivariate Anal 99: 834–857

Aneiros-Pérez G, Vieu P (2009) Automatic estimation procedure in partial linear model with functional data. To appear in Statistical papers

Benhenni K, Ferraty F, Rachdi M, Vieu P (2007) Local smoothing regression with functional data. *Computational Statistics* 22: 353–369

Cancelo JR, Espasa A, Grafe R (2008) Forecasting the electricity load from one day to one week ahead for the Spanish system operator. International Journal of Forecasting 24: 588–602

Conejo A, Plazas M, Espínola R, Molina B (2005) Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 20: 1035–1042

Ferraty F, Vieu P (2004) Nonparametric models for functional data, with applications in regression, curve discrimination and time series prediction. J Nonparametric Statist 16: 111–127

Ferraty F, Vieu P (2006) Nonparametric Functional Data Analysis. Springer, New York

Analyse procuste de courbes fermées dans le plan appliquée à l'étude de la forme des otolithes de rougets

NÉRINI* David, MORAT Fabien et LETOURNEUR Yves

* Adresse pour correspondance:
Centre d'Océanologie de Marseille, UMR LMGEM 6117 CNRS
Campus de Luminy, Case 901, 13288 Marseille Cedex 09
e-mail: nerini@univmed.fr

Résumé

Les otolithes sont des concrétions calcaires qui siègent dans l'oreille interne de la plupart des poissons. Ces "pierres d'oreille" ont un rôle physiologique primordial puisqu'elles participent aux fonctions mécano-réceptrices d'audition et d'équilibration des poissons. L'étude des otolithes n'est pas un sujet novateur en océanographie et la bibliographie recèle de nombreuses études ayant trait à ces objets : détermination d'espèces, estimation de l'âge, de la croissance des individus, estimation de stocks, détermination des régimes alimentaires de prédateurs, etc. Les otolithes, de par leur croissance particulière, enregistrent les traits d'histoire de vie des individus (âge, éléments chimiques, reproduction, ...) et sont souvent décrits comme de véritables "boites noires".

Nous nous intéressons dans ce travail à la possibilité de localiser géographiquement la provenance d'animaux de la même espèce en étudiant la forme de leurs otolithes. Il s'agit d'une espèce de Rouget (Mullus barbatus) échantillonnée en divers sites de Méditerranée Occidentale, de Mer Egée et de Mer Noire. Nous disposons pour cela d'un échantillon d'images 2D d'otolithes dont on extrait la bordure extérieure. Les objets ainsi créés peuvent être vus comme des courbes fermées dans le plan. A partir d'éléments empruntés d'une part, à la géométrie différentielle et d'autre part à l'Analyse de Données Fonctionnelles, nous construisons un espace de fonctions dans lequel l'étude de la forme de ces objets va pouvoir être menée en s'affranchissant des effets de taille, de rotation, de réflexion et de translation entre objets. Nous proposons des solutions pratiques pour mener l'analyse procuste généralisée de

ces courbes fermées dans le cadre fonctionnel.

Nous montrons également qu'il est possible d'appliquer les méthodes classiques d'étude et de représentation de la variabilité d'un ensemble de courbes (ACP fonctionnelle) généralisées à des courbes dans le plan. Enfin, une analyse discriminante construite à partir de l'échantillon de courbes fermées montre qu'il est possible de déterminer la provenance des animaux dans 87% des cas.

Références

Do Carmo, M., P., (1976). *Differential geometry of curves and surfaces*, Prentice-Hall Dryden, I., L., Mardia, K., V., (1998). *Statistical Shape Analysis*, J. Wiley & Sons

Gower, J., C., Dijksterhuis, G., B., (2004). *Procrustes Problems*, Oxford Statistical Science Series

Kaziska, D., Srivastava, A., (2007). Gait-Based Human Recognition by Classification of Cyclostationary Processes on Nonlinear Shape Manifolds, *J. Am. Stat. Ass.*, **102**, 1114-1128

Ramsay J. O., Silverman B. W., (2005). *Functional data analysis*, 2d ed., Springer, New-York

Small, C., G., (1996). *The Statistical Theory of Shape*, Springer, New-York

Méthodes récursives en estimation et prévision non paramétriques

AMIRI Aboubacar

Adresse pour correspondance:
Université d'Avignon et des Pays de Vaucluse, Laboratoire d'Analyse
Non Linéaire et Géométrie (EA 2151), F-84018 Avignon
e-mail: aboubacar.amiri@univ-avignon.fr

Résumé

Nous nous intéressons dans cet exposé aux méthodes d'estimation non paramétrique à partir d'estimateurs récursifs à noyau ainsi qu'à leurs applications à la prévision.

Considérons un processus à temps discret $\{X_k : k \geq 1\}$ et on souhaite prédire la valeur de X_{n+h} (prévision à l'horizon $h \in \mathbb{N}^*$).

Plusieurs méthodes existent dans la littérature, cependant il n'existe pas de meilleure méthode que les autres dans toutes les situations.

La méthode de prévision par k-moyennes mobiles, par exemple, consiste à prendre comme prévision la moyenne des observations des k périodes précédentes. La prévision est renouvelée de période en période. Cette méthode est simple d'utilisation, avec pour avantage d'atténuer suffisamment les fluctuations de la série tout en préservant son allure générale, cependant elle s'avère moins performante lorsque l'on prend un nombre élevé de données.

La méthode du lissage exponentiel, quant à elle, prend en compte la prévision de la période antérieure. A cette prévision, on augmente l'écart subi, pondéré d'un coefficient a compris entre 0 et 1. Ceci se traduit, pour $h = 1$, par

$$\widehat{X}_{n+1} = \widehat{X}_n + a(X_n - \widehat{X}_n) \quad (0 < a < 1).$$

Le prédicteur est donc défini par:

$$\widehat{X}_{n+1} = (1-a)\sum_{j=0}^{n-1} a^j X_{n-j}, 0 < a < 1,$$

et peut se calculer de manière récursive par la relation :

$$\widehat{X}_{n+1} = a\widehat{X}_n + (1-a)X_{n+1} \quad (0 < a < 1). \tag{0.1}$$

On constate à partir de (0.1) que la valeur prédite à l'instant $(n+1)$ est une moyenne pondérée entre la valeur estimée faite en n et la dernière observation de la série. L'avantage de cette récursivité est que l'on n'a pas à relisser de nouveau le processus lorsqu'une nouvelle observation s'ajoute à la série. Ce qui n'est pas négligeable car cela permet de réduire considérablement le temps de calcul. Tout comme la méthode de prévision par moyenne mobile, le lissage exponentiel est simple d'utilisation et facilement compréhensible mais son principal inconvénient réside sur le choix de la constante de lissage a . Elle est également moins efficace pour prédire des longues séries. Pour des prévisions à long terme on privilégiera les méthodes de Box-Jenkins [6].

Encore plus robustes que les méthodes de Box-Jenkins, et ayant l'avantage d'une mise en pratique très facile, les méthodes dites non paramétriques sont plus récemment apparues pour tenter d'apporter un nouveau regard sur le problème de la prévision. Le principe de la prévision non paramétrique repose sur le fait que le problème de la prévision peut être vu comme un cas particulier de l'estimation de la régression dans le sens où si l'on suppose que le processus est Markovien, strictement stationnaire et d'ordre 1, alors prédire X_{n+1} revient à estimer l'espérance conditionnelle $E(X_{n+1}/X_n)$ par $\widehat{X}_{n+1} = \widehat{r}_{n-1}(X_n)$, où $\widehat{r}_n(x)$ est l'estimateur de la régression à noyau de $E(X_1/X_0 = x)$ basé sur les observations (X_i, X_{i+1}) pour $i = 1, \dots n-1$.

L'objectif principal de cet exposé est d'améliorer les performances du prédicteur non paramétrique par noyau, en terme de temps de calcul en utilisant des estimateurs récursifs à noyau de la régression pour la construction du prédicteur. Pour cela nous avons besoin d'estimer la régression de manière récursive et donc nous devons d'abord estimer la densité de manière recursive, d'où la répartition de notre travail en trois parties.

Nous introduisons d'abord une famille paramétrique d'estimateurs récursifs de la densité indéxée par un paramètre $l \in [0, 1]$. Leur comportement asymptotique en fonction du paramètre l va nous amener à introduire des critères de comparaison basés sur les biais, variance et erreur quadratique asymptotiques. Pour ces critères, nous comparons nos estimateurs entre eux et aussi comparons notre famille d'estimateurs à l'estimateur non récursif de la densité de Parzen-Rozenblatt. Ensuite, nous définissons à partir de notre famille d'estimateurs de la densité une famille d'estimateurs récursifs à noyau de la fonction de régression. Nous étudions également ses propriétés asymptotiques en fonction du paramètre l. Nous utilisons enfin les résultats obtenus sur l'estimation de la régression pour construire un prédicteur non paramétrique par noyau. Des exemples d'application seront donnés enfin, pour valider la performance (notamment en temps de calcul) de notre méthode de prévision

References

[1] **Ahmad**, I. and **Lin**, P.E. (1976). Nonparametric sequential estimation of a multiple regression function, Bulletin of Mathematical Statitistcs 17, 63-75.

[2] **Amiri**, A. (2009). Sur une famille paramétrique d'estimateurs séquentiels de la densité pour un processus fortemement mélangeant, C. R. Acad. Sci. Paris, Ser. I, 347, 309-314.

[3] **Bosq**, D. (1998). Nonparametric statistics for stochastic processes lecture. Estimation and prediction, Lecture notes in statistics, 2nd ed. Springer-Verlag, New York.

[4] **Bosq**, D. and **Blanke**, D. (2007). Inference and prediction in large dimensions, Wiley Series in Probability and Statistics ISBN 978-0-470-08147-1.

[5] **Bosq**, D. et **Lecoutre**, J.P. (1987). Théorie de l'estimation fonctionnelle, ed. Economica.

[6] **Box**, G.E.P and **Jenkins**, G.M. (1976). Time series analysis, forcasting and control. Holden-Day San Francisco.

[7] **Carbon** , M. (1993). Nonparametric vs parametric forecasting in time series: a computational point view, Applied Stochastic Models and Data Analysis 9, 215-229.

[8] **Carbon**, M. et **Francq**, C. (1995). Estimation non paramétrique de la densité et de la régression - prévision non paramétrique, La Revue de Modulad, ISSN 1145-895X , No. 15, 1-25.

[9] **Collomb**, G. (1981). Estimation non-paramétrique de la régression: Revue bibliographique. International Statistics Review, Vol 49, No. 1, 75-93.

[10] **Collomb**, G. (1981). From nonparametric regression to nonparametric prediction: Survey of the mean square error and original results on the predictogram, Lecuture notes in statistcs, 16, 182-204.

[11] **Collomb**, G. (1984). Propriétés de convergence presque complète du prédicteur à noyau, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 66, 441-460.

[12] **Davies**, H.L. (1973). Strong consistency of a sequential estimator of a probability density function, Bull. Math. Statist. 15, 49-54.

[13] **Deheuvels**, P. (1974). Conditions Nécessaires et Suffisantes de Convergence Ponctuelle Presque Sûre et Uniforme Presque Sûre des Estimateurs de la Densité, Comptes rendus de l'Academie des Sciences de Paris vol. 278, 1217-1220

[14] **Deheuvels**, P. (1974). Sur une famille d'estimateurs de la densité d'une variable aléatoire, Comptes Rendus de l'Academie des Sciences de Paris Serie A, 276, 1013-1015.

[15] **Deheuvels**, P. (1977). Estimation non paramtrique de la densité par histogrammes généralisés, Revue de Statistique Appliquées, XXV, 5-42.

[16] **Devroye**, L. (1981). On the almost everywhere convergence of nonparamertic regression fuction estimates, Ann. Stat. 9, 1310-1319.

[17] **Devroye**, L., **Gyorfi**, L. (1985). Nonparametric Density Estimation: the L_1 View, Wiley, New York, Russian translation: Mir, Moscow, 1988.

[18] **Devroye**, L. and **Wagner**, T. J. (1980). On the L_1 Convergence of Kernel Estimators of Regression Functions With Application in Discrimination, Z. Wahrscheinlichkeitstheorie verw. Gebiete 51, 15-25.

[19] **Devroye**, L. and **Wagner**, T. J. (1980). Distribution-free consistency rsults in nonparametric discrimination and regression function estimation, Ann. Stat., 8, 231-239.

[20] **Greblecki**, W., **Krzyzak**, A. and **Pawlak**, M. (1984). Distribution-free pointwise consistency of regression estimate, Ann. Stat., 12, 1570-1575.

[21] **Greblecki**, W. and **Pawlak**, M. (1987). Necessary and Sufficient Consistency Conditions for a Recursive Kernel Regression Estimate, Journal of Mult. Anal., 23, 67-76.

[22] **Györfi**, L. (1978). Recent results on nonparametric regression estimate and multiple classification, Prob. Contr. Inform. Theory, 10, 509-512.

[23] **Györfi**, L., **Kholer**, M., **Krzyzak**, A. and **Walk** H. (1998). Weak and strong unviersal consistency of semi-recursive kernel and partitioning regression estimates, Statist. Decisions, 16, 1-18.

[24] **Györfi**, L., **Kholer**, M., **Krzyzak**, A. and **Walk** H. (2002). A Distribution-Free Theory of Nonparametric Regression, Springer-Verlag New york. **Isogai**, E. (1984). Joint asymptotic normality of nonparametric recursive density estimators at a finite number of distinct points, J. Japan Statist. Soc. 14, no. 2, 125-135.

[25] **Krzyzak**, A. (1992). Global Convergence of the Recursive Kernel Regression Estimates with Applications in Classification and Nonlinear System Estimation, IEEE Transaction in Information Therory, 38, 1323-1338.

[26] **Krzyzak**, A. and **Pawlak**, M. (1984). Almost everywhere convergence of a recursive regression function estimate and classification, IEEE Transaction in Information Therory, 30, 91-93.

[27] **Liang**, H.Y. and **Baek**, J.II. (2004). Asymptotic normality of recursive density estimates under some dependence assuptions, Metrika, 60, 155-166.

[28] **Masry**, E. (1983). Probability density estimation from sampled data, IEEE Transaction in Information Therory, 29, 696-709.

[29] **Masry**, E. (1986). Recursive Probability Density Estimation for Weakly Dependent Stationary Processes, IEEE Transaction in Information Therory, 32, no 2, 254-267.

[30] **Masry**, E. (1987). Almost Sure convergence of Recursive Density Estimators for stationary Mixing Processes, Statistics and probability Letters 5, 249-254.

[31] **Masry**, E. and **Györfi**, L. (1987). Strong Consistency and Rates for Recursive Probability Density Estimators of Stationary Processes, Journal Multivariate Analysis, Vol 22, 79- 93.

[32] **Roussas**, G.G. (1990). Asymptotic normality of the kernel estimate under dependance conditions: Application to hasard rate, J. Stat. Plann. Inference, 25, 81-104.

[33] **Roussas**, G.G. (1992). Ecxact rates of almost sure convergence of a recursive kernel estimate of a probability density function: Application to regression and hazard rate estimation, Nonparametric Statistics, Vol. 1, 171-195.

[34] **Roussas**, G.G. and **Tran**, L.T. (1992). Asymptotic Normality of the Recursive Kernel Regression Estimate Under Dependence Conditions, The Annals of Statistics, Vol. 20. No. 1., 98-120.

[35] **Tran**, L.T (1989). Recursive Density Estimation Under Dependence. IEEE Transactions on Information Theory, Vol 35 No. 5, 1103-1008.

[36] **Walk**, H. (2001). Strong Universal Pointwise Consistency of Recursive Regression Estimates, Ann. Inst. Stat. Math. Vol 53, Nř 4 691-707 .

[37] **Wang**, L. and **Liang**, H.Y. (2004). Strong uniform convergence of the recursive regression estimators under ϕ-mixing conditions, Metrika 59, 245- 261.

[38] **Wegman**, E. and **Davies**, H.I. (1979). Remarks on Some Recursive Estimators of a Probability Density, The Annals of Statistics , Vol 7, No. 2, 316-327.

[39] **Wertz**, W. (1985). Sequential and Recursive Estimators of the Probability Density, Statistics 16, 277-295.

[40] **Wolverton**, C. and **Wagner**, T.J. (1969). Recursive Estimates of Probability Densities. IEEE Transactions on Systemes Sciences and Cybernetics Vol 5, 307-308.

[41] **Yamato**, H. (1972). Sequential estimation of a continuous propability density function and mode. Bull. Math. Statist. Jap., Vol. 14, 1-12.

Étude spectrale des processus stationnaires multidimensionnels et analyse en composantes principales dans le domaine des fréquences

Soutenance de thèse de Doctorat

CABRAL Emmanuel

* Adresse pour correspondance:
Équipe de Statistique et Probabilités
IMT, Université Paul Sabatier
118 Route de Narbonne
31062 Toulouse Cedex 09
e-mail: cabral@cict.fr

Résumé

Les méthodes statistiques multidimensionnelles ont connu ces dernières décennies des développements très importants tant au niveau théorique (notamment en statistique opératorielle) qu'au niveau des applications (comme, par exemple, le traitement des données fonctionnelles). L'approche dans le domaine des fréquences (et de son outil majeur la transformée de Fourier) permet notamment d'envisager le traitement de fonctions aléatoires et de processus stochastiques particuliers (stationnaires par exemple). Dans l'étude d'un phénomène aléatoire pouvant être modélisé par un processus p-dimensionnel $(X_t)_{t \in T}$, où chaque X_t est à valeurs dans $\mathbb{C}^p$, il peut être intéressant, d'une part, de bien maîtriser les outils spectraux associés à $(X_t)_{t \in T}$ et, d'autre part, dans un but de clarification, d'obtenir un processus q-dimensionnel ($q < p$) résumant le mieux possible $(X_t)_{t \in T}$ en un certain sens.

Les méthodes permettant de résumer une fonction aléatoire continue stationnaire p-dimensionnelle, définie sur $T = \mathbb{Z}$, $\mathbb{R}$, $[-\pi, \pi[$ ou, plus généralement, sur un groupe abélien localement compact existent, elles s'appuient sur un critère étroitement lié à la stationnarité, leur mise en oeuvre est l'Analyse en Composantes Principales (ACP) de la mesure aléatoire associée à $(X_t)_{t \in T}$. Cette ACP, appelée aussi ACP dans le domaine des fréquences de $(X_t)_{t \in T}$, consiste en l'ACP de chacune des composantes spectrales de $(X_t)_{t \in T}$.

Nous pouvons présenter nos contributions en les quatre points suivants:

(*i*) Premièrement, il est bien connu qu'à tout processus stationnaire, on peut lui associer différents outils spectraux tels que la mesure aléatoire, la mesure spectrale à valeurs projecteurs et un opérateur unitaire. Nous réalisons alors un travail de synthèse sur les produits tensoriel et de convolution de mesures aléatoires et spectrales. Cet investissement permet de répondre, dans la pratique, à des problèmes d'interpolation, d'identification de processus spatial ou encore de tranformations de Fourier inverses.

(*ii*) Deuxièmement, étant donné un processus stationnaire périodique, on s'intéresse à la spécificité de ses outils spectraux. Nous définissons une notion de quasi-périodicité et quantifions la proximité entre les mesures aléatoires associées à un processus quasi-périodique et sa version périodique. Nous étudions alors la proximité des ACP, dans le domaine des fréquences, correspondantes.

(*iii*) Ensuite, nous nous intéressons, en troisième lieu, aux outils spectraux associés à un processus cyclostationnaire et, particulièrement, à la mesure aléatoire associée. De plus, nous nous intéressons au produit tensoriel de processus cyclostationnaires en étudiant les propriétés des outils spectraux associés.

(*iv*) Enfin, nous avons porté une attention particulière sur les ACP, dans le domaine des fréquences, centrée et non centrée en établissant une relation explicite entre ces analyses. Quelques applications et perspectives sont proposées en guise de conclusion.

Références

Boudou, A., Cabral, E.N., Romain, Y. (2008) Recent results on random and spectral measures with some applications to Statistics. In *Functional and Operatorial Statistics*, pp 77-83, Contrib. Statist., Physica-Verlag, Springer, Heidelberg.

Boudou, A., Cabral, E.N., Romain, Y. (2010) Centered and non-centered principal component analyses in the frequency domain. *Statistics and Probability letters* 80, 96-103.

Feature Selection in Signal Regression by Blockwise Boosting

GERTHEISS Jan

Adresse pour correspondance:
Department of Statistics
Ludwig-Maximilians-Universität Munich
Akademistr. 1, D-80799 Munich, Germany
e-mail: `jan.gertheiss@stat.uni-muenchen.de`
web: `http://www.statistik.lmu.de/~gertheiss`

Abstract

Signal regression means predicting a scalar quantity from a functional regressor, which has been extensively studied in the chemometrics community, with an excellent summary of tools by Frank and Friedman (1993). Nowadays signal regression is also a topic of research in computational biology, particularly in proteomics (see for example Yasui et al., 2003, Tibshirani et al., 2004, or Goeman, 2008). With the recent surge of interest in functional data, signal regression may be embedded into the framework of functional data, as outlined by Ramsay and Silverman (2005).

More formally, let the data be given by $(y_i, x_i(t))$, $i = 1, \ldots, n$, where y is the response variable and $x(t), t \in I$ denotes a function defined on an interval $I \subset \mathbb{R}$, also called the signal. The regression problem with response y and predictor $x(t)$ can be solved in a nonparametric way (as e.g. done by Ferraty and Vieu, 2006, or Ferraty and Vieu, 2009), or parametrically. A functional linear model for scalar responses has the form

$$y_i = \beta_0 + \int_I x_i(t)\beta(t)dt + \varepsilon_i, \tag{0.2}$$

where $\beta(.)$ is a parameter function and ε_i with $E(\varepsilon_i) = 0$ represents a noise variable, cf. Ramsay and Silverman (2005). The naive approach of fitting by least squares, frequently yields a perfect fit of the data, but with poor predictive value. A more promising approach is based

on regularization with roughness penalties where

$$L(\beta) = \sum_{i=1}^{n} \left\{ y_i - \beta_0 - \int_I x_i(s)\beta(s)ds \right\}^2 + \lambda \int_I |\beta^{(m)}(t)|^q dt$$

is minimized, with $\beta^{(m)}$ denoting the mth derivative. The discretized form of the functional linear model (0.2) is given by

$$y_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i, \tag{0.3}$$

where $x_{ij} = x_i(t_j), \beta_j = \beta(t_j)$ for values $t_1 < \cdots < t_p, t_j \in I$. For simplicity's sake, take values $t_1, \ldots, t_p$ as equidistant, $t_{j+1} - t_j = \Delta$. In the discretized form, derivatives of the coefficient function are replaced by their discrete analogs, i.e. differences of adjacent coefficients.

If response y is not metric but binary, a better model than (0.2), resp. (0.3), is the logistic model

$$P(y_i = 1|x_i) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}, \text{ with}$$

$$\eta_i = \beta_0 + \int_I x_i(t)\beta(t)dt \approx \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j.$$

>From the viewpoint of interpretability, an interesting aspect of signal regression is feature selection. In general, the objective of feature selection is to determine which predictors effect the response. In the case of functional regressors it means to identify the relevant intervals in the signals' domain. In a discretized setting, these intervals turn into groups of measurements $x_j = x(t_j)$ at adjacent observation points t_j.

A possible way how feature selection can be carried out is the application of Boosting techniques. Boosting has been introduced by Freund and Schapire (1996) as a classification tool, but can be used for regression as well (see for example Bühlmann and Yu, 2003). In general, the regression function is based on an ensemble of so-called weak (or base) learners $f^{(m)}(x)$. In common componentwise Boosting (see for example Bühlmann, 2006, or Bühlmann and Hothorn, 2007) the base learner f is a function of only one component of the (p-dimensional) vector $x = (x_1, \ldots, x_p)^T$. That means, in each Boosting iteration one component needs to be selected. By contrast, we do not select single variables but groups $x^{(s)}$ of k adjacent predictors, i.e. observation points (see Tutz and Gertheiss, 2009, or Gertheiss and Tutz, 2009). Let $x^{(s)}$ denote the vector $(x_s, \ldots, x_{s+k-1})^T$, so that one has $s = 1, \ldots, p - k + 1$ groups of adjacent measurement points. For the regression function f we fit linear combinations of these (k adjacent) predictors, i.e. $f(x) = \beta^T x^{(s)}$. That means, $f^{(m)}$ from above is represented by $\beta^{(m)}$ and a distinct block $x^{(s)}$. The selection step now refers to blocks instead of single components. Hence the procedure is called *blockwise* Boosting. Since adjacent measurements are highly correlated, simple least squares estimation of β is not recommended. So we penalize the sum of squared differences between adjacent coefficients β_j, i.e. discretized first derivatives. This makes good sense, since features are assumed to cover intervals, and measurements at adjacent observation points should be linked to the response in a similar way.

In the talk blockwise Boosting for metric and binary responses is presented. It is investigated in simulation studies and applied to real world data from chemometrics (Christensen et al., 2004), proteomics (Petricoin et al., 2002) and sleep studies. Finally, some alternatives using L1-regularization are shortly discussed.

References

Bühlmann, P. and Yu, B. (2003). Boosting with the L2 Loss – Regression and Classification, *Journal of the American Statistical Association*, 98, 324–339.

Bühlmann, P. (2006). Boosting for high-dimensional Linear Models, *The Annals of Statistics*, 34, 559–583.

Bühlmann, P. and Hothorn, T. (2007). Boosting Algorithms – Regularization, Prediction and Model Fitting, *Statistical Science*, 22, 477–505.

Christensen, J., Nørgaard, L., Heimdal, H., Pedersen, J.G. and Engelsen, S.B. (2004). Rapid Spectroscopic Analysis of Marzipan – Comparative Instrumentation, *Journal of Near Infrared Spectroscopy*, 12, 63–75.

Dettling, M. and Bühlmann, P. (2003). Boosting for Tumor Classification with Gene Expression Data, *Bioinformatics*, 19, 1061–1069.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis*, New York: Springer.

Ferraty, F. and Vieu, P. (2009). Additive Prediction and Boosting for Functional Data, *Computational Statistics & Data Analysis*, 53, 1400–1413.

Frank, I.E. and Friedman, J.H. (1993). A Statistical View of Some Chemometrics Regression Tools, *Technometrics*, 35, 109–148.

Freund, Y. and Schapire, R.E. (1996). Experiments with a New Boosting Algorithm, *Proceedings of the Thirteenth International Conference on Machine Learning*, 148–156, San Francisco, CA: Morgan Kaufmann.

Friedman, J.H., Hastie T. and Tibshirani, R. (2001). Additive Logistic Regression – A Statistical View of Boosting, *The Annals of Statistics*, 28, 337–407.

Gertheiss, J. and Tutz, G. (2009). Supervised Feature Selection in Mass Spectrometry-based Proteomic Profiling by Blockwise Boosting, *Bioinformatics*, 25, 1076–1077.

Goeman, J.J. (2008). Autocorrelated Logistic Ridge Regression for Prediction Based on Proteomics Spectra, *Statistical Applications in Genetics and Molecular Biology*, 7, Article 10.

Krämer, N. (2006). Boosting for Functional Data, *COMPSTAT 2006 – Proceedings in Computational Statistics*, 1121–1128, Heidelberg: Physica.

Marx, B.D. and Eilers, P.H.C. (1999). Genaralized Linear Regression on Sampled Signals and Curves – A P-Spline Approach, *Technometrics*, 41, 1–13.

Petricoin, E.F., Ornstein, D.K., Paweletz, C.P., Ardekani, A.M., Hackett, P.S., Hitt, B.A., Velassco, A., Trucco, C., Wiegand, L., Wood, K., Simone, C.B., Levine, P.J., Lineham, W.M., Emmert-Buck, M.R., Steinberg, S.M., Kohn, E.C. and Liotta, L.A. (2002). Serum Proteomic Patterns for Detection of Prostate Cancer, *Journal of the National Cancer Institute*, 94, 1576–1578.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis* (2nd ed.), New York: Springer.

Slawski, M., zu Castell, W. and Tutz, G. (2009). Feature Selection Guided by Structural Information, *Department of Statistics, LMU Munich, Technical Report*, 51, `http://epub.ub.uni-muenchen.de/10251/`.

Tibshirani, R., Hastie, T., Narasimhan, B., Soltys, S., Shi, G., Koong, A. and Le, Q.-T. (2004). Sample Classification from Protein Mass Spectrometry by 'Peak Probability Contrasts', *Bioinformatics*, 20, 3034–3044.

Tutz, G. and Gertheiss, J. (2009). Feature Extraction in Signal Regression: A Boosting Technique for Functional Data Regression, *Journal of Computational and Graphical Statistics* (to appear).

Yasui, Y., Pepe, M., Thomson, M.L., Adam, B.-L., Wrigth Jr., G.L., Qu, Y., Potter, J.D., Winget, M., Thornquist, M. and Feng, Z. (2003). A Data-analytic Strategy for Protein Biomarker Discovery – Profiling of high-dimensional Proteomic Data for Cancer Detection, *Biostatistics*, 4, 449–463.

Zou, H. and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net, *Journal of the Royal Statistical Society B*, 67, 301–320.

.

On the almost sure convergence of the uth geometric conditional quantile with functional covariate

CHAOUCH Mohamed

Adresse pour correspondance:
EDF R&D, ICAME-SOAD, 1 Av. Général de Gaulle,
92141 Clamart, France
e-mail: mohamed.chaouch@edf.fr

Résumé

Geometric quantiles as defined by Chaudhuri (1996) are an extension of univariate quantiles in the multivariate set-up that uses the geometry of multivariate data clouds. A very important application of geometric quantiles is the detection of outliers in multivariate data by means of quantile contours. In this work, we study a nonparametric estimation of the uth geometric conditional quantile when the covariate is a functional random variable. We propose an estimator based on a kernel method and we establish the almost sure convergence rate under the probability measure's concentration property on small balls of the functional variable. An algorithm to compute geometric conditional quantile estimates is also developed. Then our estimation procedure is evaluated by means of some simulated data.

Références

Brown, B. M. (1983), Statistical use of the spatial median. Journal of Royal Statistical Society, B, **45**, 25-30.

Chakraborty, B. (2001). On affine equivariant multivariate quantiles. The Institute of Statistical Mathematics, **53**, 80-403.

Chaouch, M. and Goga, C. (2010). Design-Based estimation for geometric quantiles. To appear in Computational statistics and Data Analysis.

Chaouch, M., Cardot, H. and Cénac, P. (2010). Stochastic approximation to the multivariate and the functional median. Submitted in Compstat 2010.

Chaouch, M., Gannoun, A. et Saracco, J. (2009). Estimation de quantile géométriques conditionnels et non conditionnels. Journal de la Société Francaise de Statistique, **150,2**, 1-27.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach. The Annals of Statistics, **20**, 897-916.

Chaudhuri, P. (1996). On a geometric notation of quantiles for multivariate data. Journal of the American Statistical Association **91**, 862-872.

Cheng, Y. and De Gooijer, J.G., (2007). On the uth geometric conditional quantile. Journal of Statistical Planning and Inference, **137**, 1914-1930.

De Gooijer, J. G., Gannoun, A. and Zerom D. (2006). A multivariate quantile predictor. Communications in Statistics-Theory and Methods, **35**, 133-147.

Haldane, J.B.S. (1948). Note on the median of a multivariate distribution. Biometrika **35**, 414-415.

Kemperman, J.H.B. (1987). The median of a finite measure on a Banach space, In: Dodge, Y. (Ed.), Statistical Data Analysis Based on the L_1 Norm and Related Methods, North-Holland, Amesterdam, 217-230.

Koltchinskii, V. (1997). M-estimation, convexity and quantiles. The annals of Statistics, **25**, 435-477.

Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications. Statistica Neerlandica, **56**, 214-232.

Small, C.G. (1990). A survey of multidimensional medians. International Statistical Review, **58**, 263-277.

Estimation adaptative de quantiles conditionnels

CRAMBES Christophe *, BRUNEL Elodie

Adresse pour correspondance:
Université Montpellier 2, Case Courrier 051, Place Eugène Bataillon, 34095
MONTPELLIER Cedex
e-mail: ccrambes@math.univ-montp2.fr

Résumé

On considère le problème de l'estimation non-paramétrique de quantiles conditionnels (voir [5] pour une large vue d'ensemble des problèmes d'estimation de quantiles conditionnels). On se place dans la cadre suivant: étant donné un réel $\alpha \in]0;1[$, $(X_i, Y_i)_{i=1,\ldots,n}$ sont des observations i.i.d. du couple de variables aléatoires réelles (X, Y) avec

$$Y_i = g_\alpha(X_i) + \varepsilon_i,$$

où $g_\alpha(x)$ est le quantile conditionnel d'ordre α défini par

$$\mathbb{P}(Y \leq g_\alpha(x)|X = x) = \alpha,$$

et où $\varepsilon_1, \ldots, \varepsilon_n$ sont des erreurs i.i.d., indépendantes des X_i. On s'intéresse alors à l'estimation non-paramétrique de la fonction g_α sur la base des observations $(X_i, Y_i)_{i=1,\ldots,n}$. Plusieurs méthodes ont été proposées (voir [7], [3], [4], [6]), basées sur la caractérisation du quantile conditionnel $g_\alpha(x)$ comme solution du problème de minimisation

$$\min_{a \in \mathbb{R}} \mathbb{E}\left[\ell_\alpha\left(Y - a\right)|X = x\right],$$

où ℓ_α est la fonction définie par $\ell_\alpha(u) = |u| + (2\alpha - 1)u$. Cette approche consiste à estimer la fonction g_α dans un certain espace de fonctions (trigonométriques, splines, ondelettes, ...). On propose dans ce travail une méthode permettant de sélectionner automatiquement la dimension de l'espace de fonctions. Ceci se fait en considérant le contraste

$$\gamma_n(t) = \frac{1}{n} \sum_{i=1}^{n} \ell_\alpha \left(Y_i - t(X_i)\right),$$

auquel on rajoute une pénalité choisie de façon adéquate de façon à sélectionner la dimension de l'espace de fonctions en assurant un équilibre entre biais et variance de l'estimateur. Ce type de méthodologie a déjà été développé dans le cadre de l'estimation de la moyenne conditionnelle (voir [1], [2]). On donne ensuite un résultat de consistence pour cet estimateur adaptatif, et des simulations sont effectuées pour analyser son comportement en pratique.

References

[1] Y. Baraud, *Model selection for regression on a random design*, ESAIM P&S **6** (2002), pp. 127-146.

[2] A.R. Barron, L. Birgé, and P. Massart, *Risk bounds for model selection via penalization*, Probab. Theory Relat. Fields **113** (1999), pp. 301-413.

[3] J. Fan, T. -C. Hu, and Y. Truong, *Robust nonparametric function estimation*, Scand. J. of Stat., **21** (1994), pp. 433-446.

[4] X. He, and P. Shi, *Convergence rate of B-spline estimators of nonparametric conditional quantile functions*, J. Nonparam. Statist., **3** (1994), pp. 299-308.

[5] R. Koenker, *Quantile regression.* Econometric Society Monographs, Cambridge, 2005.

[6] R. Koenker, P. Ng and S. Portnoy, *Quantile smoothing splines*, Biometrika, **81** (1994), pp. 673-680.

[7] M. Lejeune, and P. Sarda, *Quantile regression: a nonparametric approach*, Comp. Stat and Data An., **6** (1988), pp. 229-239.

Application of MCMC to change point detection

ANTOCH Jaromír

Adresse pour correspondance:
Charles University in Prague
Department of Probability and Mathematical Statistics
Sokolovská 83, CZ-18675 Praha 8, Czech Republic
e-mail: jaromir.antoch@mff.cuni.cz

Résumé

Change point detection is a topic of interest of many applied and theoretical statisticians since seventies. This lecture will present a less usual approach to the change point estimation, i.e., the use of Bayesian statistics and MCMC approach assuming statistical models with random parameters for modelling the data. Construction of the models is illustrated on the analysis of well known temperature series from Klementinum. Markov chains are generated from the posterior distribution of these parameters using MCMC and desired results are derived from obtained chains. The posterior distribution of the change point and other parameters are estimated from the random samples generated by the combination of the Metropolis-Hastings algorithm and the Gibbs sampler.

Analyzed data comes from Klementinum, a historical building in the central part of Prague. Meteorological station, where the temperature is measured since 1775, is placed in one of its towers. For our analysis we had available year averages of these measurements during the period from 1775 till 1992.

Throughout the lecture we will considered three models with random parameters suitable for description of our observations.

1. Piecewise constant expected value.

2. Two-phase linear model with a jump.

3. Two-phase linear model with a gradual change.

The sequence of average year temperatures measured in Klementinum was already analyzed by many other statisticians using different approaches. We will make a comparison of some of these results.

Keywords: Change point estimation, Markov chain Monte Carlo, Metropolis-Hastings algorithm, Gibbs sampler, Bayesian statistics, Klementinum temperature series.

Références

Antoch J. and Legát D. (2008) Application of MCMC to change point detection. *Aplications of Mathematics*, **53(4)**, 281 – 296, 2008.

Antoch J. and Hušková M. (1999) Estimators of changes. Asymptotics, Nonparametrics and Time Series. Marcel Dekker, Basel, 533 – 577.

Barry D. and Hartigan J. (1993) A Bayesian analysis of change-point problems. JASA **88**, 309 – 319.

Carlin B.P, Gelgand A.E. and Smith A.F.M. (1992) *Hierarchical Bayesian analysis of change point problems.* Applied Statiistics **41**, 389 – 405.