

Functional data analysis in spaces of surfaces

Analisi di dati funzionali in spazi di superfici

Laura M. Sangalli

Abstract The talk presents a novel functional data analysis technique for surface estimation and spatial smoothing, at the interface between statistics and numerical analysis.

Abstract *Il seminario presenta una nuova tecnica di analisi di dati funzionali per la stima accurata di superfici e campi spaziali, all'interfaccia tra statistica ed analisi numerica.*

Key words: penalized regression, partial differential equations, finite elements.

1 Surface estimation and spatial smoothing via regression models with partial differential regularizations

The talk presents a novel functional data analysis technique for accurate surface estimation and spatial smoothing. The proposed class of models are penalized regression models with regularizing terms involving partial differential operators. In simpler context of curve estimation and univariate smoothing problems, the idea of regularization with ordinary differential operators has already proved to be very effective and is in general playing a central role in the functional data analysis literature. See, e.g., [12]. Also in the more complex case of surface estimation and spatial smoothing, some methods use roughness penalties involving simple forms of partial differential operators. A classical example is given by thin-plate-splines, while more recent proposals are offered for instance by [13, 16, 8]; see also the applications in [9, 6, 1]. Finally, although in a different framework, the use of simple form of (stochastic) PDEs is also at the core of the Bayesian spatial models introduced by

Laura M. Sangalli

MOX - Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano
e-mail: laura.sangalli@polimi.it

[10] and more generally the larger literature on Bayesian inverse problems [15] and data assimilation in inverse problems [5].

Regression models with partial differential regularizations merge advanced statistical methodology with numerical analysis techniques. Thanks to the combination of potentialities from these two scientific areas, the proposed class of models have important advantages with respect to classical statistical techniques for bidimensional smoothing and for surfaces and spatial fields estimation. Regression models with partial differential regularizations are able to efficiently deal with data distributed over irregularly shaped domains, with complex boundaries, strong concavities and interior holes [14]. Moreover, they can comply with specific conditions at the boundaries of the problem domain [14, 3], which is fundamental in many applications to obtain meaningful estimates. The proposed models can also deal with data scattered over Riemannian manifold domains [7], only few methods existing in literature for this type of data structures. Moreover, regression models with partial differential regularizations have the capacity to incorporate problem-specific priori information about the spatial structure of the phenomenon under study [4, 3, 2], with a very flexible modeling that allows naturally for anisotropy and non-stationarity. Space-varying covariate information is also included in the models via a semiparametric framework. The estimators have a penalized regression form, they are linear in the observed data values, and have good inferential properties. The use of advanced numerical analysis techniques, and specifically of finite elements (see, e.g., [11]), makes the models computationally very efficient.

During the talk the method will be illustrated in various applied contexts, including demographic data and medical imaging data.

Acknowledgements The talk is based on joint work with Laura Azzimonti, Bree Ettinger, Fabio Nobile, Simona Perotto, Jim Ramsay, Piercesare Secchi, Matthieu Wilhelm. This research has been funded by the research program Dote Ricercatore Politecnico di Milano - Regione Lombardia, project “Functional data analysis for life sciences”, and by the starting grant *FIRB Futuro in Ricerca*, MIUR Ministero dell’Istruzione dell’Università e della Ricerca, research project “Advanced statistical and numerical methods for the analysis of high dimensional functional data in life sciences and engineering” (<http://mox.polimi.it/users/sangalli/firbSNAPLE.html>).

References

1. Augustin, N.H., Trenkel, V.M., Wood, S.N. and Lorance, P.: Space-time modelling of blue ling for fisheries stock management Issue. *Environmetrics*, Vol. 24, Part 2, pp. 109–119 (2013)
2. Azzimonti L.: Blood flow velocity field estimation via spatial regression with PDE penalization, PhD Thesis, Politecnico di Milano (2013)
3. Azzimonti, L., Nobile, F., Sangalli, L. M., Secchi, P.: Mixed Finite Elements for spatial regression with PDE penalization. TechRep. MOX 20/2013, Dipartimento di Matematica, Politecnico di Milano (2013)
4. Azzimonti, L., Sangalli, L. M., Secchi, P., Domanin, M., Nobile, F.: Blood flow velocity field estimation via spatial regression with PDE penalization. TechRep. MOX 19/2013, Dipartimento di Matematica, Politecnico di Milano (2013)

5. D'Elia, M., Perego, M., and Veneziani, A.: A Variational Data Assimilation Procedure for the Incompressible Navier-Stokes Equations in Hemodynamics. *SIAM Journal of Scientific Computing*, Vol. 52, Part 4, pp. 340–359. 2012, Volume 52, Issue 2, pp 340-359 (2012)
6. Ettinger, B., Guillas, S. and Lai, M.-J.: Bivariate splines for ozone concentration forecasting. *Environmetrics*, Vol. 23, Part 4, pp. 317–328 (2012)
7. Ettinger, B., Perotto, S., Sangalli, L. M. (2012). Spatial regression models over two-dimensional manifolds. Tech. Rep. MOX 54/2012, Dipartimento di Matematica, Politecnico di Milano.
8. Guillas, S. and Lai, M.: Bivariate splines for spatial functional regression models. *Journal of Nonparametric Statistics*, Vol. 22, Part 4, pp. 477–497 (2010)
9. Marra, G., Miller, D. and Zanin, L.: Modelling the spatiotemporal distribution of the incidence of resident foreign population. *Statistica Neerlandica*, Vol. 66, pp. 133–160 (2012)
10. Lindgren, F., Rue, H., and Lindström, J.: An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Ser. B, Statistical Methodology*, Vol. 73, pp. 423–498, with discussions and a reply by the authors (2011)
11. Quarteroni, A. Numerical Models for Differential Problems. *Springer* (2013)
12. Ramsay, J. O. and Silverman, B. W.: Functional Data Analysis, 2nd edn. *Springer* (2005)
13. Ramsay, T.: Spline Smoothing over Difficult Regions. *Journal of the Royal Statistical Society Ser. B, Statistical Methodology*, Vol. 64, pp. 307–319 (2002)
14. Sangalli, L. M., Ramsay, J. O., Ramsay, T. O.: Spatial spline regression models. *Journal of the Royal Statistical Society Ser. B, Statistical Methodology*, Vol. 75, Part 4, pp. 681–703 (2013)
15. Stuart, A.: Inverse problems: a Bayesian perspective. *Acta Numerica*, Vol. 19, pp. 451–559 (2010)
16. Wood, S. N., Bravington, M. V., and Hedley, S. L.: Soap film smoothing. *Journal of the Royal Statistical Society Ser. B, Statistical Methodology*, Vol. 70, pp. 931–955 (2008)