Spatio-temporal spline regression models

Mara Bernardi, Gabriele Mazza, Laura Sangalli

We extend the spatial spline regression model described in ([3]) to time dependent data and we implement an algorithm to compute the solution to the estimation problem.

Capitolo 1

Descrizione del modello

1.1 Caso senza covariate

Dati e modello

Siano $\{\underline{p}_i = (x_i, y_i); i = 1, ..., n\}$ un insieme di n punti spaziali in un dominio limitato $\Omega \subset \mathbb{R}^2$ e siano $\{t_j; j = 1, ..., m\}$ un insieme di m istanti temporali in un intervallo $[0, T] \subset \mathbb{R}$. In questi punti ed istanti osserviamo i dati: siano quindi z_{ij} i valori della variabile reale nel punto $\underline{p}_i$ al tempo t_j.

Supponiamo che le osservazioni $\{z_{ij}; i = 1, ..., n; j = 1, ..., m\}$ provengano da una funzione $f(\underline{p}, t)$, con l'aggiunta di un rumore:

$$z_{ij} = f(\underline{p}_i, t_j) + \epsilon_{ij} \quad i = 1, ..., n \ \ j = 1, ..., m \ \ , \tag{1.1}$$

dove $\{\epsilon_{ij}; i = 1, ..., n; j = 1, ...m\}$ sono residui indipendenti identicamente distribuiti di media nulla e varianza σ^2.

L'obiettivo del modello è stimare la funzione spazio-temporale $f(\underline{p}, t)$ dai dati, minimizzando il seguente funzionale:

$$J_{\underline{\lambda}}(f) = \sum_{i=1}^{n} \sum_{j=1}^{m} \big(z_{ij} - f(\underline{p}_i, t_j)\big)^2 \ +$$

$$+ \ \lambda_S \int_0^T \int_\Omega \big(\triangle f\big)^2 d\Omega \ dt \ + \ \lambda_T \int_\Omega \int_0^T \Big(\frac{\partial^2 f}{\partial t^2}\Big)^2 dt \ d\Omega. \tag{1.2}$$

Il primo termine di $J_{\underline{\lambda}}(f)$ considera la minimizzazione dello scarto quadratico tra i dati e la funzione f calcolata nei corrispondenti punti spaziali e istanti temporali. Tuttavia in aggiunta il funzionale presenta due termini di penalizzazione, che nel processo di minimizzazione cercano di rendere liscia e regolare la funzione rispettivamente in spazio e tempo.

Il termine della penalizzazione in spazio comprende l'integrale sull'intervallo $[0, T]$ dell'integrale sul dominio spaziale Ω del quadrato del laplaciano della funzione f. Come è noto, il laplaciano esprime la curvatura della

funzione, ed è quindi una misura di quanto la funzione è liscia in spazio. L'analogo significato in tempo è rappresentato dalla derivata seconda in t, che nell'ultimo termine della penalizzazione è integrata prima sull'intervallo temporale $[0, T]$, e in seguito sul dominio spaziale Ω.

I due termini λ_S e λ_T sono rispettivamente i pesi della penalizzazione in spazio e in tempo. La scelta di $\underline{\lambda}$, vettore (λ_S, λ_T) deve essere molto accurata. Infatti, valori troppo bassi per i due termini causerebbero una stima vicina all'interpolazione dei dati (poiché darebbero più peso al termine con gli scarti quadratici), mentre valori troppo elevati porterebbero ad avere una funzione f fin troppo liscia e quindi distante dai dati. Per questo motivo sarà dato ampio spazio alla scelta di $\underline{\lambda}$.

DESCRIZIONE DI DOVE VIVE LA FUNZIONE F

Riduzione finito-dimensionale del problema

Per poter risolvere il problema numericamente è necessaria una riduzione finito-dimensionale della funzione f. Rappresentiamo quindi f con un opportuno sviluppo di basi separate in spazio e tempo.

Siano $\{\varphi_i(\underline{p}); i = 1, ..., N\}$ un insieme di N basi spaziali definite sul dominio Ω. Non è necessario che le basi siano esattamente n, è possibile averne anche un numero maggiore be a set of n spatial basis functions and $\{\psi_j(t); j = 1, ..., m\}$ be a set of m temporal basis functions. The spatio-temporal field $f(\underline{p}, t)$ we want to estimate can be written as follows:

$$f(\underline{p}, t) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \ \varphi_i(\underline{p}) \ \psi_j(t). \tag{1.3}$$

We can then rewrite the functional $J_\lambda(f)$ in a matricial form.
We define the vector $\underline{z}$ of observed values and the vector $\underline{c}$ containing the coefficients of the basis expansion of the spatio-temporal field f:

$$\underline{z} = \begin{bmatrix} z_{11} \\ \vdots \\ z_{1m} \\ z_{21} \\ \vdots \\ z_{2m} \\ \vdots \\ z_{nm} \end{bmatrix} \qquad \underline{c} = \begin{bmatrix} c_{11} \\ \vdots \\ c_{1m} \\ c_{21} \\ \vdots \\ c_{2m} \\ \vdots \\ c_{nm} \end{bmatrix}. \tag{1.4}$$

We define the matrix Φ of the evaluations of the spatial basis functions in the spatial points $\{\underline{p}_i; i = 1, ..., n\}$ and the matrix Ψ of the evaluations of

the temporal basis functions in the temporal instants $\{t_j; j = 1, ..., m\}$:

$$\Phi = \begin{bmatrix} \varphi_1(\underline{p}_1) & \varphi_2(\underline{p}_1) & \cdots & \varphi_n(\underline{p}_1) \\ \varphi_1(\underline{p}_2) & \varphi_2(\underline{p}_2) & \cdots & \varphi_n(\underline{p}_2) \\ \vdots & \vdots & \cdots & \vdots \\ \varphi_1(\underline{p}_n) & \varphi_2(\underline{p}_n) & \cdots & \varphi_n(\underline{p}_n) \end{bmatrix}$$

$$\Psi = \begin{bmatrix} \psi_1(t_1) & \psi_2(t_1) & \cdots & \psi_m(t_1) \\ \psi_1(t_2) & \psi_2(t_2) & \cdots & \psi_m(t_2) \\ \vdots & \vdots & \cdots & \vdots \\ \psi_1(t_n) & \psi_2(t_n) & \cdots & \psi_m(t_m) \end{bmatrix}$$

Let Π be the Kronecker product of Φ and Ψ:

$$\Pi = \Phi \otimes \Psi.$$

Finally, we can write the functional J in (1.2) in the following way:

$$J = (\underline{z} - \Pi\underline{c})^T(\underline{z} - \Pi\underline{c}) + \underline{c}^t S \underline{c} \tag{1.5}$$

where S is the matrix containing the penalty term. We obtain the matrix S by discretizing the penalty term of J in (1.2) in the following way:

$$\lambda_S \sum_{j=1}^{m} \int_{\Omega_T} \left(\triangle f(\underline{p}, t_j) \right)^2 d\underline{p} + \lambda_T \sum_{i=1}^{n} \int_0^T \left(\frac{\partial^2 f(\underline{p}_i, t)}{\partial t^2} \right)^2 dt, \tag{1.6}$$

where Ω_T is a triangulation of the domain Ω.
This last expression can in turn be written as follows:

$$\lambda_S \, \underline{c}^T (I_m \otimes S_S) \, \underline{c} \; + \; \lambda_T \, \underline{c}^T (S_T \otimes I_n) \, \underline{c},$$

where I_n and I_m are the identity matrices of dimension $n \times n$ and $m \times m$ respectively and S_S and S_T are the discretizations of the integrals in (1.6). This leads to the definition of S as:

$$S = \lambda_S \, (S_S \otimes I_m) \; + \; \lambda_T \, (I_n \otimes S_T).$$

The expressions for matrices S_S and S_T are obtained substituting the basis expansion of f (1.3) in equation (1.6). The first term in equation (1.6) leads to the computation of the following expression for the (k,l)-th element of the matrix S_S:

$$S_{S_{k,l}} = \int_\Omega \triangle\varphi_k \triangle\varphi_l \quad k = 1, ..., n \;\; l = 1, ..., n.$$

Defining $g = \triangle\varphi_l$, we obtain

$$S_{S_{k,l}} = \int_\Omega \triangle\varphi_k g \; , \qquad \int_\Omega gv = \int_\Omega \triangle\varphi_l v.$$

Using the Green formula and exploiting the Neumann boundary conditions we can write:

$$S_{S_{k,l}} = -\int_\Omega \nabla \varphi_k \nabla g \ , \qquad \int_\Omega gv = -\int_\Omega \nabla \varphi_l \nabla v.$$

We define the vector of basis functions and the vectors of partial derivatives of the basis functions

$$\underline{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix}$$

$$\underline{\varphi}_x = \begin{bmatrix} \partial\varphi_1/\partial x \\ \partial\varphi_2/\partial x \\ \vdots \\ \partial\varphi_n/\partial x \end{bmatrix} \qquad \underline{\varphi}_x = \begin{bmatrix} \partial\varphi_1/\partial y \\ \partial\varphi_2/\partial y \\ \vdots \\ \partial\varphi_n/\partial y \end{bmatrix} \tag{1.7}$$

and the matrices

$$R_0 = \int_\Omega \underline{\varphi}\underline{\varphi}^T$$

$$R_1 = \int_\Omega (\underline{\varphi}_x\underline{\varphi}_x^T + \underline{\varphi}_y\underline{\varphi}_y^T).$$

Letting v be a basis element and $\underline{g}$ be the coefficients of the expansion of the function g on the basis, we obtain

$$S_S = -R_1\underline{g} \ , \qquad R_0\underline{g} = -R_1$$

and finally

$$S_S = R_1 R_0^{-1} R_1.$$

The second term in equation (1.6) leads to the following definition of matrix S_T:

$$S_T = \begin{bmatrix} \int_0^T \psi_1''(t)\psi_1''(t)dt & \int_0^T \psi_1''(t)\psi_2''(t)dt & \dots & \int_0^T \psi_1''(t)\psi_m''(t)dt \\ \int_0^T \psi_2''(t)\psi_1''(t)dt & \int_0^T \psi_2''(t)\psi_2''(t)dt & \dots & \int_0^T \psi_2''(t)\psi_m''(t)dt \\ \vdots & \vdots & \dots & \vdots \\ \int_0^T \psi_m''(t)\psi_1''(t)dt & \int_0^T \psi_m''(t)\psi_2''(t)dt & \dots & \int_0^T \psi_m''(t)\psi_m''(t)dt \end{bmatrix}.$$

In order to obtain an estimate of the coefficients vector $\underline{c}$, we have to minimize the functional J in (1.5). Imposing that the first derivative of J equals zero, we obtain the following estimate for $\underline{c}$:

$$\hat{\underline{c}} = (\Pi^T\Pi + S)^{-1}\Pi^T\underline{z}$$

Choosing the smoothing parameters

The values of the smoothing parameters λ_S and λ_T can be chosen via the minimization of the generalized cross-validation (GCV):

$$GCV(\underline{\lambda}) = \frac{nm}{nm - \text{tr}(H)} D(\hat{\underline{c}})$$

where $\underline{\lambda}$ is the vector (λ_S, λ_T), nm is the number of data, H is the hat matrix that maps the vector of observed values $\underline{z}$ to the vector of fitted values $\hat{\underline{z}}$:

$$H = \Pi(\Pi^T \Pi + S)^{-1} \Pi^T,$$

and D is the deviance of the model:

$$D(\hat{\underline{c}}) = (\underline{z} - \hat{\underline{z}})^T (\underline{z} - \hat{\underline{z}}) = (\underline{z} - H\hat{\underline{c}})^T (\underline{z} - H\hat{\underline{c}}).$$

Model with covariates

The model described above can be easily extended to the case when covariates are available. In this case, the model (1.1) becomes:

$$z_{ij} = \underline{w}_{ij}^T \underline{\beta} + f(\underline{p}_i, t_j) + \epsilon_{ij} \quad i = 1, ..., n \;\; j = 1, ..., m \;,$$

where $\underline{w}_{ij}$ is a vector of p covariates associated to the observation z_{ij} and $\underline{\beta}$ is a vector of regression coefficients. Consequently, the functional J in equation (1.5) becomes:

$$J = (\underline{z} - W\underline{\beta} - \Pi\underline{c})^T (\underline{z} - W\underline{\beta} - \Pi\underline{c}) + \underline{c}^t S \underline{c} \;,$$

where W is the $nm \times p$ matrix containing the vectors $\{\underline{w}_{ij}; i = 1, ..., n; j = 1, ..., m\}$. We want to estimate the vector of regression coefficients $\underline{\beta}$ and the vector $\underline{c}$ of coefficients of the basis expansion of the spatio-temporal field f. Therefore, we compute the first partial derivatives of the functional J with respect to $\underline{\beta}$ and $\underline{c}$:

$$\frac{\partial}{\partial \underline{\beta}} J = -2W^T \underline{z} + 2W^T \Pi \underline{c} + 2W^T W \underline{\beta} \;,$$

$$\frac{\partial}{\partial \underline{c}} J = -2\Pi^T \underline{z} + 2\Pi^T W \underline{\beta} + 2(\Pi^T \Pi + S)\underline{c} \;.$$

Imposing that the two expressions above equals zero we obtain the two following equations:

$$\begin{cases} W^T W \hat{\underline{\beta}} = W^T (\underline{z} - \Pi\hat{\underline{c}}) \\ (\Pi^T \Pi + S)\hat{\underline{c}} = \Pi^T (\underline{z} - W\hat{\underline{\beta}}) \end{cases} .$$

Solving this system of two equations leads to explicit solution of the estimation problem.

Application to Venice waste data

We apply the method described to a dataset of measurements of the annual amount of waste produced per capita in 720 cities in the province of Venice. The measurements cover the period from 1997 to 2011.

For the spatial basis functions, we use linear finite elements on a triangulation of the domain. For the temporal basis functions, we use spline functions of order three.

The analysis has been performed with the software R.

The following figures shows the estimated spatio-temporal field at a fixed time instant and at three spatial points.
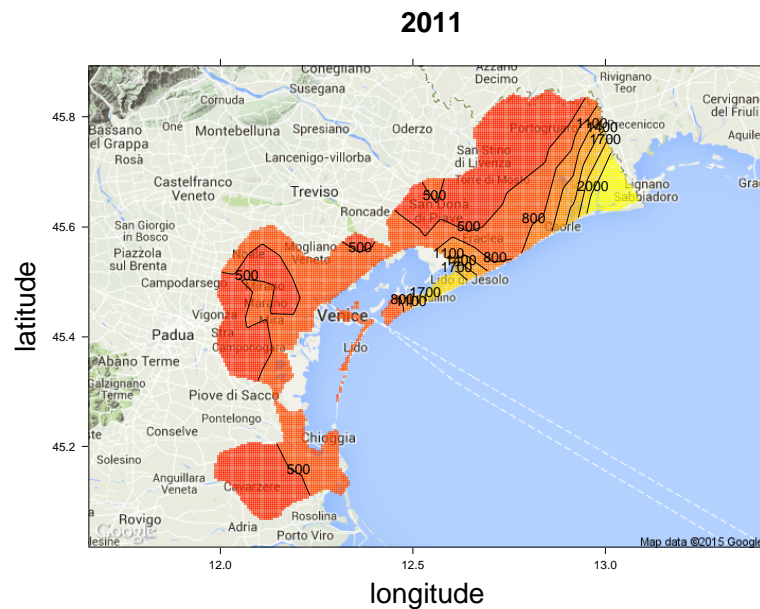

Figura 1.1: Estimated spatio-temporal field of the Venice waste data at a fixed time instant: year 2011.
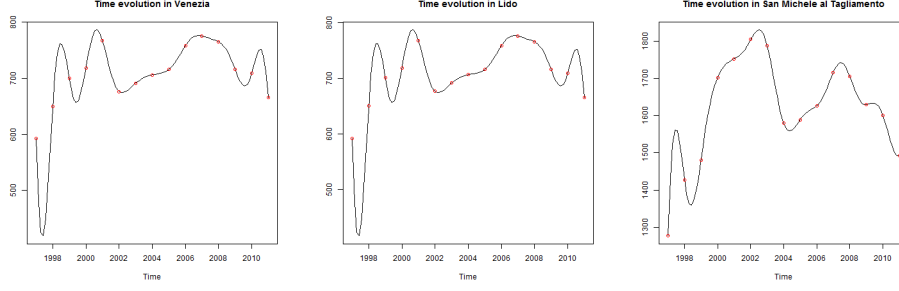
Figura 1.2: Estimated spatio-temporal field of the Venice waste data at three fixed spatial points: Venice, Lido and San Michele al Tagliamento.

Comparison with other spatio-temporal prediction methods

We compare our method with two other approaches to spatio-temporal fields estimation. The first method consists of a generalized additive mixed model with a tensor product of two smoothers in space and time; this algorithm is presented by Marra et al. in ([2]) and by Augustin et al. in ([1]). The second method is spatio-temporal kriging with a separable variogram marginally exponential both in time and space, whose parameters are estimated from the empirical variogram.

We apply the three methods to simulated data on a C-shaped domain. We use the spatial test function f that was used in ([4]) and we construct a spatio-temporal test function g in the following way:

$$g(\underline{p}, t) = f(\underline{p}) \cos(t).$$

Bibliografia

[1] Nicole H. Augustin, Verena M. Trenkel, Simon N. Wood, Pascal Lorance, *Space-time modelling of blue ling for fisheries stock management*, Environmetrics, 24, 109–119, (2013)

[2] Giampiero Marra, David L. Miller, Luca Zanin, *Modelling the spatio-temporal distribution of the incidence of resident foreign population*, Statistica Neerlandica, 66, 133–160, (2012)

[3] Laura M. Sangalli, James O. Ramsay, Timothy O. Ramsay, *Spatial spline regression models*, Journal of the Royal Statistical Society: Series B, 75, 681–703, (2013)

[4] Simon N. Wood, Mark W. Bravington, Sharon L. Hedley, *Soap film smoothing*, Journal of the Royal Statistical Society: Series B, 70, 931–955, (2008)