Journal of Multivariate Analysis 146 (2016) 261–268

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Relative-error prediction in nonparametric functional statistics: Theory and practice

Jacques Demongeot [a], Ali Hamie [a], Ali Laksaci [b], Mustapha Rachdi [c,*]

[a] Univ. Grenoble-Alpes, Laboratoire AGIM FRE 3405 CNRS, Faculté de Médecine de Grenoble, Université J. Fourier, 38700 La Tronche, France

[b] Laboratoire de Statistique et Processus Stochastiques, Université Djillali Liabès, Sidi Bel-Abbès, BP. 89, Sidi Bel-Abbès 22000, Algeria

[c] Univ. Grenoble-Alpes, Laboratoire AGIM FRE 3405 CNRS, UPMF, UFR SHS, BP. 47, 38040 Grenoble Cedex 09, France

ARTICLE INFO

Article history:
Received 19 November 2014
Available online 9 October 2015

AMS subject classifications:
62G05
62G08
62G20
62G35
62G07
62G32
62G30
Secondary 62E20

Keywords:
Mean square relative error
Nonparametric estimation
Functional data
Regression operator
Positive responses
Asymptotic normality
Small ball property
Stock price
Economic data

ABSTRACT

In this paper, an alternative kernel estimator of the regression operator of a scalar response variable Y given a random variable X taking values in a semi-metric space is considered. The constructed estimator is based on the minimization of the mean squared relative error. This technique is useful in analyzing data with *positive* responses, such as stock prices or life times. Least squares or least absolute deviation are among the most widely used criteria in statistical estimation for regression models. However, in many practical applications, especially in treating, for example, the stock price data, the size of the relative error rather than that of the error itself, is the central concern of the practitioners. This paper offers then an alternative to traditional estimation methods by considering the minimization of the least absolute relative error for operatorial regression models. We prove the strong and the uniform consistencies (with rates) of the constructed estimator. Moreover, the mean squared convergence rate is given and the asymptotic normality of the proposed estimator is proved. Finally, supportive evidence is shown by simulation studies and an application on some economic data was performed.

1. Introduction

In this paper, we aim to advance the modelization of the link between a positive real random variable Y (i.e., the positive response variable) and a functional explanatory variable X (i.e., the covariable).

Noting that, currently, the statistical analysis of the functional data has become usual, due to the progress of computational tools which offer the opportunity to observe phenomena in a real time monitoring. Such continuous data (or functional data, curves, surfaces, …) may occur in different fields of applied sciences, for example in chemometrics

* Corresponding author.
E-mail addresses: jacques.demongeot@agim.eu (J. Demongeot), ali.hamie@agim.eu (A. Hamie), alilak@yahoo.fr (A. Laksaci), mustapha.rachdi@agim.eu, mustapha.rachdi@upmf-grenoble.fr (M. Rachdi).

http://dx.doi.org/10.1016/j.jmva.2015.09.019

(e.g., spectrometric curves), in econometrics (e.g., the stock market index), in biomechanics (e.g., human movements) or in medicine (e.g., electro-cardiograms/electro-encephalograms).

Then, the modelization of the functional variables is becoming more and more popular since the publication of the monograph by Ramsay and Silverman (1997) on the functional data analysis. For an overview on this topic, we refer readers to the monographs by Ramsay and Silverman [26], Ferraty and Vieu [15] and Horvàth and Kokoszka [17]. We return to Cuevas [9] and Bongiorno et al. [5] for an overview and more recent advances in functional data analysis. In this context, studying the co-variability, between a scalar variable and a functional one is an important subject, and there are several ways to explain it.

Let $\mathcal{F}$ be a semi-metric space equipped with a semi-metric d. We consider n pairs of independent random variables (X_i, Y_i) for $i = 1, \dots, n$ that we assume drawn from the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}_+^*$ where $\mathbb{R}_+^*$ is the set of the strictly positive real numbers. A common nonparametric modelization of this relationship is based on the following consideration:

$$Y = m(X) + \varepsilon, \tag{1}$$

where m is the regression operator defined from $\mathcal{F}$ to $\mathbb{R}_+^*$ and ε is a random error variable. Recall that the operator m is usually estimated by minimizing the expected squared loss function $E\left[(Y - m(X))^2|X\right]$. However, this loss function which is considered as a measure of the prediction performance may be unadapted to some situations. Indeed, the use of the least square regression is translated as treating all variables, in the study, as having an equal weight. Thus, the presence of outliers can lead to irrelevant results. So, in this paper we circumvent the limitations of the classical regression by estimating the operator m with respect to the minimization of the following mean squared relative error (MSRE):

$$E\left[((Y - m(X))/Y)^2 \left|X\right.\right] \quad \text{for } Y > 0. \tag{2}$$

This criterium is clearly a more meaningful measure of the prediction performance than the least square error, in particular, when the range of predicted values is large. Moreover, the solution of (2) can be explicitly expressed by the ratio of the first two conditional inverse moments of Y given X. In fact, in order to construct the regression estimator allowing to the best MSRE prediction, we assume that the first two conditional inverse moments of Y given X, that is $g_\gamma(x) := E(Y^{-\gamma}\left|X = x\right.)$ for $\gamma = 1, 2$, exist and are finite almost-surely (a.s.). Then, one can show easily (cf. [23]) that the best mean squared relative error predictor of Y given X is:

$$r(X) = E(Y^{-1}\left|X\right.)/E(Y^{-2}\left|X\right.) = g_1(X)/g_2(X), \quad \text{a.s.} \tag{3}$$

Thus, we estimate the regression operator r which minimizes the MSRE by:

$$\widetilde{r}(x) = \sum_{i=1}^{n} Y_i^{-1} K(h^{-1} d(x, X_i)) / \sum_{i=1}^{n} Y_i^{-2} K(h^{-1} d(x, X_i)) \tag{4}$$

where K is a kernel and $h = h_{K,n}$ is a sequence of positive real numbers.

Although the MSRE is frequently used as a measure of the performance in practice, especially for time series forecasting, the theoretical properties of this regression analysis were not widely studied in the statistical literature and even less in the functional framework. Indeed, the first consideration of the MSRE as a loss function, in the estimation method, is given by Narula and Wellington [22]. Since this work, criteria based on minimizing the sum of absolute relative errors (ARE) and the sum of squared relative errors (SRE) were proposed in different areas (cf. for instance, [27,19] for some models in the software engineering, [6] for some examples in medicine or [7] for some financial applications). Notice that, most of the recent results, in this estimation method, are devoted to linear or multiplicative regression models (cf. [29] for recent advances and references). The nonparametric treatment of these models has not yet been fully explored. As far as we know, only the paper by Jones et al. [18] has paid attention to the study of the nonparametric prediction via the relative error regression method. In fact, these authors studied the asymptotic properties of an estimator minimizing the sum of the squared relative errors by considering both estimation methods (the kernel method and the local linear approach).

Notice that, in the functional data analysis setting, the ordinary least squares regression, which is based on the minimization of the mean squared error, is one of the most popular models. Among various nonparametric estimation methods, for this model, we mention the classical kernel technique which is studied in [15]. In this last, it is established the pointwise almost-complete convergence (a.co.) of the kernel estimator of the regression operator. Then, asymptotic normality results for the same estimator, in the strong mixing case, have been established in [21]. Recently, in [13] it is given the uniform version of the almost-complete convergence rate, in the i.i.d. case. Motivated by its superiority over the classical kernel method, the local linear smoothing technique has been considered, in the functional data setting, by many authors (cf. for instance, [3,11] and references therein). Also, in Crambes et al. [8] it is given the L_p-consistency of a family of M-estimators of the regression operator with a functional regressor, in both cases (1) i.i.d. and (2) strong mixing conditions. An alternative estimation method based on the k nearest-neighbors procedure (kNN) has been used by Kudraszow and Vieu [20] in which the uniform almost-complete consistency of the constructed estimator is proved in the i.i.d case. Recently, Amiri et al. [1] have stated the L_2 and the almost-sure consistencies of a family of recursive kernel estimators of the functional regression.