Modelling data observed irregularly over space and regularly in time

Chrysoula Dimitriou-Fakalou [1]

[1] Department of Statistics, London School of Economics, London, WC2A 2AE, UK

Abstract: When the data has been collected regularly over time and irregularly over space, it is difficult to impose an explicit auto-regressive structure over the space, as it is over time. We study a phenomenon on a number of fixed locations and, for each location, the process forms an auto-regressive time series. The dependence over space is reflected by the covariance matrix of the noise process, which is 'white' in time but not over the space. We consider the asymptotic properties of our inference methods when the number of observations in time only tends to infinity. We refer to Hjellvik and Tjøstheim (1999) and Zhang, Yao, Tong, and Stenseth (2003), who used similar settings.

Keywords: Predictor; Multivariate; Gaussian; Likelihood.

1 Spatial modelling

For any N locations $\mathbf{s}_1, \cdots, \mathbf{s}_N$, we consider the zero-mean random vector

$$\mathbf{Y} = [Y(\mathbf{s}_1), \cdots, Y(\mathbf{s}_N)]^\tau \tag{1}$$

and we let

$$\gamma_{j,k} = E\{Y(\mathbf{s}_j)Y(\mathbf{s}_k)\},\ \ j,k = 1, \cdots, N. \tag{2}$$

We assume that $\gamma_{j,j} < \infty$, for all $j = 1, \cdots, N$, and that the matrix

$$\mathbf{V} \equiv Var\{\mathbf{Y}\} \tag{3}$$

is positive-definite. We also write down the inverse variance matrix

$$\mathbf{V}^{-1} \equiv \begin{bmatrix} a_{1,1} & -a_{1,2} & \cdots & -a_{1,N} \\ -a_{1,2} & a_{2,2} & \cdots & -a_{2,N} \\ \vdots & & \ddots & \\ -a_{1,N} & -a_{2,N} & & a_{N,N} \end{bmatrix} \tag{4}$$

and let the $(q \times 1)$ parameter vector

$$\mathbf{a} = [a_{1,2}, \cdots, a_{1,N}, a_{2,3}, \cdots, a_{2,N}, \cdots, a_{N-1,N}, a_{1,1}, \cdots, a_{N,N}]^\tau, \tag{5}$$

where $q = N(N+1)/2$. We consider the parameter space, such that the following condition is true.

(C1) For any $\mathbf{a} \in \mathcal{A}$, it holds that all the eigenvalues of $\mathbf{V}$ are positive and finite.

For observations $\{Y(\mathbf{s}_j),\ j = 1, \cdots, N\}$, we may write down the Gaussian likelihood as a function of the parameters

$$L(\mathbf{a}) \equiv \frac{1}{(\sqrt{2\pi})^N} |\mathbf{V}^{-1}|^{1/2} \exp\{-\frac{1}{2}[\sum_{j=1}^{N} a_{j,j}\, Y(\mathbf{s}_j)^2 - 2 \sum_{\substack{j,k=1,\\ j<k}}^{N} a_{j,k}\, Y(\mathbf{s}_j)Y(\mathbf{s}_k)]\}, \tag{6}$$

for any $\mathbf{a} \in \mathcal{A}$.

1.1 Interpretation of the inverse covariance matrix and best linear predictors

We try now to interpret the elements of $\mathbf{V}^{-1}$ in a way that favors the fact that we are dealing with N locations, which cannot be ordered naturally. For any location $j = 1, \cdots, N$, we define

$$Y^*(\mathbf{s}_j) \equiv \sum_{\substack{k=1,\\ k\neq j}}^{N} \beta_{j,k}\, Y(\mathbf{s}_k), \tag{7}$$

to be the best linear predictor of $Y(\mathbf{s}_j)$, based on all other sites available $k = 1, \cdots, N,\ k \neq j$, in the sense that

$$E\{Y(\mathbf{s}_j) - Y^*(\mathbf{s}_j)\}^2 \equiv \min_{\{\psi_{j,k}\}} E\{Y(\mathbf{s}_j) - \sum_{\substack{k=1,\\ k\neq j}}^{N} \psi_{j,k}\, Y(\mathbf{s}_k)\}^2. \tag{8}$$

Then it holds that

$$Cov\{Y(\mathbf{s}_j) - Y^*(\mathbf{s}_j), Y(\mathbf{s}_k)\} = 0,\ k = 1, \cdots, N,\ k \neq j, \tag{9}$$

or, similarly,

$$\gamma_{j,k} - \sum_{\substack{m=1,\\ m\neq j}}^{N} \beta_{j,m} \cdot \gamma_{m,k} = 0,\ k = 1, \cdots, N,\ k \neq j. \tag{10}$$

On the other hand, we can define the prediction variances as

$$\begin{aligned} \nu_j &\equiv Var\{Y(\mathbf{s}_j) - Y^*(\mathbf{s}_j)\} = Cov\{Y(\mathbf{s}_j) - Y^*(\mathbf{s}_j), Y(\mathbf{s}_j)\} \\ &= \gamma_{j,j} - \sum_{\substack{m=1,\\ m\neq j}}^{N} \beta_{j,m} \cdot \gamma_{m,j},\ j = 1, \cdots, N. \end{aligned} \tag{11}$$

Thus, we may re-write (10) and (11) as

$$\gamma_{j,k}\{1/\nu_j\} - \sum_{\substack{m=1,\\ m\neq j}}^{N} \gamma_{m,k} \cdot \beta_{j,m}\{1/\nu_j\} \;=\; 0,\; j,k = 1,\cdots,N,\; k \neq j, \quad (12)$$

$$\gamma_{j,j}\{1/\nu_j\} - \sum_{\substack{m=1,\\ m\neq j}}^{N} \gamma_{m,j} \cdot \beta_{j,m}\{1/\nu_j\} \;=\; 1,\; j = 1,\cdots,N, \quad (13)$$

respectively. Equations (12) and (13) imply the decomposition

$$\mathbf{V}^{-1} = \mathbf{\Lambda}^{-1}\mathbf{B}, \quad (14)$$

where $\mathbf{\Lambda}$ is a diagonal matrix with (j,j)-th element equal to $\nu_j,\ j = 1,\cdots,N$, and $\mathbf{B}$ is a matrix with 1 on the main diagonal and elements $-\beta_{j,k}$ at the j-th row and k-th column, $j,k = 1,\cdots,N,\ j \neq k$. For Gaussian random variables, the best linear predictors are conditional expectations and the prediction variances are conditional variances and such a decomposition has been performed by Besag (1975).

2 Spatio-temporal modelling

Let $\{Y_t(\mathbf{s}_j),\ t \in \mathcal{Z},\ j = 1,\cdots,N\}$ be a real-valued process observed over time and on N fixed locations. We consider for every location $j = 1,\cdots,N$, and for fixed positive integer p, a causal auto-regression model

$$Y_t(\mathbf{s}_j) \equiv b_{1,j}\ Y_{t-1}(\mathbf{s}_j) + \cdots + b_{p,j}\ Y_{t-p}(\mathbf{s}_j) + u_t(\mathbf{s}_j), \quad (15)$$

where

$$\mathbf{u}_t \equiv [u_t(\mathbf{s}_1),\cdots,u_t(\mathbf{s}_N)]^\tau \sim WN(\mathbf{0},\mathbf{V}). \quad (16)$$

We stack the parameter vectors

$$\mathbf{b}_j \equiv [b_{1,j},\cdots,b_{p,j}]^\tau,\ j = 1,\cdots,N, \quad (17)$$

and

$$\mathbf{b} \equiv [\mathbf{b}_1^\tau,\cdots,\mathbf{b}_N^\tau]^\tau. \quad (18)$$

All the processes $\{Y_t(\mathbf{s}_j),\ t \in \mathcal{Z}\},\ j = 1,\cdots,N$, which have been defined by (15) and (16) form a multivariate auto-regression of order p. We consider the parameter space, such that the following condition is true.

(C2) For any $\mathbf{b}_j \in \mathcal{B},\ j = 1,\cdots,N$, a causal auto-regression (15) is defined.

For positive integer $T > p$, we collect observations $\{Y_t(\mathbf{s}_j),\ t = 1-p,\cdots,T,\ j = 1,\cdots,N\}$ and wish to estimate the true parameter vector

$[\mathbf{b}_0^\tau, \mathbf{a}_0^\tau]^\tau$ that has generated the observations. We assume that the parameter vector is an inner point of the parameter space. For any $\mathbf{b}_j \in \mathcal{B},\ j = 1, \cdots, N$, we define

$$u_t(\mathbf{s}_j, \mathbf{b}_j) \equiv Y_t(\mathbf{s}_j) - b_{1,j}\ Y_{t-1}(\mathbf{s}_j) - \cdots - b_{p,j}\ Y_{t-p}(\mathbf{s}_j),\ t \in \mathcal{Z}. \qquad (19)$$

Then, we may write down the Gaussian likelihood

$$L(\mathbf{a}, \mathbf{b}) \equiv \frac{1}{(\sqrt{2\pi})^{NT}} |\mathbf{V}^{-1}|^{T/2} \qquad \exp\{-\frac{1}{2}\sum_{t=1}^{T}[\sum_{j=1}^{N} a_{j,j}\ u_t(\mathbf{s}_j, \mathbf{b}_j)^2 - 2\sum_{\substack{j,k=1,\\ j<k}}^{N} a_{j,k}\ u_t(\mathbf{s}_j, \mathbf{b}_j) u_t(\mathbf{s}_k, \mathbf{b}_k)]\}, \qquad (20)$$

for all $\mathbf{b}_j \in \mathcal{B},\ j = 1, \cdots, N$, and $\mathbf{a} \in \mathcal{A}$.

2.1 Estimators

We define our Gaussian likelihood estimators $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}} \equiv [\hat{\mathbf{b}}_1^\tau, \cdots, \hat{\mathbf{b}}_N^\tau]^\tau$, such that (20) is maximized. We may now state the next theorem.
Theorem. Under conditions (C1) and (C2), if

$$\mathbf{u}_t \sim IID(\mathbf{0}, \mathbf{V}),$$

then as $T \to \infty$ (N fixed), it holds that
(i) the estimators are consistent,
(ii) the estimators are asymptotically normal,
(iii) the estimators $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are asymptotically independent.

Acknowledgments: The paper is a part of the thesis titled 'Statistical inference for spatial and spatio-temporal processes', which will be submitted for the London School of Economics and the University of London. Therefore, Miss Dimitriou-Fakalou would like to thank the Leverhulme Trust, which funded her PhD as a part of the project titled 'Spatial and spatio-temporal modelling', as well as her supervisor Professor Q. Yao.

References

Besag, J. (1975). Statistical Analysis of Non-Lattice Data. *The Statistician*, **24**, 179-195.

Hjellvik, V. and Tjøstheim, D. (1999). Modelling Panels of Intercorrelated Autoregressive Time Series. *Biometrika*, **86**, 573-590.

Zhang, W., Yao, Q., Tong, H. and Stenseth N.C. (2003). Smoothing for Spatio-Temporal Models and its Application in Modelling Muskrat-Mink Interaction. *Biometrics*, **59**, 813-821.