SPATIAL REGRESSION AND PREDICTION WITH INVERSE REGRESSION METHOD

JEAN-MICHEL LOUBES[♭] AND ANNE-FRANÇOISE YAO [§]

Abstract. In this paper, we propose a dimension reduction model for spatially dependent variables. Namely, we investigate an extension of the inverse regression method of Li [15] under strong mixing condition. Li's method is based on estimation of the matrix of covariance of the expectation of the explanatory given the dependent variable, called the inverse regression. Then, we study, under strong mixing condition, the weak and strong consistency of this estimate, using a kernel estimate of the inverse regression. We provide the asymptotic behaviour of this estimate. A spatial predictor based on this dimension reduction approach is also proposed.

1. Introduction

Spatial data analysis is a growing field over the last decade, with various applications in several domains such as soil science, geology, oceanography, econometrics, epidemiology, forestry and many others (see for example [18], [7] or [8] for exposition, methods and applications). In nonparametric spatial estimation, the estimator is, as in the *i.i.d.* observations case, penalized by the dimension of the regressor, which is often called in statistics, "*the curse of dimensionality*". For general references, we refer to Tran [19], Tran and Yakowitz [20], Carbon, Hallin and Tran [5], Carbon et al. [6], Hallin et al. [9], Carbon [3]) or for nonparametric spatial regression see e.g. Biau and Cadre [1], Lu and Chen ([16], [17]), Hallin et al. [10], Carbon, Francq and Tran [4].

Dimension reduction methods are classically used to overcome this issue. Observing an i.i.d sample $Z_i = (X_i, Y_i)$ the aim is to estimate the regression function $m(x) = \mathbf{E}(Y|X = x)$. In the dimension reduction framework, one assumes that there exist Φ an orthonormal matrix $d \times D$, with D as small as possible, and $g : \mathbb{R}^D \to \mathbb{R}$, an unknown function such that the function $m(.)$ can be written as

$$m(x) = g(\Phi\,.X). \tag{1}$$

Model 1 conveys the idea that "less information on X" $(\Phi\,.X)$ provides as much information on $m(.)$ as X. The function g is the regression function of Y given the D dimensional vector $\Phi.X$. Estimating the matrix Φ and then the function g (by nonparametric methods) provides an estimator which converges faster than the initial nonparametric estimator. Φ is uniquely defined under orthogonal transformation and an estimation of this later is done through an estimation of his range $\mathrm{Im}(\Phi^T)$ (where Φ^T is the transpose of Φ) called *Effective Dimensional Reduction space* (EDR).

Various methods for dimension reduction exist in the literature for *i.i.d* observations. For example we refer to the classical theory of principal component analysis, leading to the multiple linear regression, or the generalized linear model in [2], using data transformations leading to linear approximation of the model. The additive models (e.g. Hastie and Tibshirani [12]) deals with methods based on the derivative estimation of the gradient of the regression function $\mathbf{E}(Y|X = x)$ developped in [11], [13] or [21].

In this paper, we focus on the Inverse Regression method, proposed by Li in [15]: if X *is such that for all vector b in $\mathbb{R}^d$, there exists a vector B of $\mathbb{R}^D$ such that* $\mathbf{E}(b^T X|\Phi.X) = B^T(\Phi.X)$ (this later condition is satisfied as soon as X is elliptically distributed), then, if Σ denotes the variance of X, the

---

Key words and phrases. Kernel estimator; Spatial regression; Random fields; Strong mixing coefficient; Dimension reduction.

§ Corresponding author : University Aix-Marseille 2, Campus de Luminy, case 901, 13288 Marseille cedex 09. anne-francoise.yao@univmed.fr.

♭ University Toulouse 3, Institut de Mathématiques de Toulouse.

space $\mathrm{Im}(\Sigma^{-1}\mathbf{var}(\mathbf{E}(X|Y))$ is included into the *EDR space*. Moreover, the two spaces coincide if the matrix $\Sigma^{-1}\mathbf{var}(\mathbf{E}(X|Y))$ is of full rank (recall that $\mathrm{Rank}(\Sigma^{-1}\mathbf{var}(\mathbf{E}(X|Y))) = D$).

Hence, the estimation of the *EDR space* is essentially based on the estimation of the matrix of the *inverse regression* $\mathbf{E}(X|Y)$, where Σ is also estimated using a classical empirical estimator.

In his initial version, Li suggested an estimator based on the regressogram estimate of $\mathbf{E}(X|Y)$ but drawbacks of the regressogram lead other authors to suggest alternatives based on the nonparametric estimation of $\mathbf{E}X|Y$, see for instance [14] or [22] which enable to recover the optimal rate of convergence in $\sqrt{n}$.

In this work, we deal with dimension reduction for spatial data. We focus on a kernel estimate for the inverse regression for spatial dependent data under strong mixing conditions, studied in [1, 6, 4]and provide a consistent estimator of the EDR space. We also use the properties of the inverse regression method to build a *dimension reduction predictor* which correspond to the *nonparametric predictor* of [1]. Indeed, in [1] nonparametric spatial predictor based on kernel method is proposed. It is an interesting alternative to parametric predictor methods such as the krigging method since it does not requires any underlying model. It only requires the knowledge of the number of the neighbors. We will see that the property of the inverse regression method provides a way of estimating this number.

References

[1] G. Biau and B. Cadre. Nonparametric spatial prediction. *Statistical Inference for Stochastic Processes*, 7:327–349, 2004.
[2] D. R. Brillinger. *A Generalized Linear Model with "Gaussian" regressor variables. In A Festschrift for Eric L. Lehmann (P. J. Bickel, K. A. Doksum and J. A. Hodges, eds.* Wadsworth, Belmont, C, 1983.
[3] M. Carbon. Polygone des frÃ©quences pour des champs alÃ©atoires. *Comptes. Rendus. Acad. Sci. Paris I*, 342(9):693–696, 2006.
[4] M. Carbon, C. Francq, and L. T. Tran. Kernel regression estimation for random fields. *Journal of Statistical Planning and Inference*, 137(Issue 3):778–798, 2007.
[5] M. Carbon, M. Hallin, and L. T. Tran. Kernel density estimation for random fields. *Statistics and Probability Letters*, 36:115–125, 1996.
[6] M. Carbon, L. T. Tran, and B. Wu. Kernel density estimation for random fields: The l_1 theory. *Nonparametric Statistics*, 6:157–170, 1997.
[7] N. A.C. Cressie. *Statistics for Spatial Data.* New-York, 1991.
[8] X. Guyon. *Random Fields on a Network - Modeling, Statistics, and Applications.* Springer, New-York, 1995.
[9] M. Hallin, Z. Lu, and L. T. Tran. Kernel density estimation for spatial processes: The l_1 theory. *J. of Mult. Anal.*, 88(1):61–75, 2004.
[10] M. Hallin, Z. Lu, and L. T. Tran. Local linear spatial regression. *Ann. of Stat.*, 32(6):2469–2500, 2004.
[11] W. Hardle and T. M. Stoker. Investigating smooth multiple regression by method of average derivatives. *J. Am. Statist. Ass.*, 84:986–995, 1989.
[12] T. J. Hastie and R. Tibshirani. Generalized additive models (with discussion). *Statist. Sci.*, 1:297–318, 1986.
[13] M. Hirstache, A. Judisky, J. Polzehl, and V. Spokoiny. Structure adaptative for dimension reduction. *Ann of Stat*, 29(6):1537–1566, 2001.
[14] T. Hsing. Nearest-neighbor inverse regression. *Ann. of Stat*, 27(2):697–731, 1999.
[15] K. C. Li. Sliced inverse regression for dimension reduction, with discussions. *Journal of the American Statistical Association*, 86:316–342, 1991.
[16] Z. Lu and X. Chen. Spatial nonparametric regression estimation: Non-isotropic case. *Acta Mathematicae Applicatae Sinica, English Series*, 18(4):641–656, 2002.
[17] Z. Lu and X. Chen. Spatial kernel regression estimation: Weak concistency. *Stat and Probab. Lett*, 68:125–136, 2004.
[18] B. Ripley. *Spatial Statistics.* Wiley, New-York, 1981.
[19] L. T. Tran. Kernel density estimation on random fields. *Journal of Multivariate Analysis*, 34:37–53, 1990.
[20] L. T. Tran and S. Yakowitz. Nearest neighbor estimators for random fields. *Journal of Multivariate Analysis*, 44:23–46, 1993.
[21] Y. Xia, H. Tong, W. Li, and L. et Zhu. An adaptive estimation of dimension reduction space. *JR Stat. Soc., Ser. B*, 64(3):363–410, 2002.
[22] L.Z. Zhu and TK. Fang. Asymptotics for kernel estimate of sliced inverse regression. *Annals of Statistics*, 24:1053–1068, 1996.