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:

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

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:

$$\begin{aligned} J_{\underline{\lambda}}(f) = \sum_{i=1}^{n} \sum_{j=1}^{m} \big(z_{ij} - f(\underline{p}_i, t_j)\big)^2 \; + \\ + \; \lambda_S \int_0^T \int_\Omega \Big(\triangle f\Big)^2 d\Omega \; dt \; + \; \lambda_T \int_\Omega \int_0^T \Big(\frac{\partial^2 f}{\partial t^2}\Big)^2 dt \; d\Omega. \end{aligned} \tag{1.2}$$

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 $\underline{\lambda}$.

Spazio funzionale per f

Sviluppo in funzioni di base della funzione f

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.

Sia $\{\varphi_i(\underline{p}); i = 1, ..., N\}$ un insieme di N basi spaziali definite sul dominio Ω e $\{\psi_j(t); j = 1, ..., M\}$ un insieme di basi temporali definite sull'intervallo $[0, T]$. Da questi due insiemi di basi marginali in spazio e tempo si costruisce la funzione f, con tutti i prodotti incrociati disponibili tra spazio e tempo:

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

Nel codice implementato sono disponibili elementi finiti in spazio e splines in tempo, perciò il dominio Ω sarà sostituito dalla sua triangolazione Ω_T.

La funzione $f(\underline{p}, t)$ può essere identificata semplicemente con i valori dei coefficienti $\{c_{ij}; i = 1, ..., N \ j = 1, ..., M\}$. Quindi l'obiettivo della stima sarà il vettore contenente questi coefficienti:

$$\underline{c} = \begin{bmatrix} c_{11} \\ \vdots \\ c_{1M} \\ c_{21} \\ \vdots \\ c_{2M} \\ \vdots \\ c_{NM} \end{bmatrix}. \tag{1.4}$$

e si dimostrerà che sarà soluzione di un sistema lineare.

Discretizzazione dei termini di penalizzazione di J

Dopo aver fissato le basi dello sviluppo di f, è necessario riscrivere in forma discreta anche il funzionale J in (1.2). La parte più complessa da trattare di J è rappresentata dai due termini di penalizzazione in spazio e tempo, che si semplifica considerando la seguente discretizzazione:

$$\lambda_S \sum_{j=1}^{M} \int_{\Omega_T} \Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i)\Big)^2 d\Omega + \lambda_T \sum_{i=1}^{N} \int_0^T \Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2}\Big)^2 dt, \qquad (1.5)$$

La scelta di questa discretizzazione è giustificata dalla separabilità delle due penalizzazioni in J e dalla forma di f nell'espressione in funzioni di base in 1.3. Infatti f è ricavata combinando le basi dei modelli marginali in spazio e in tempo, e si può ricavare che:

$$f(\underline{p}, t) = \sum_{j=1}^{M} m_{S_j}(\underline{p})\psi(t) = \sum_{i=0}^{N} \varphi_i(\underline{p}) m_{T_i}(t)$$

dove

$$m_{S_j}(\underline{p}) = \sum_{i=0}^{N} c_{ij}\varphi_i(\underline{p}) \qquad \forall j = 1...M$$

$$m_{T_i}(t) = \sum_{j=0}^{M} cij\psi_j(t) \qquad \forall i = 1...N$$

cioè da f si possono ricavare due insiemi di funzioni marginali fissando rispettivamente l'indice in tempo o in spazio.

Inoltre lo smoothing marginale spaziale ha il suo termine di penalizzazione (l'integrale su Ω_T del laplaciano, che indicheremo con J_S) così come lo smoothing in tempo (l'integrale su $[0, T]$ della derivata seconda, che indicheremo con J_T). Tali modelli marginali sono già stati trattati con le stesse basi dei casi che saranno analizzati in seguito (elementi finiti in spazio e B-spline in tempo) e sono validi, quindi per discretizzare J secondo quanto indicato in 1.5 basta applicare J_S e J_T alle restrizioni marginali risultanti dallo sviluppo di f:

$$\lambda_S \sum_{j=1}^{M} J_S(m_{S_j}(\underline{p})) + \lambda_T \sum_{i=1}^{N} J_T(m_{T_i}(t)) =$$
$$= \lambda_S \sum_{j=1}^{M} \int_{\Omega_T} \Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i)\Big)^2 d\Omega + \lambda_T \sum_{i=1}^{N} \int_0^T \Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2}\Big)^2 dt$$

Dopo questa costruzione (analoga a quella riportata in [2]) non resta che calcolare i due termini con gli integrali in 1.5.

L'integrale $\int_{\Omega_T}(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i))^2 d\Omega$ implica la creazione della matrice S_S, il cui (k,l)-mo elemento è:

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

Per avere una forma utile a livello computazionale di questa espressione è necessario considerare $g = \triangle\varphi_l$ e un insieme di funzioni test v, discretizzate in spazio con le stesse funzioni di base di f:

$$g = \sum_{i=1}^{N} g_i \varphi_i$$

$$v = \sum_{i=1}^{N} v_i \varphi_i \qquad v \in \mathbb{R}^N$$

e si ha:

$$S_{S_{k,l}} = \int_{\Omega} \triangle\varphi_k g \;, \qquad \int_{\Omega} gv = \int_{\Omega} \triangle\varphi_l v \qquad \forall v \in \mathbb{R}^N$$

Usando la formula di Green ed eliminando gli integrali di bordo grazie alle condizioni di Neumann, si ricava:

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

Di conseguenza, se si definiscono i vettori con le funzioni di base in spazio e i vettori con le loro derivate parziali

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

$$\underline{\varphi}_x = \begin{bmatrix} \partial\varphi_1/\partial x \\ \partial\varphi_2/\partial x \\ \vdots \\ \partial\varphi_n/\partial x \end{bmatrix} \qquad \underline{\varphi}_x = \begin{bmatrix} \partial\varphi_1/\partial y \\ \partial\varphi_2/\partial y \\ \vdots \\ \partial\varphi_n/\partial y \end{bmatrix} \tag{1.6}$$

e le matrici

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

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

si ha, per l'arbitrarietà di v:

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

dove $\underline{g}$ contiene i coefficienti dell'espansione in base di g. Quindi in conclusione

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

Per passare all'integrale, però, occorre considerare anche i coefficienti del vettore $\underline{c}_j = \begin{bmatrix} c_{1j} & c_{2j} & \dots & c_{Nj} \end{bmatrix}^T$ e si ha:

$$\int_{\Omega_T} \Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i) \Big)^2 d\Omega = \underline{c}_j^T S_S \underline{c}_j.$$

Riguardo al termine $\int_0^T (\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2})^2 dt$, la semplificazione è molto più semplice, poichè se si considera la matrice

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

e il vettore $\underline{c}_i = \begin{bmatrix} c_{i1} & c_{i2} & \dots & c_{iM} \end{bmatrix}^T$ allora si ha:

$$\int_0^T \Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2} \Big)^2 dt = \underline{c}_i^T S_T \underline{c}_i$$

Ora che sono state ricavate le forme quadratiche associate ai due integrali, per completare lo studio di 1.5 è necessario solamente introdurre opportuni prodotti di Kronacker e l'uso del vettore $\underline{c}$:

$$\lambda_S \sum_{j=1}^{M} \int_{\Omega_T} \Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i) \Big)^2 d\Omega + \lambda_T \sum_{i=1}^{N} \int_0^T \Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2} \Big)^2 dt =$$
$$= \lambda_S \; \underline{c}^T (I_M \otimes S_S) \; \underline{c} \; + \; \lambda_T \; \underline{c}^T (S_T \otimes I_N) \; \underline{c}$$

dove I_M and I_N sono matrici identità di dimensioni $M \times M$ e $N \times N$ rispettivamente. Di conseguenza se

$$S = \lambda_S \; (S_S \otimes I_M) \; + \; \lambda_T \; (I_N \otimes S_T).$$

allora

$$\lambda_S \sum_{j=1}^{M} \int_{\Omega_T} \Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i) \Big)^2 d\Omega + \lambda_T \sum_{i=1}^{N} \int_0^T \Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2} \Big)^2 dt = \underline{c}^T S \underline{c}$$

Soluzione

Per avere una forma matriciale della versione discreta di $J_{\underline{\lambda}}(f)$, cioè di

$$J^D_{\underline{\lambda}}(f) = \sum_{i=1}^{n}\sum_{j=1}^{m}\big(z_{ij} - f(\underline{p}_i, t_j)\big)^2 + \\ + \lambda_S \sum_{j=1}^{M}\int_{\Omega_T}\Big(\triangle(\sum_{i=1}^{N} c_{ij}\varphi_i)\Big)^2 d\Omega + \lambda_T \sum_{i=1}^{N}\int_0^T\Big(\frac{\partial^2 \sum_{j=1}^{M}(c_{ij}\psi_j)}{\partial t^2}\Big)^2 dt$$

occorre definire il vettore $\underline{z}$ dei valori osservati

$$\underline{z} = \begin{bmatrix} z_{11} \\ \vdots \\ z_{1m} \\ z_{21} \\ \vdots \\ z_{2m} \\ \vdots \\ z_{nm} \end{bmatrix} \tag{1.7}$$

e le matrici Φ (con le valutazioni delle basi spaziali nei punti $\{\underline{p}_i; i = 1, ..., n\}$) e Ψ (con le valutazioni delle basi temporali $\{t_j; j = 1, ..., m\}$):

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

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

Sia Π il prodotto di Kronecker Φ e Ψ:

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

Allora

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

Quindi si avrà:

$$J^D_{\underline{\lambda}}(f) = (\underline{z} - \Pi\underline{c})^T(\underline{z} - \Pi\underline{c}) + \underline{c}^t S\underline{c} \tag{1.8}$$

Una volta che è stata ricavata questa forma, per risolvere il problema di minimo è sufficiente derivare questa espressione, e si ritrova:

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

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):

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

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 $\underline{\hat{z}}$:

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

and D is the deviance of the model:

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

1.2 Caso con covariate

Il modello si estende facilmente se si prevede che il dato possa essere influenzato da covariate. Il modello di (1.1) diventa:

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

dove $\underline{w}_{ij}$ è il vettore delle p covariate associate a z_{ij} e $\underline{\beta}$ è il vettore dei coefficienti di regressione. Di conseguenza, il funzionale discreto di (1.8) diventa:

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

dove W è la matrice $nm \times p$ con i vettori $\{\underline{w}_{ij}; i = 1, ..., n; j = 1, ..., m\}$.

Per ricavare la soluzione occorre derivare questa espressione rispetto a $\underline{\beta}$ e $\underline{c}$:

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

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

Imponendo che le derivate siano uguali a zero si hanno le seguenti equazioni:

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

che ricordano le equazioni usate per la regressione e per il modello senza covariate, con la differenza che in questo caso a $\underline{z}$ è sottratto in entrambi i casi la parte spiegata dal termine di modello a cui non si riferiscono $\underline{\hat{\beta}}$ e $\underline{\hat{c}}$ rispettivamente

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)