

Capitolo 1

Descrizione del modello

In questo capitolo viene descritto nel dettaglio il modello proposto per l'analisi dei dati in spazio e tempo. Il modello è un'estensione al caso tempo-variante di quanto riportato in [3] e si distingue in due casi, in base alla presenza di possibili covariate.

1.1 Caso senza covariate

Dati e modello

Siano $\{p_i = (x_i, y_i); i = 1, \dots, n\}$ un insieme di n punti spaziali in un dominio limitato $\Omega \subset \mathbb{R}^2$ e siano $\{t_j; j = 1, \dots, m\}$ 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 p_i al tempo t_j .

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

$$z_{ij} = f(p_i, t_j) + \varepsilon_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, m, \quad (1.1)$$

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

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

$$J_{\underline{\lambda}}(f) = \sum_{i=1}^n \sum_{j=1}^m (z_{ij} - f(p_i, t_j))^2 + \lambda_S \int_0^T \int_{\Omega} (\Delta f)^2 d\Omega dt + \lambda_T \int_{\Omega} \int_0^T \left(\frac{\partial^2 f}{\partial t^2} \right)^2 dt d\Omega. \quad (1.2)$$

Il primo termine $J_{\underline{\lambda}}(f)$ tende a minimizzazione dello scarto quadratico tra i dati e la funzione f calcolata nei corrispondenti punti spaziali e istanti temporali. Tuttavia non è l'unico, in quanto in aggiunta il funzionale presenta

due termini di smoothing il cui obiettivo nel processo di minimizzazione è di rendere liscia 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 di f in spazio. Notoriamente il laplaciano esprime la curvatura locale della funzione, quindi il suo integrale è un ottimo modo per avere 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 $\begin{bmatrix} \lambda_S \\ \lambda_T \end{bmatrix}$ 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 $\underline{\lambda}$ (sez. 1.3).

Spazio funzionale per f

Prima di procedere con la analisi del modello, occorre definire uno spazio funzionale per $f(\underline{p}, t)$ coerente con l'analisi e in cui siano definite le derivate necessarie. Sia definita la funzione a tempo fissato

$$f_t(\underline{p}) : \underline{p} \rightarrow f(\underline{p}, t) \quad \underline{p} \in \Omega \quad t \in [0, T] .$$

Come primo passo, si richiede che

$$f_t(\underline{p}) \in H_0^2(\Omega) = \{h \in H^2(\Omega) | \partial_\nu h = 0 \text{ su } \partial\Omega\} \quad \forall t \in [0, T]$$

cioè le restrizioni della funzione a tempo fissato appartengono allo spazio di Sobolev $H^2(\Omega)$ con condizioni di Neumann alla frontiera.

Successivamente si impone la condizione sul tempo, per poter garantire la derivabilità:

$$\left(\int_0^T (\|f_t(\underline{p})\|_{H^2}^2) dt \right) < \infty$$

Grazie a questa costruzione è stato ricavato lo spazio funzionale per f , che garantisce sufficiente regolarità alla funzione per poter definire il funzionale di 1.2. Questo spazio sarà indicato con $L^2(0, T; H_0^2(\Omega))$.

Sviluppo in funzioni di base di $f(\underline{p}, t)$

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

Sia $\{\varphi_i(\underline{p}); i = 1, \dots, N\}$ un insieme di N basi spaziali definite sul dominio Ω e $\{\psi_j(t); j = 1, \dots, M\}$ 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 i due insiemi:

$$f(\underline{p}, t) = \sum_{i=1}^N \sum_{j=1}^M c_{ij} \varphi_i(\underline{p}) \psi_j(t). \quad (1.3)$$

Nel codice implementato sono disponibili basi per elementi finiti in spazio e B-splines in tempo. Per coerenza con queste scelte da questo punto in avanti il dominio spaziale non sarà più insicato con Ω , ma con la sua triangolazione Ω_T .

La prima conseguenza dello sviluppo in basi di $f(\underline{p}, t)$ è che risulta possibile identificarla semplicemente con i valori dei coefficienti $\{c_{ij}; i = 1, \dots, N, j = 1, \dots, M\}$. Quindi l'obiettivo della stima sarà il vettore contenente questi coefficienti:

$$\underline{c} = \begin{bmatrix} c_{11} \\ \vdots \\ c_{1M} \\ c_{21} \\ \vdots \\ c_{2M} \\ \vdots \\ c_{NM} \end{bmatrix} \quad (1.4)$$

e si dimostrerà che sarà soluzione di un sistema lineare.

Discretizzazione dei termini di smoothing del funzionale $J_{\underline{\lambda}}(f)$

Dopo aver fissato le basi dello sviluppo di f e identificato la funzione con il vettore \underline{c} , è necessario riscrivere in forma matriciale anche il funzionale $J_{\underline{\lambda}}(f)$ in 1.2 per poter risolvere il problema. La parte più complessa da trattare di $J_{\underline{\lambda}}(f)$ è rappresentata dai due termini di smoothing in spazio e tempo, che si semplifica considerando la seguente discretizzazione:

$$\begin{aligned} J_{\underline{\lambda}}^D(f) = & \sum_{i=1}^n \sum_{j=1}^m (z_{ij} - f(\underline{p}_i, t_j))^2 + \\ & + \lambda_S \sum_{j=1}^M \int_{\Omega_T} \left(\Delta \left(\sum_{i=1}^N c_{ij} \varphi_i(\underline{p}) \right) \right)^2 d\Omega + \lambda_T \sum_{i=1}^N \int_0^T \left(\frac{\partial^2 \sum_{j=1}^M (c_{ij} \psi_j(t))}{\partial t^2} \right)^2 dt. \end{aligned} \quad (1.5)$$

La scelta di questa discretizzazione è giustificata dalla separabilità dei termini di smoothing in spazio e tempo di $J_{\underline{\lambda}}(f)$ e dalla forma di f nell'espressione

in funzioni di base in 1.3. Infatti f è derivata dalle combinazioni delle basi dei modelli marginali in spazio e in tempo, e si può anche scrivere nelle seguenti forme:

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

dove

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

$$m_{T_i}(t) = \sum_{j=1}^M c_{ij} \psi_j(t) \quad \forall i = 1 \dots N ,$$

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 in spazio, che indicheremo con J_S) così come lo smoothing in tempo (l'integrale su $[0, T]$ della derivata seconda in tempo, che indicheremo con J_T). Tali modelli marginali sono già stati trattati in letteratura con le stesse basi dei casi che saranno analizzati in seguito (elementi finiti in spazio e B-splines in tempo), quindi per discretizzare gli ultimi due termini di $J_\lambda(f)$ secondo quanto indicato in 1.5 basta applicare J_S e J_T alle restrizioni marginali risultanti dallo sviluppo di f :

$$\lambda_S \sum_{j=1}^M J_S(m_{S_j}(\underline{p})) + \lambda_T \sum_{i=1}^N J_T(m_{T_i}(t)) =$$

$$= +\lambda_S \sum_{j=1}^M \int_{\Omega_T} \left(\Delta \left(\sum_{i=1}^N c_{ij} \varphi_i(\underline{p}) \right) \right)^2 d\Omega + \lambda_T \sum_{i=1}^N \int_0^T \left(\frac{\partial^2 \sum_{j=1}^M (c_{ij} \psi_j(t))}{\partial t^2} \right)^2 dt .$$

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

L'integrale $\int_{\Omega_T} (\Delta(\sum_{i=1}^N c_{ij} \varphi_i))^2 d\Omega$ si può calcolare ricorrendo nuovamente alle funzioni marginali in spazio m_{S_j} . Fissato j , si ha:

$$\int_{\Omega_T} (\Delta(\sum_{i=1}^N c_{ij} \varphi_i))^2 d\Omega = \int_{\Omega_T} (\Delta m_{S_j}(\underline{p}))^2 d\Omega .$$

Saranno omessi per semplicità gli argomenti delle funzioni. Per avere una forma utile a livello computazionale di questa espressione è necessario considerare $g = \Delta m_{S_j}$ e approssimarla esprimendola con le basi in spazio:

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

Inoltre, occorre definire un insieme di funzioni test v , secondo l'espansione di basi in spazio e al variare dei coefficienti delle basi:

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

Fatto ciò, si potrà considerare la seguente quantità:

$$\int_{\Omega} \Delta m_{S_j} g d\Omega$$

con

$$\int_{\Omega} g v d\Omega = \int_{\Omega} \Delta m_{S_j} v d\Omega \quad \forall v \in \mathbb{R}^N .$$

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

$$\int_{\Omega} g v d\Omega = - \int_{\Omega} \nabla m_{S_j} \nabla g d\Omega .$$

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

$$\begin{aligned} \underline{\varphi} &= \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix} \\ \underline{\varphi}_x &= \begin{bmatrix} \partial \varphi_1 / \partial x \\ \partial \varphi_2 / \partial x \\ \vdots \\ \partial \varphi_n / \partial x \end{bmatrix} \quad \underline{\varphi}_y = \begin{bmatrix} \partial \varphi_1 / \partial y \\ \partial \varphi_2 / \partial y \\ \vdots \\ \partial \varphi_n / \partial y \end{bmatrix} \end{aligned} \quad (1.6)$$

e le matrici

$$\begin{aligned} R_0 &= \int_{\Omega} \underline{\varphi} \underline{\varphi}^T \\ R_1 &= \int_{\Omega} (\underline{\varphi}_x \underline{\varphi}_x^T + \underline{\varphi}_y \underline{\varphi}_y^T) . \end{aligned}$$

si ha, per l'arbitrarietà di v :

$$\int_{\Omega} \Delta m_{S_j} g d\Omega = - \underline{c}_j^T R_1 \underline{g} , \quad R_0 \underline{g} = - R_1 \underline{c}_j$$

dove \underline{g} e \underline{v} contengono i coefficienti dell'espansione in base di g e v rispettivamente, e $\underline{c}_j = [c_{1j} \ c_{2j} \ \dots \ c_{Nj}]^T$. Quindi in conclusione, se

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

si avrà:

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

Si noti che la matrice S_S non dipende dai coefficienti di \underline{c}_j ed è simmetrica poichè R_0 e R_1 lo sono.

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 $\underline{c}_i = [c_{i1} \ c_{i2} \ \dots \ c_{iM}]^T$ allora si ha:

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

Anche la matrice S_T è simmetrica.

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} :

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

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

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

A causa della simmetria dei termini con cui è costruita, S è simmetrica.

Soluzione del problema di stima

Per avere una forma matriciale della versione discreta di

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

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

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

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

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

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

Sia Π il prodotto di Kronecker Φ e Ψ :

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

Questa matrice consente di poter dire che

$$\begin{bmatrix} f(\underline{p}_1, t_1) \\ \vdots \\ f(\underline{p}_1, t_m) \\ f(\underline{p}_2, t_1) \\ \vdots \\ f(\underline{p}_2, t_m) \\ \vdots \\ f(\underline{p}_n, t_m) \end{bmatrix} = \Pi \underline{c} , \quad (1.8)$$

quindi si ha:

$$J_{\underline{\lambda}}^D(f) = (\underline{z} - \Pi \underline{c})^T (\underline{z} - \Pi \underline{c}) + \underline{c}^t S \underline{c} . \quad (1.9)$$

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

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

porre uguale a zero, da cui

$$(\Pi^T \Pi + S) \underline{c} = \Pi^T \underline{z} \quad (1.10)$$

e si ha in conclusione

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

Grazie a questo risultato è possibile ricavare i valori stimati della funzione f nei punti dati:

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

dove H rappresenta l'analogo della *hat matrix* del classico modello di regressione in questo caso.

Proprietà statistiche di $\hat{\underline{c}}$

Il modello di partenza indicato in 1.1 può essere scritto anche in forma matriciale

$$\underline{z} = \Pi \underline{c} + \underline{\varepsilon} . \quad (1.11)$$

A causa delle proprietà statistiche di $\underline{\varepsilon}$

$$\mathbb{E}[\underline{\varepsilon}] = \underline{0} \quad \text{Var}[\underline{\varepsilon}] = \sigma^2 I_{nm}$$

sia ha

$$\mathbb{E}[\underline{z}] = \Pi \underline{c} \quad \text{Var}[\hat{\underline{c}}] = \sigma^2 I_{nm}$$

e quindi è immediato ricavare per lo stimatore $\hat{\underline{c}}$ (grazie alle proprietà simmetria di S):

$$\mathbb{E}[\hat{\underline{c}}] = (\Pi^T \Pi + S)^{-1} \Pi^T \Pi \underline{c} \quad \text{Var}[\underline{z}] = \sigma^2 (\Pi^T \Pi + S)^{-1} \Pi^T \Pi (\Pi^T \Pi + S)^{-1}.$$

Non è stata ipotizzata la gaussianità per $\underline{\varepsilon}$, ma se fosse ipotizzata anche $\hat{\underline{c}}$ sarebbe gaussiano. Grazie a questa ulteriore ipotesi sarebbe possibile elaborare (con una data significatività) una regione di confidenza per $\hat{\underline{c}}$ e quindi una banda di confidenza per la funzione stimata f .

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:

$$z_{ij} = \underline{w}_{ij}^T \underline{\beta} + f(\underline{p}_i, t_j) + \varepsilon_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, m, \quad (1.12)$$

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 1.9 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 $\{\underline{w}_{ij}; i = 1, \dots, n; j = 1, \dots, m\}$.

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

$$\begin{aligned} \frac{\partial}{\partial \underline{\beta}} J &= -2W^T \underline{z} + 2W^T \Pi \underline{c} + 2W^T W \underline{\beta}, \\ \frac{\partial}{\partial \underline{c}} J &= -2\Pi^T \underline{z} + 2\Pi^T W \underline{\beta} + 2(\Pi^T \Pi + S) \underline{c}. \end{aligned}$$

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

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

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 $\hat{\underline{\beta}}$ e $\hat{\underline{c}}$ rispettivamente.

A questo punto si possono ricavare le soluzioni. Si ha:

$$\hat{\underline{c}} = [\Pi^T \Pi + S + \Pi^T W (W^T W)^{-1} W^T \Pi]^{-1} \Pi^T [I - W (W^T W)^{-1} W^T] \underline{z} = A Q \underline{z}$$

con $A = [\Pi^T \Pi + S + \Pi^T W (W^T W)^{-1} W^T \Pi]^{-1} \Pi^T$ e $Q = [I - W (W^T W)^{-1} W^T]$, matrice molto importante nel caso di regressione lineare, poiché proietta il vettore dei dati nel sottospazio ortogonale allo spazio generato dalle colonne della matrice disegno, ricavando così il vettore dei residui. Questa matrice si ritrova anche in questo caso, ed è ancora valide le sue proprietà:

- Q è idempotente, cioè $QQ = Q$;
- Q è simmetrica;
- a causa del fatto che proietta nel sottospazio ortogonale di $\text{Col}(W)$, QW risulta essere la matrice nulla di opportune dimensioni.

Infine, la stima di $\hat{\underline{\beta}}$ si ricava dalla stima ottenuta per $\hat{\underline{c}}$:

$$(W^T W)^{-1} W^T (I - \Pi A Q) \underline{z}$$

In modo analogo al caso senza covariate, è necessario ricavare anche per questo caso la *hat matrix*:

$$\hat{\underline{z}} = \Pi \hat{\underline{c}} + W \hat{\underline{\beta}} = [\Pi A Q + W (W^T W)^{-1} W^T (I - \Pi A Q)] \underline{z} = H \underline{z}.$$

Proprietà statistiche di $\hat{\underline{c}}$ e $\hat{\underline{\beta}}$

Anche in questo caso è possibile ricavare valore atteso e varianza degli stimatori ottenuti, ed è utile in quanto consente di verificare la significatività delle covariate. Per farlo però, è necessario avere la forma matriciale del modello indicato in 1.12:

$$\underline{z} = \Pi \underline{c} + W \underline{\beta} + \underline{\varepsilon}. \quad (1.13)$$

Di nuovo si ha

$$\mathbb{E}[\underline{\varepsilon}] = \underline{0} \quad \text{Var}[\underline{\varepsilon}] = \sigma^2 I_{nm}$$

e di conseguenza

$$\mathbb{E}[\underline{z}] = \Pi \underline{c} + W \underline{\beta} \quad \text{Var}[\underline{z}] = \sigma^2 I_{nm}.$$

Grazie a ciò e alle proprietà ricavate per la matrice Q è possibile ricavare che:

$$\mathbb{E}[\hat{\underline{c}}] = A Q \Pi \underline{c} \quad \text{Var}[\hat{\underline{c}}] = \sigma^2 A Q A^T.$$

Per $\hat{\underline{\beta}}$ i calcoli sono più complessi, ma si semplificano grazie alle proprietà indicate in precedenza per la matrice Q . Si ritrova:

$$\mathbb{E}[\hat{\underline{\beta}}] = \underline{\beta} + (W^T W)^{-1} W^T (I - \Pi A \Pi) \underline{c}$$

$$\text{Var}[\hat{\underline{\beta}}] = \sigma^2 (W^T W)^{-1} + \sigma^2 (W^T W)^{-1} W^T \Pi A Q A^T \Pi^T W (W^T W)^{-1}.$$

Come nel caso senza covariate, anche ora si potrebbe stimare la gaussianità degli stimatori se fosse ipotizzata per $\underline{\varepsilon}$. Questo permette di elaborare intervalli di confidenza per le componenti di $\underline{\beta}$ e di verificare la significatività delle covariate.

1.3 Stima di σ^2 e scelta dei parametri λ_S e λ_T

Quanto riportato di seguito è valido indipendentemente dal fatto che siano inserite nel modello le covariate. Perciò è valido per entrambi i modelli proposti in precedenza.

Stima di σ^2

Stimare la varianza dell'errore è necessario se si vuole fare inferenza sugli stimatori ed è molto semplice se si conoscono i gradi di libertà equivalenti del modello. Ma questi si ricavano dalla *hat matrix*:

$$\text{EDF} = \text{tr}(H) .$$

Con questo valore si calcola la stima della varianza, usando i residui e il numero totale di dati:

$$\hat{\sigma}^2 = \frac{1}{nm - \text{tr}(H)} (\underline{z} - \hat{\underline{z}})^T (\underline{z} - \hat{\underline{z}})$$

Parametri di smoothing

I parametri λ_S e λ_T hanno un ruolo rilevante nella stima della soluzione, poichè scelgono quanto peso dare allo smoothing della funzione in spazio e tempo. Quindi è opportuno che siano fissati accuratamente prima della stima della soluzione.

Secondo quanto indicato in [2], la scelta corretta si ha con il valore di $\underline{\lambda}$ che realizza il minimo dell'indice di *generalized cross validation*

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

dove D è la devianza del modello. Si ha:

$$D(\hat{\underline{c}}, \hat{\underline{\beta}}) = 2\sigma^2 (l_{\text{sat}} - l(\hat{\underline{c}}, \hat{\underline{\beta}})) ,$$

dove l è la logverosimiglianza del modello, che si ipotizza gaussiano, valutata rispettivamente nel suo massimo (valore di saturazione) e in corrispondenza dei valori stimati. Non è difficile dimostrare che, sia nel caso con covariate che senza covariate, si ha:

$$D(\hat{\underline{c}}, \hat{\underline{\beta}}) = (\underline{z} - \hat{\underline{z}})^T (\underline{z} - \hat{\underline{z}})$$

Di conseguenza, il miglior $\underline{\lambda}$ può essere scelto come valore che minimizza

$$GCV(\underline{\lambda}) = \frac{nm}{nm - \text{tr}(H)} (\underline{z} - \hat{\underline{z}})^T (\underline{z} - \hat{\underline{z}}) .$$

Bibliografia

- [1] Nicole H. Augustin, Verena M. Trenkel, Simon N. Wood, Pascal Lorange, *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)