MOX-Report No. 53/2020

Some numerical test on the convergence rates of regression with differential regularization

Arnone, E.; Kneip, A.; Nobile, F.; Sangalli, L. M.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Published in Functional and High-Dimensional Statistics and
Related Fields, Springer, 2020, 11–18
Journal version available at https://doi.org/10.1007/978-3-030-47756-1_3

Some numerical test on the convergence rates of regression with differential regularization

Eleonora Arnone[1], Alois Kneip[2], Fabio Nobile[3], Laura M. Sangalli[1]

[1] MOX– Dipartimento di Matematica, Politecnico di Milano
[2] Universität Bonn
[3] École polytechnique fédérale de Lausanne

Abstract

We numerically study the bias and the mean square error of the estimator in Spatial Regression with Partial Differential Equation (SR-PDE) regularization. SR-PDE is a novel smoothing technique for data distributed over two-dimensional domains, which allows to incorporate prior information formalized in term of a partial differential equation. This technique also enables an accurate estimation when the shape of the domain is complex and it strongly influences the phenomenon under study.

1 Introduction

Spatial functional statistic is a field of research of strong interest in recent years, due to the fact that spatially dependent functional data are increasingly available in many applied fields, such as biology, life science, environmental science and engineering (see [7, 17] for a review on the recent proposed methods).

In this work, we numerically investigate the asymptotic properties of the estimator in Spatial Regression with Partial Differential Equation regularization (SR-PDE) introduced in [18, 20, 3]. SR-PDE is a penalized regression method, that includes the penalty term the misfit from a linear Partial Differential Equation (PDE). This allow a great flexibility of the method. In particular, the PDE in the regularizing term enables the modelling of anisotropy and non-stationarity of the phenomenon under study. Moreover, thanks to the use of the finite element method, it allows to consider domain of complex shape, such as domains with strong concavities, that affect the phenomenon under study, and to impose boundary conditions. Smoothing is a fundamental step in most analyses involving functional data [19, 10, 14]. In this respect, the considered SR-PDE method provides a versatile tool for the smoothing of functional data observed over two-dimensional domains.

Other regularized least-square smoothers have been proposed that can deal with complex domains, such as bivariate splines over triangulations [15, 11, 8, 16], soap film smoothing [24], and low-rank thin-plate spline approximations [23, 21]. All these methods have isotropic regularizing terms. Among the methods mentioned above, the only one that can comply with boundary conditions is soap film smoothing. The asymptotic properties of bivariate splines over triangulations are investigated in [16]. To the best of our knowledge, no results on large sample properties is available for any of the other methods.

The study of the asymptotic properties of classical penalized regression estimators is a well established literature that dates back to the 80s (see, e.g., [9] and references therein). The arguments used to prove the study the bias and the MSE of thin-plate-splines and of smoothing splines [4, 5, 6, 12, 13], however, exploit the existence of an explicit closed form of the Green functions of the differential operator in the regularizing term. Due to the more complex penalty considered by SR-PDE, and moreover, due to the presence of boundary conditions which enable to deal with domains of complex shape, a closed form for the Green functions of the differential operator in the regularizing term is not available for SR-PDE. In addition, as already mentioned, the estimation problem is solved by means of finite elements, with a mixed formulation. This is very convenient from a computational point of view, but makes the analysis of the asymptotic properties much more involved. In [2] a first attempt to study the bias of the infinite dimensional estimator with respect to the smoothing parameter is presented, while the finite element estimator is studied letting the discretization becomes more and more fine, but fixing the number of observations.

In this work, instead, we want to study the asymptotic behavior of the estimator when the number of observations increases to infinity. Next section presents the estimator, while the last section reports some simulation studies that investigate the rates for the bias and the mean square error of the estimator.

2 Spatial Regression with PDE penalization

Let $\Omega \subset \mathbb{R}^2$ a bounded domain, with boundary $\partial\Omega \in C^2$ or polygonal. Consider n observations $z_i \in \mathbb{R}$, for $i = 1, \ldots, n$, located at points $\mathbf{p}_i = (x_i, y_i) \in \Omega$. Assume that:

$$z_i = f_0(\mathbf{p}_i) + \varepsilon_i$$

where $f_0 : \Omega \to \mathbb{R}$ is the field we wish to estimate, and ε_i are independent errors with zero mean and finite variance σ^2.

Denote by $H^2(\Omega)$ the Sobolev space of functions in $L^2(\Omega)$ with derivatives up to the 2-th order in $L^2(\Omega)$, and let V_α the space H^2 with Dirichlet or Neumann boundary conditions, that is

$$V_\alpha = V_\alpha^{\text{dir}} = \{f \in H^2(\Omega) : \; f = \alpha \text{ on } \partial\Omega\}$$

or

$$V_\alpha = V_\alpha^{\text{neu}} = \{f \in H^2(\Omega) : \; \frac{\partial f}{\partial \nu} = \alpha \text{ on } \partial\Omega\}$$

where ν denotes the normal versor to the boundary $\partial\Omega$, and α is the value imposed on the boundary. SR-PDE solves the following estimation problem:

$$\hat{f} = \operatorname*{argmin}_{f \in V_\alpha} \frac{1}{n} \sum_{i=1}^{n} (f(\mathbf{p}_i) - z_i)^2 + \lambda_n \int_\Omega (Lf - u)^2 \qquad (1)$$

where

$$L(\mathbf{p})f = -\mathrm{div}(\mathbf{K}(\mathbf{p})\nabla f) + \mathbf{b}(\mathbf{p}) \cdot \nabla f + c(\mathbf{p})f$$

is a second order linear elliptic operator and the PDE $Lf = u$ partially describes the phenomenon under study. The smoothing parameter $\lambda_n > 0$ controls the relative weight of the two terms in the functional in (1): a data fidelity term, given by the sum of square errors, and a model fidelity term, the differential regularization, defined as the $L^2(\Omega)$-norm of the misfit with respect to the PDE. We explicitly highlight the dependence of the smoothing parameter with respect to n, since as the number of data locations increases less regularization is needed. We thus expect to let λ_n go to zero as n goes to infinity.

The SR-PDE estimator defined in (1) cannot be computed analytically, we thus have to compute an approximated solution. Figure 1 shows the discretization process. We first introduce a triangulation of the domain Ω and then we define a finite element basis over the triangulation. Each finite element basis is a piecewise linear function over the triangulation, which take value one at a node of the triangulation and zero at all the other nodes. We approximate (1) in the finite element space, and in particular we obtain an approximate solution of the problem solving a linear system. For an accurate description of the discretization see [3].

In this work we restrict our attention to the special case in which the finite element basis is linear (i.e. each basis is a piecewise linear function) and the triangulation is such that the vertices of the triangles are in correspondence of the data locations $\mathbf{p}_i$. This is a standard setting in many applications.

3 Numerical study of asymptotic properties

As shown in [22], the best rate of convergence for general penalized regression estimators over a 2-dimensional domain is

$$\text{MSE} \sim n^{-\frac{p}{p+1}}$$

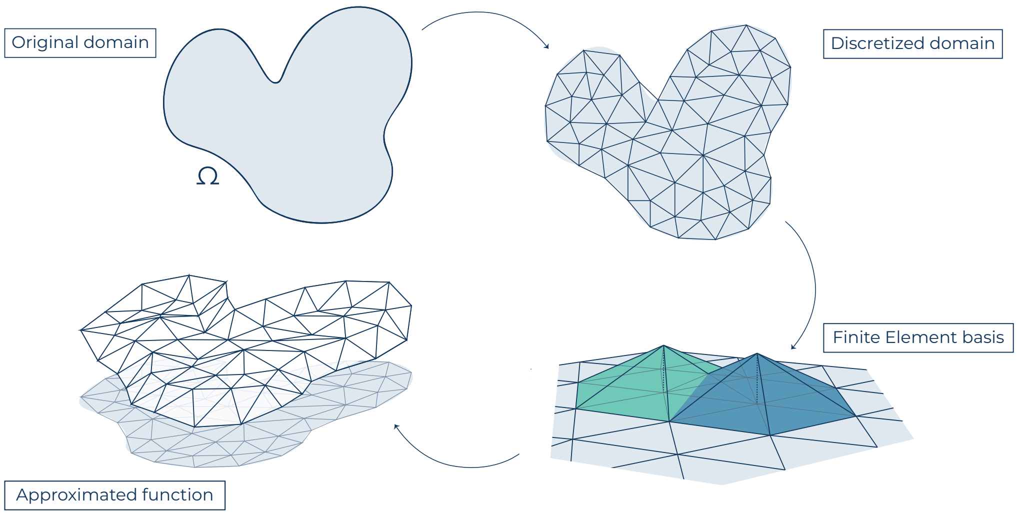

Figure 1: The discretization process. Starting from the original domain (top, left), a polygonal approximation is given and a triangulation is defined (top, right). The linear finite element basis (bottom, right) is introduced over the triangulation and a piecewise linear approximation (bottom, left) of the function of interest is computed.

and is achieved choosing

$$\lambda_n \sim n^{-\frac{p}{2(p+1)}}$$

where p is the number of existing derivatives of the function f_0 that we want to estimate. Since the estimator of SR-PDE is searched in the space $H^2(\Omega)$, in our simulations we set $p = 2$ and let λ_n decrease as $n^{-1/3}$. We thus expect to observe a rate of convergence for the bias of the estimator of order $n^{-1/3}$ and for the MSE of order $n^{-2/3}$.

We consider four different simulation settings that are characterized by different boundary conditions (b.c.): Dirichlet exact b.c., Dirichlet wrong b.c., Neumann exact b.c. and Neumann wrong b.c.. In this way, we can also explore the effect of different boundary conditions on the rate of decay of the error. Exact b.c. corresponds to a complete knowledge of α, that is of the phenomenon at the boundary, while wrong b.c. corresponds to no-knowledge of the behavior at the boundary. The error is computed in the discrete norm on the data locations. We use the same spatial domain and the same test function considered in the first chapter of [1], where the convergence is studied in the case of exact Dirichlet boundary conditions.

Figures 2 and 3 show the bias of the SR-PDE estimator with respect to the number of observations n in case of Dirichlet and Neumann boundary conditions respectively. To compute the bias the method is applied to the exact data, without adding any noise at the evaluations. We can observe that both in the Dirichlet and Neumann case the expected rate of convergence is achieved in case

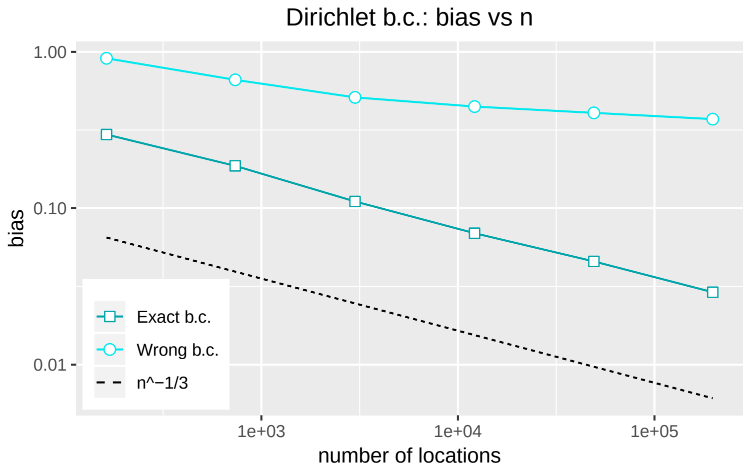

Figure 2: Test functions without noise; exact and wrong Dirichlet boundary conditions. Convergence rates of the bias of the finite element estimator with respect to the number of observations n, with $\lambda_n = n^{-2/3}$.

of exact boundary conditions. The rate on decay of the bias is strongly influenced by wrong Dirichlet boundary conditions, as we can observe from Figure 2 the error is practically non decreasing for large values of n. Wrong Neumann boundary conditions still affect the rate of decay of the bias, however, as we can observe from Figure 3, the bias is still decreasing for large values of n.

Figures 4 and 5 show the MSE of the SR-PDE estimator with respect to the number of observations n in case of Dirichlet and Neumann boundary conditions respectively. To compute the MSE a gaussian uncorrelated noise is added to the exact data. As for the bias, we observe that wrong Dirichlet boundary conditions strongly affect the performance of the estimator. The expected rate is achieved in the exact Dirichlet and in the wrong Neumann case. In the exact Neumann case the rate of convergence seems to be faster than expected, this may be due to the fact that, even if the estimator is searched in the space $H^2(\Omega)$, the true f_0 has more than two derivatives.

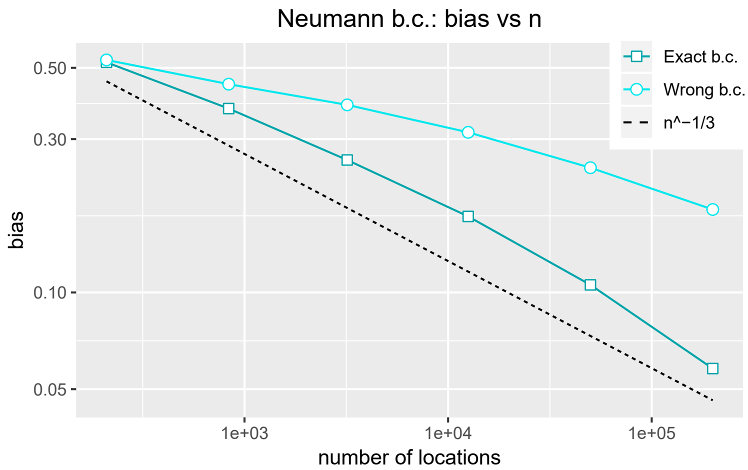

Figure 3: Test functions without noise; exact and wrong Neumann boundary conditions. Convergence rates of the bias of the finite element estimator with respect to the number of observations n, with $\lambda_n = n^{-2/3}$.

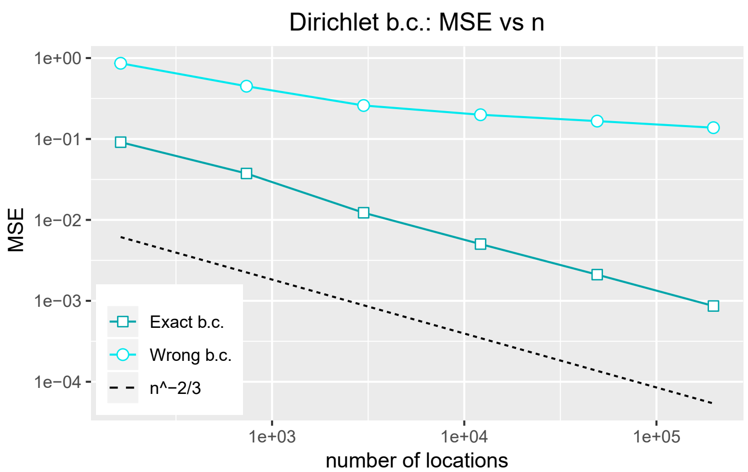

Figure 4: Data with noise; exact and wrong Dirichlet boundary conditions. Convergence rates of the MSE of the finite element estimator with respect to the number of observations n, with $\lambda_n = n^{-2/3}$.

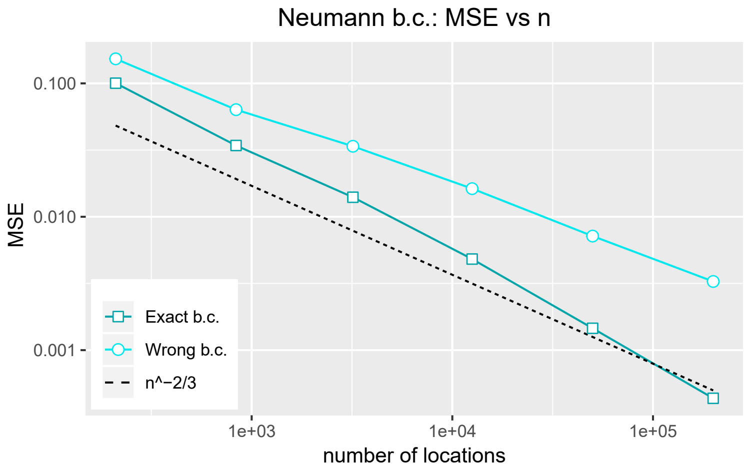

Figure 5: Data with noise; exact and wrong Neumann boundary conditions. Convergence rates of the MSE of the finite element estimator with respect to the number of observations n, with $\lambda_n = n^{-2/3}$.

4 Future directions

We have numerically investigated the rate of decay of the bias and the MSE of the SR-PDE estimator, showing that the optimal rate of convergence can be achieved when Dirichlet or Neumann exact boundary conditions are enforced. We also have shown that wrong Neumann boundary conditions affect the rate of decay of the error, that however continue to decay for large values of n. The empirical results displayed in this work support the consistency of SR-PDE estimator. We are currently working on proving the consistency theoretically.

We have here considered a standard choice of the discretization of the domain, with a finite element basis for each data location. However, the SR-PDE does not impose this restriction. An interesting future development is the study of the rate of convergence when the finite element basis is not constrained to the data locations, in order to have a finer or coarser triangulation of the domain, that may not directly be linked to the number of observations.

References

[1] Arnone, E.: Regression with PDE penalization for modelling functional data with spatial and spatio-temporal dependence. Doctoral dissertation (2018)

[2] Azzimonti L., Nobile F., Sangalli L. M., Secchi P.: Mixed Finite Elements for Spatial Regression with PDE Penalization. SIAM/ASA Journal on Uncertainty Quantification **2(1)**, 305–335 (2014)

[3] Azzimonti L., Sangalli L. M., Secchi P., Domanin M., Nobile F.: Blood flow velocity field estimation via spatial regression with PDE penalization. Journal of the American Statistical Association **110(511)**, 1057–1071 (2015)

[4] Cox, D. D.: Asymptotics for M-type smoothing splines. The Annals of Statistics, 530–551 (1983)

[5] Cox, D. D.: Multivariate smoothing spline functions. SIAM Journal on Numerical Analysis, **21(4)**, 789–813 (1984)

[6] Cucker, F., Zhou, D. X: Learning theory: an approximation theory viewpoint (Vol. 24). Cambridge University Press. (2007)

[7] Delicado, P., Giraldo, R., Comas, C., Mateu, J.: Statistics for spatial functional data: some recent contributions. Environmetrics, **21(3-4)**, 224–239 (2010)

[8] Ettinger, B., Guillas, S., Lai, M. J.: Bivariate splines for ozone concentration forecasting. Environmetrics, **23(4)**, 317–328 (2012)

[9] Eubank, R. L.: Nonparametric regression and spline smoothing. CRC press. (1999)

[10] Ferraty, F., Vieu, P.: Nonparametric functional data analysis: theory and practice. Springer Science & Business Media. (2006)

[11] Guillas, S., Lai, M. J.: Bivariate splines for spatial functional regression models. Journal of Nonparametric Statistics, **22(4)**, 477–497 (2010)

[12] Györfi, L., Kohler, M., Krzyzak, A., Walk, H.: A distribution-free theory of nonparametric regression. Springer Science & Business Media. (2006)

[13] Huang, J. Z: Local asymptotics for polynomial spline regression. The Annals of Statistics, **31(5)**, 1600–1635 (2003)

[14] Kokoszka, P., Reimherr, M.: Introduction to functional data analysis. Chapman and Hall/CRC. (2017)

[15] Lai, M. J., Schumaker, L. L.: Spline functions on triangulations (No. 110). Cambridge University Press. (2007)

[16] Lai, M. J., Wang, L.: Bivariate penalized splines for regression. Statistica Sinica, **23(3)**, 1399–1417 (2013)

[17] Mateu, J., Romano, E.: Advances in spatial functional statistics. Stoch Environ Res Risk Assess **31** 1–6 (2017)

[18] Ramsay, T.: Spline smoothing over difficult regions. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **54(2)**, 307–319 (2002)

[19] Ramsay, J., Silverman, B.: Functional Data Analysis. (Second edition) Springer, New York (2005)

[20] Sangalli, L. M. and Ramsay, J. O. and Ramsay, T. O.: Spatial spline regression models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **75(4)**, 681–703 (2013)

[21] Scott-Hayward, L. A. S., Mackenzie, M. L., Donovan, C. R., Walker, C. G., Ashe, E.: Complex region spatial smoother (CReSS). Journal of Computational and Graphical Statistics, **23(2)**, 340–360 (2014)

[22] Stone, C. J.: Optimal global rates of convergence for nonparametric regression. The annals of statistics, 1040–1053 (1982)

[23] Wang, H., Ranalli, M. G.: Low-rank smoothing splines on complicated domains. Biometrics, **63(1)**, 209–217 (2007)

[24] Wood, S. N., Bravington, M. V., Hedley, S. L.: Soap film smoothing. Journal of the Royal Statistical Society: Series B (Statistical Methodology), **70(5)**, 931–955 (2008)

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

49/2020 Bonaventura,L.; Garres Diaz,J.
Flexible and efficient discretizations of multilayer models with variable density

50/2020 Bonaventura,L.; Gomez Marmol, M.
The TR-BDF2 method for second order problems in structural mechanics

51/2020 Ferraccioli, F.; Sangalli, L. M.; Arnone, E.,; Finos, L.
A functional data analysis approach to the estimation of densities over complex regions

52/2020 Arnone, E.; Kneip, A.; Nobile, F.; Sangalli, L. M.
Some first results on the consistency of spatial regression with partial differential equation regularization

48/2020 Clerici, F.; Ferro, N.; Marconi, S.; Micheletti, S.; Negrello, E.; Perotto, S.
Anisotropic adapted meshes for image segmentation: application to 3D medical data

46/2020 Bucelli, M.; Salvador, M.; Dede', L.; Quarteroni, A.
Multipatch Isogeometric Analysis for Electrophysiology: Simulation in a Human Heart

45/2020 Gatti, F.; Menafoglio, A.; Togni, N.; Bonaventura, L.; Brambilla, D.; Papini, M; Longoni, L.
A novel dowscaling procedure for compositional data in the Aitchison geometry with application to soil texture data

44/2020 Masci, C.; Ieva, F.; Paganoni A.M.
EM algorithm for semiparametric multinomial mixed-effects models

47/2020 Sangalli, L.M.
A novel approach to the analysis of spatial and functional data over complex domains

42/2020 Miglio, E.; Parolini, N.; Quarteroni, A.; Verani, M.; Zonca, S.
A spatio-temporal model with multi-city mobility for COVID-19 epidemic