C. R. Acad. Sci. Paris, Ser. I 357 (2019) 912–916

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistics

Admissibility results under some balanced loss functions for a functional regression model

Résultats d'admissibilité dans une classe de fonctions de perte équilibrées dans un modèle de régression fonctionnelle

Kouider Djerfi [a], Fethi Madani [b], Idir Ouassou [c]

[a] *Université Djilali Liabès de Sidi Belabbès, Algeria*
[b] *Laboratory of Stochastic Models, Statistic and Applications, University of Tahar Moulay, Saida, Algeria*
[c] *National School of Applied Sciences, University of Cadi Ayyad, Marrakech, Morocco*

A R T I C L E I N F O

Article history:
Received 19 March 2019
Accepted after revision 25 October 2019
Available online 15 November 2019

Presented by the Editorial Board

A B S T R A C T

We consider the problem of the nonparametric estimation in a functional regression model $Y = r(X) + \varepsilon$, with Y a real random variable response and X representing a functional variable taking values in a semi-metric space. The aim of this note is to find conditions of admissibility of Stein-type estimators of such a model under a class of balanced loss functions. Our method is to compare the risk with that obtained in the case of a quadratic loss.

R É S U M É

On considère le problème de l'estimation non paramétrique dans un modèle de régression fonctionnelle $Y = r(X) + \varepsilon$, où Y est une variable aléatoire réelle et X est une variable fonctionnelle à valeurs dans un espace semi-métrique. Le but de cette note est de trouver les conditions d'admissibilité des estimateurs de type Stein de ce modèle dans une classe de fonctions de perte équilibrées. Notre méthode consiste à comparer le risque avec celui obtenu dans le cas d'une perte quadratique.

1. Introduction

The technological developments in information processing have made possible the real-time monitoring of many processes in different fields including stock markets, audience rating, meteorology, biochemistry, or medicine. As a consequence, there is an increasing availability of functional data and several interesting applications and case studies have been published. In practice, the use of functional data is often preferable to that of large finite-dimensional vectors obtained by

E-mail addresses: djorfik@gmail.com (K. Djerfi), fethi.madani@univ-saida.dz (F. Madani), i.ouassou@uca.ma (I. Ouassou).

https://doi.org/10.1016/j.crma.2019.10.012

discrete approximations of the functions. The reason is that, whereas the smoothness properties can be easily handled in a functional framework, they are likely to cause collinearity problems that invalidate many standard multivariate (finite-dimensional) techniques. On the other hand, since the beginning of the 1960s, nonparametric statistical methods have been developed intensively, whereas a lot of scientists today collect samples of curves and other functional observations. For an introduction and applications of this field, the books by Ramsay and Dalzell [18], Bosq [4] and Ferraty and Vieu [11] provide some basic methods of analysis along with diverse case studies in several areas, including criminology, economics, archeology, and neurophysiology. For additional case studies or theoretical developments, one can see also, e.g., Benhenni et al. ([1], [2]), Boente and Fraiman [3], Ramsay and Silverman ([19], [20]), Rachdi and Vieu [17], and references therein. For a survey on recent advances in Functional Data Analysis (FDA) and High-Dimensional Statistics (HDS) with recent selected papers in related fields, we can see [9] and [14].

In this work, we are concerned with the problem of the nonparametric estimation of the regression operator with functional data. More precisely, we are interested in the regression model:

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

where Y is the response variable taken value in $\mathbb{R}$, the explanatory variable X is of functional type and the error ε is assumed to be normally distributed with zero mean and variance σ_ϵ^2.

Note that, in the parametric case, Stein in [21] proved the inadmissibility of the usual estimator $Y = (Y_1, \ldots; Y_p)$ of a normal mean $\theta = (\theta_1, \ldots; \theta_p)$ under quadratic loss function, and showed that the estimators of the form:

$$(1 - \frac{b}{a + \|Y\|^2})Y$$

dominated Y for a (respectively b) sufficiently small (respectively large) when $p \geq 3$. A few years later, James and Stein in [15] gave a new proof of the result of [21], where other distributions and other loss functions are considered. Since then, much research has been devoted to improving upon the best invariant estimator of a location vector θ by relaxing the normality assumption, using different loss functions, or considering more general estimators. To quote only a few of them, Bradwein et al. ([5], [6]) considered a generalization of James–Stein estimators under spherical symmetry distributions and, in another work, they present the shrinkage estimators of the location parameter for certain spherically symmetric distributions. The same approach is given by Cellier et al. [8], [7] under elliptically symmetric distributions. Fourdrinier et al. [12] focused on the estimation of a loss function for a spherically symmetric distribution in the general linear model and present in [13] an expository development of loss estimation with substantial emphasis on the setting where the distributional context is normal and its extension to the case where the underlying distribution is spherically symmetric.

Our major objective is to give necessary or sufficient conditions to establish the admissibility of Stein-type estimators; for this we consider a class of balanced loss functions proposed firstly by Zellner in [22], Dey et al. in [10], and Jozani et al. in [16]. But, in all this literature, they use only the quadratic or the weighted quadratic balanced loss function. In our case, more class of balanced loss functions are considered and illustrated with various examples. The main idea of the method used is based on the comparison of the risk calculated under the different balanced loss functions proposed with the risk obtained by the quadratic loss function. Consequently, we are interested in Stein's improvement of the kernel estimator $\hat{r}(x)$ in the model (1) using the balanced loss function

$$L(r(x), \hat{r}_g(x)) = \omega\phi(\hat{r}(x), \hat{r}_g(x)) + (1 - \omega)\phi(r(x), \hat{r}_g(x)) \tag{2}$$

where $0 < \omega < 1$, $\phi(s, t)$ is a differentiable function from $\mathbb{R}^2$ to $\mathbb{R}$ and g is a positive operator from $E \times \mathbb{R}^n$ to $\mathbb{R}$, where (E, d) is a semi-metric space. The Stein-type estimator $\hat{r}_g$ is defined by $\hat{r}_g(x) = \hat{r}(x) - g(x, Y)$.

Contrary to what has been shown by Dey et al. in [10], in some cases the question about admissibility and dominance may depend mainly on ω on the one hand. On the other hand, we note that results for balanced loss functions may be inferred directly from the corresponding results for quadratic and weighted quadratic loss functions, as demonstrated by Jozani et al. in [16].

2. Model and estimators

We consider the regression model (1). From the sample $(X_1, Y_1), \ldots, (X_n, Y_n)$, we estimate the regression operator r by:

$$\hat{r}(x) = \sum_{i=1}^{n} Y_i W_i(x) = (Y_1, \ldots, Y_n)^t (W_1(x), \ldots, W_n(x)) := Y^t W$$

where $W_i(x) = \frac{\Delta_i(x)}{\sum_{i=1}^n \Delta_i(x)}$, $\Delta_i(x) = K\left(\frac{d(x, X_i)}{h_n}\right)$ and K is a real function defined on $\mathbb{R}^+$. The bandwidth h_n is such that: $\lim_{n \to \infty} h_n = 0$.

For the convergence of the estimator $\hat{r}(x)$, we can see [11] for more detailed proofs.

3. Balanced loss functions and associated risk

The balanced loss function is given by the following formula:

$$L(r(x), \hat{r}_g(x)) = \omega\phi(\hat{r}(x), \hat{r}_g(x)) + (1-\omega)\phi(r(x), \hat{r}_g(x))$$

where $\hat{r}_g(x) = \hat{r}(x) - g(x, Y)$ is the Stein-type estimator.

The risk of any estimator $\hat{r}_g$ will be denoted by

$$\mathcal{R}(r(x), \hat{r}_g(x)) = \mathbb{E}\{L(r(x), \hat{r}_g(x))\}.$$

Along this work, we consider distance-based balanced loss functions, which means that we have:

$$\phi(s, t) = \psi(s - t)$$

where the representing function ψ is a positive function such that $\psi(0) = 0$. So the expression for the difference of the risk is:

$$\begin{aligned}\triangle\mathcal{R}(x) &= \mathcal{R}(r(x), \hat{r}_g(x)) - \mathcal{R}(r(x), \hat{r}(x)) \\ &= \omega\mathbb{E}\{\psi(g(x, Y))\} + (1-\omega)\mathbb{E}\{\psi(r(x) - \hat{r}(x) + g(x, Y)) - \psi(r(x) - \hat{r}(x))\}.\end{aligned}$$

The method used later is to compare the difference in risk $\triangle\mathcal{R}(x)$ with that obtained by considering the quadratic loss function $L_0(r(x), \hat{r}_g(x)) = (r(x) - \hat{r}_g(x))^2$. If we note $\mathcal{R}_0(x)$ the risk obtained relatively to the quadratic loss function and $\triangle\mathcal{R}_0(x)$ the difference in risk, we have:

$$\begin{aligned}\triangle\mathcal{R}_0(x) &= \mathcal{R}_0(r(x), \hat{r}_g(x)) - \mathcal{R}_0(r(x), \hat{r}(x)) \\ &= \mathbb{E}\{(g^2(x, Y) - 2(\hat{r}(x) - r(x))g(x, Y)\}.\end{aligned}$$

4. Admissibility and dominance. Main results

The objective in this section is to show the influence of the choice of the balanced loss function on the possibility of improvement of the usual estimator (the kernel estimator). We begin with examples where the choice of BLF leads to finding a solution to this problem of improvement by shrinkage; then we cite examples of the opposite case where the usual estimator can not be dominated by another estimator obtained by contraction. The following lemma gives us an explicit formula of the difference in risk $\triangle\mathcal{R}_0(x)$.

Lemma 4.1. *For every differentiable function $g : E \times \mathbb{R}^n \longrightarrow \mathbb{R}$, we have:*

$$\triangle\mathcal{R}_0(x) = \mathbb{E}\{g^2(x, Y)\} - 2\sigma_\epsilon \sum_{i=1}^{n} W_i(x)\mathbb{E}\{\frac{\partial(g(x, Y))}{\partial Y_i}\}$$

where $\partial g(x, Y)/\partial Y_i$ is the partial derivative of $g(x, Y)$ with respect to the i-th variable Y_i.

We need to apply the last lemma for the quadratic balanced loss function in order to have sufficient conditions on function g such that the estimator $\hat{r}_g$ dominate $\hat{r}$ under this balanced loss function through the following theorem.

Theorem 1. *Let $g : E \times \mathbb{R}^n \to \mathbb{R}$ be a positive differentiable function. A sufficient condition for the estimator $\hat{r}(x) - g(x, Y)$ to dominate $\hat{r}(x)$ under the quadratic balanced loss function is:*

$$g^2(x, Y) - 2(1-\omega)\sigma_\epsilon \prec W(x), \nabla g(x, Y) \succ \leq 0. \tag{3}$$

Sketch of the proof of Theorem 1. The basic ideas of this proof are as follows.

(i) We use the quadratic balanced function given by:

$$L(r(x), \hat{r}_g(x)) = \omega(g(x, Y))^2 + (1-\omega)(r(x) - \hat{r}(x) + g(x, Y))^2$$

to explicit the associated difference in risk by the form:

$$\triangle\mathcal{R}(x) = \mathbb{E}\{g^2(x, Y)\} - 2(1-\omega)\mathbb{E}\{(\hat{r}(x) - r(x))g(x, Y)\}.$$

(ii) We get sufficient conditions for the estimator $\hat{r}(x) - g(x, Y)$ to dominate $\hat{r}(x)$ by applying the formula

$$\mathbb{E}\{(\hat{r}(x) - r(x))g(x, Y)\} = \sigma_\epsilon \sum_{i=1}^{n} W_i(x)\mathbb{E}[\nabla_i g(x, Y)].$$

In order to save space, the proof may be obtained on simple request to one of the authors.

5. Examples

In this section and for the purpose of illustrating our results, we propose some examples by considering different classes of balanced loss functions.

Example 1 *(Weighted quadratic balanced loss).*
Let us define the weighted quadratic balanced loss function by:

$$\begin{aligned} L(r(x), \hat{r}_g(x)) &= \omega\tau(\hat{r}(x) - \hat{r}_g(x))^2 + (1-\omega)\tau(r(x) - \hat{r}_g(x))^2 \\ &= \omega\tau(g(x, Y))^2 + (1-\omega)\tau(r(x) - \hat{r}(x) + g(x, Y))^2, \qquad \tau > 0. \end{aligned} \tag{4}$$

A sufficient condition for the estimator $\hat{r}(x) - g(x, Y)$ to dominate $\hat{r}(x)$ under (4) is

$$\tau g^2(x, Y) - 2(1-\omega)\tau\sigma_\epsilon \prec W(x), \nabla g(x, Y) \succ \leq 0. \tag{5}$$

Example 2 *(Logistic balanced loss).*
We define the logistic balanced loss function defined by:

$$\begin{aligned} L(r(x), \hat{r}_g(x)) &= \omega(-\ln\frac{1}{2}(1 + \mathrm{e}^{-g^2(x,Y)})) \\ &\quad + (1-\omega)(-\ln\frac{1}{2}(1 + \mathrm{e}^{-(r(x)-\hat{r}(x)+g(x,Y))^2})). \end{aligned} \tag{6}$$

Under this logistic balanced loss function (6), the kernel estimator $\hat{r}(x)$ is admissible.

Example 3 *(Reflected normal balanced loss).*
We define the reflected normal balanced loss function defined by:

$$\begin{aligned} L(r(x), \hat{r}_g(x)) &= \omega(1 - \mathrm{e}^{-ag^2(x,Y)}) \\ &\quad + (1-\omega)(1 - \mathrm{e}^{-a(r(x)-\hat{r}_g(x))^2}). \end{aligned} \tag{7}$$

For the regression model (1), and under this reflected normal balanced loss function (7), the kernel estimator $\hat{r}(x)$ is admissible.

Example 4 *(Linex balanced loss).*
Under the Linex balanced function:

$$\begin{aligned} L(r(x), \hat{r}_g(x)) &= \omega(\mathrm{e}^{ag(x,Y)} - ag(x, Y) - 1) \\ &\quad + (1-\omega)(\mathrm{e}^{a(r(x)-\hat{r}(x)+g(x,Y))} - a(r(x) - \hat{r}(x) + g(x, Y)) - 1). \end{aligned} \tag{8}$$

A necessary condition for the estimator $\hat{r}_g(x)$ to dominate the kernel estimator $\hat{r}(x)$ is:

$$ag(x, Y) < \ln(\frac{1}{\omega}).$$

Example 5 *(Absolute value balanced loss).*
The absolute value balanced loss function is:

$$L(r(x), \hat{r}_g(x)) = \omega g(x, Y) + (1-\omega)|r(x) - \hat{r}(x) + g(x, Y)|. \tag{9}$$

For the regression model (1) and under the absolute value balanced loss function (9) with $\frac{1}{2} \leq \omega < 1$, the kernel estimator $\hat{r}(x)$ is admissible.

6. Concluding remarks

1. In this note, we show the impact of the choice of the loss function on the admissibility of the Stein-type estimator of the regression operator. In some cases, the usual estimator is dominated by the Stein estimator, whereas in other choices of the loss function, the situation is similar to the quadratic case, as proved in results obtained by James and Stein in [21].

2. We can without difficulty project the results of this paper on the problem of the parametric regression model. This idea stems from the simple fact that the explanatory variable X in the model (1) is of functional type (taking values in a semi-metric space), so one can have analogue results for the parametric model studied in [10] or any general multivariate linear model.
3. The nonparametric methods used for improving the usual estimators (smoothing methods and cross-validation idea) leads to choose between small bias and large variance (or the contrary) with respect to interesting arguments in the posed problem. But as all can see, Stein's method (shrinkage) leads to minimize the risk (and so the variance) without taking the bias into account.

Acknowledgements

The authors thank the referee for his/her detailed comments and suggestions.

References

[1] K. Benhenni, F. Ferraty, M. Rachdi, P. Vieu, Local smoothing regression with functional data, Comput. Stat. 22 (22) (2007) 353–369.
[2] K. Benhenni, S. Hedli-Griche, M. Rachdi, P. Vieu, Consistency of the regression estimator with functional data under long memory conditions, Stat. Probab. Lett. 78 (8) (2008) 1043–1049, https://doi.org/10.1016/j.spl.2007.11.011.
[3] G. Boente, R. Fraiman, Kernel-based functional principal components, Stat. Probab. Lett. 48 (4) (2000) 335–345.
[4] D. Bosq, Linear Processes in Function Space: Theory and Application, Lecture Notes in Statistics, vol. 149, Springer-Verlag, New York, 2002.
[5] A.C. Brandwein, W.E. Strawderman, Generalizations of James–Stein estimators under spherical symmetry, Ann. Stat. 19 (3) (1991) 1639–1650.
[6] A.C. Brandwein, S. Ralescu, W.E. Strawderman, Shrinkage estimators of the location parameter for certain spherically symmetric distributions, Ann. Inst. Stat. Math. 45 (3) (1991) 551–563.
[7] D. Cellier, D. Fourdrinier, Shrinkage estimators under spherical symmetry for the general linear model, J. Multivar. Anal. 52 (2) (1995) 338–351.
[8] D. Cellier, D. Fourdrinier, C. Robert, Robust shrinkage estimators of the location parameter for elliptically symmetric distributions, J. Multivar. Anal. 29 (1989) 39–52.
[9] A. Cuevas, A partial overview of the theory of statistics with functional data, J. Stat. Plan. Inference 147 (2014) 1–23.
[10] D.K. Dey, M. Ghosh, W.E. Strawderman, On estimation with balanced loss functions, Stat. Probab. Lett. 45 (1999) 97–101.
[11] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis. Theory and Practice, Springer-Verlag, New York, 2006.
[12] D. Fourdrinier, M.T. Wells, Estimation of a loss function for spherically symmetric distributions in the general linear model, Ann. Stat. 23 (2) (1995) 571–592.
[13] D. Fourdrinier, M.T. Wells, On improved loss estimation for shrinkage estimators, Stat. Sci. 27 (1) (2012) 61–81.
[14] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics, J. Multivar. Anal. 146 (2016) 1–6, https://doi.org/10.1016/j.jmva.2015.12.001.
[15] W. James, C. Stein, Estimation with quadratic loss, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1: Contributions to the Theory of Statistics, Statistical Laboratory of the University of California, Berkeley, 20 June–30 July, 1960, University of California Press, Berkeley, CA, USA, 1961, pp. 361–379.
[16] M.J. Jozani, E. Marchand, A. Parsian, On estimation with weighted balanced-type loss function, Stat. Probab. Lett. 76 (2006) 773–780.
[17] M. Rachdi, P. Vieu, Nonparametric regression for functional data: automatic smoothing parameter selection, J. Stat. Plan. Inference 137 (9) (2007) 2784–2801.
[18] J. Ramsay, C. Dalzell, Some tools for functional data analysis (with discussion), J. R. Stat. Soc. B 53 (3) (1991) 539–572.
[19] J. Ramsay, B. Silverman, Functional Data Analysis, Springer Series in Statistics, Springer-Verlag, New York, 1997.
[20] J. Ramsay, B. Silverman, Applied Functional Data Analysis: Methods and Case Studies, Springer Series in Statistics, Springer-Verlag, New York, 2002.
[21] C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in: Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, vol. 1, Statistical Laboratory of the University of California, Berkeley and Los Angeles, December, 1954 and July–August, 1955, University of California Press, 1956, pp. 197–206.
[22] A. Zellner, Bayesian and non-Bayesian estimation using balanced loss functions, in: Statistical Decision Theory and Related Topics V, Springer, New York, 1994, pp. 337–390.