STAT 8025

Lecture 12: Lattice Data (II)

Dr. Emily Lei Kang

Division of Statistics & Data Science
Department of Mathematical Sciences
University of Cincinnati

University of Cincinnati

Recall the CAR model:

$$\mathrm{Y} \sim Gau(\mu, (\mathrm{I} - \mathrm{C})^{-1}\mathrm{M})$$

where

- M is $n \times n$ diagonal, $\mathrm{M} \equiv diag(\tau_1^2, \ldots, \tau_n^2)$
- $\mathrm{C} \equiv (c_{ij})$ satisfying:
  - $c_{ii} = 0$
  - $c_{ik} = 0$ for $k \notin \partial_i$
  - $\frac{c_{ij}}{\tau_i^2} = \frac{c_{ji}}{\tau_j^2}$
- Full conditional: $Y_i | Y_j, j \neq i \sim N(\mu_i + \sum_{j=1}^{n} c_{ij}(Y_j - \mu_j), \tau_i^2)$

Special cases:

- Set $\tau_i = \tau$ and $\mathrm{C} = \gamma\mathrm{H}$.
  - H is symmetric and $h_{ii} = 0$ but not necessarily 0 or 1
  - $\mathrm{Y} \sim Gau(\boldsymbol{\mu}, \tau^2(\mathrm{I} - \gamma\mathrm{H})^{-1})$
  - $Y_i | Y_j, j \neq i \sim N(\mu_i + \gamma \sum_{j=1}^n h_{ij}(Y_j - \mu_j)$
- In some references, people use the following

$$Y_i | Y_j, j \neq i \sim N(\rho \bar{Y}_i, \frac{\sigma^2}{m_i})$$

  - $\boldsymbol{\mu} = 0$
  - $\tau_i^2 = \frac{\sigma^2}{m_i}$
  - $c_{ij} = \frac{\rho}{m_i} I(i \sim j)$
  - Joint distribution? $\mathrm{Y} \sim Gau(0, (\mathrm{I} - \rho\mathrm{D}^{-1}\mathrm{A})^{-1}\sigma^2\mathrm{D}^{-1})$

$$\mathrm{Y} \sim Gau(0, \sigma^2(\mathrm{D} - \rho\mathrm{A})^{-1})$$

where $\mathrm{D} = diag(m_1, \ldots, m_n)$, and $a_{ij} = I(i \sim j)$ for $i \neq j$ and $a_{ii} = 0$.

$$Y \sim Gau(0, \sigma^2 (D - \rho A)^{-1})$$

- When $\rho = 1$, it is called the intrinsic CAR model:

$$Y \sim Gau(0, \sigma^2 (D - A)^{-1})$$

  - $D - A$ is singular. Why?
  - It can still be used as a prior. $1'Y = 0$ is enforced in MCMC.

Computing with CAR

- The CAR model leads to much simpler computation than geostat models
  - The precision matrix appears in the likelihhood
  - In the CAR model, the precision matrix is usually sparse, making computing fast even for large n
  - In MCMC updates, we can use the full conditional distributions which are simple and local

Simultaneous Auto-Regressive (SAR) Models

- Recall from the CAR model, we have

$$\text{Y} \sim Gau(\boldsymbol{\mu}, (I - C)^{-1} M),$$

where $M \equiv diag(\tau_1^2, \ldots, \tau_n^2)$, $C \equiv (c_{ij})$, and $M^{-1}(I - C)$ is positive-definite.

- We can rewrite the CAR model as

$$\text{Y} = C\text{Y} + \varepsilon \implies (I - C)\text{Y} = \varepsilon$$

Since $\text{Y} \sim Gau(\boldsymbol{\mu}, (I - C)^{-1} M)$, we have

$$\varepsilon \sim Gau(0, M(I - C)')$$

  - Notice that now the components of ε are *not* independent
  - $cov(\varepsilon, \text{Y}) = M$

Simultaneous Auto-Regressive (SAR) Models

- Now suppose that instead of letting Y induce the distribution of ε. We let ε induce a distribution Y
- Suppose

$$\varepsilon \sim Gau(0, \tilde{M}),$$

  where $\tilde{M}$ is diagonal, $(\tilde{M})_{ii} = \sigma_i^2$
- Then assume the following for Y:

$$Y_i = \sum_j b_{ij} Y_j + \varepsilon_i$$

  - Note that $B = (b_{ij})$ is a square matrix but not necessarily symmetric. The parameters represent the spatial dependence.
  - Equivalently, we have

$$Y = BY + \varepsilon$$

Simultaneous Auto-Regressive (SAR) Models

$$Y = BY + \varepsilon$$

$$\implies (I - B)Y = \varepsilon$$

$$\implies Y = (I - B)^{-1}\varepsilon$$

provided $(I - B)$ is invertible
Hence,

$$E(Y) = 0$$

$$\mathrm{var}(Y) = \mathrm{var}((I - B)^{-1}\varepsilon) = (I - B)^{-1}\tilde{M}((I - B)^{-1})')$$

$$cov(Y, \varepsilon) = (I - B)^{-1}\tilde{M}$$

SAR Model with Trend

A SAR model is usually used in a regression context, i.e. the residuals

- Assume a trend term $X\beta$,

$$\mathrm{Y} = X\beta + (I - B)^{-1}\varepsilon$$

This is a general linear model with covariance matrix Σ

$$\Sigma = (I - B)^{-1}\tilde{M}((I - B)^{-1})')$$

$$\Sigma^{-1} = (I - B)'\tilde{M}^{-1}(I - B)$$

- Notice that the SAR model Σ^{-1}, but the geostatistics model Σ
- This represents two different ways to model spatial dependence: either through Σ (geostatistical data) or through Σ^{-1} (areal data).

SAR Model for Rates

Suppose $\mathrm{r} \equiv (r_1, \ldots, r_n)'$ are raw rates, where $r_i = \frac{Y_i}{N_i}$

$$\mathrm{r} = B\mathrm{r} + \Lambda^{1/2}\varepsilon$$

where $\mathrm{var}(\varepsilon) = \sigma^2 I$; $\Lambda^{1/2} = diag\left(\frac{1}{N_i^{1/2}}\right)$

Then

$$\mathrm{r} = (I - B)^{-1}\Lambda^{1/2}\varepsilon$$

$$\mathrm{var}(\mathrm{r}) = \sigma^2 (I - B)^{-1}\Lambda((I - B)^{-1})'$$

where $\Lambda = diag\left(\frac{1}{N_i}\right)$

Equivalent Ways to Write SAR Model

1. $\mathsf{Y} = X\boldsymbol{\beta} + (I - B)^{-1}\boldsymbol{\varepsilon}$
2. $\mathsf{Y} = X\boldsymbol{\beta} + \boldsymbol{\delta}$

   $\boldsymbol{\delta} = B\boldsymbol{\delta} + \boldsymbol{\varepsilon}$

   In this formulation, the error term $\boldsymbol{\delta}$ follows a zero-mean SAR model.

- SAR models are well suited to maximum likelihood estimation but not that straightforward for MCMC fitting of Bayesian models. However, it is still feasible if some tricks are used! (such as INLA).
- Common choice $B = \rho W$

STAR Models

- SAR models have been extended to handle spatio-temporal data: The measurements Y_{it} are spatially associated at each fixed t. But, we might also want to associate, say Y_{i2} with Y_{i1} and Y_{i3}
- Define W_s that provides a spatial proximity matrix for the Y's. And let W_t define a temporal contiguity matrix for the Y's. We can define in our SAR model

$$B = \rho_s W_s + \rho_t W_t$$

- We can also introduce W_{st} to incorporate interaction between space and time
- The resulting models are referred to as spatio-temporal autoregressive (STAR) models

Non-Gaussian

- For non-Gaussian data, we can build a generalized linear model with CAR random effects. For example:

$$logit(p_i) = X_i'\boldsymbol{\beta} + \theta_i$$

$$\boldsymbol{\theta} \sim CAR$$

- It is also possible to build a spatial Markov model directly. Recall that the results in Besag (1974) are for exponential family, instead of normal distribution only.

Recall

- We study MRFs that are Gibbs distributions with exponential-family conditional distribution and non-zero potential of order 1 and 2 only
- Assume

$$p(y_i|\mathrm{y}_{-i}) = \exp[A_i(\mathrm{y}_{\partial_i})B_i(y_i) + C_i(y_i) + D_i(\mathrm{y}_{\partial_i})]$$

- **Proposition (Besag, 1974)**: If the potential functions $G^{(k)} = 0$ for $k \geq 3$, then

$$A_i(\mathrm{y}_{\partial_i}) = \alpha_i + \sum_{j=1}^{n} \theta_{ij} B_j(y_j),$$

where $\theta_{ji} = \theta_{ij}$, $\theta_{ii} = 0$, and $\theta_{ik} = 0$ for $k \notin \partial_i$. Further, if $\theta_{ij} = 0$ for all i, j, then $Y_1, \ldots, Y_n$ are independent

Non-Gaussian Auto Models

Auto logistic:

Data are $y_i = 0$ or 1; $i = 1, \ldots, n$

Recall the Q-function:

$$Q(\mathrm{y}) \equiv ln\{p(\mathrm{y})/p(\mathrm{y}_0)\}$$

and Besag (1974) shows:

$$\begin{aligned} Q(\mathrm{y}) = \quad & \textstyle\sum_{i=1}^n y_i G_i^{(1)}(y_i) + \sum_{i=1}^n \sum_{j>i}^n y_i y_j G_{ij}^{(2)}(y_i, y_j) \\ & + \cdots + y_1 \cdots y_n G_{1\ldots n}^{(n)}(y_1, \ldots, y_n) \end{aligned}$$

We assume that the Gibbs distribution contains non-null potentials up to order 2 (i.e., $G^{(k)} = 0$ for $k \geq 3$).
Then,

$$Q(\mathrm{y}) = \sum_{i=1}^{n} \alpha_i y_i + \sum_{i=1}^{n} \sum_{j>i}^{n} \theta_{ij} y_i y_j$$

where $\alpha_i \equiv G_i^{(1)}(1)$ and $\theta_{ij} \equiv G_{ij}^{(2)}(1,1)$

Now,

$$\frac{p(y_i|\mathrm{y}_{\partial_i})}{p(0_i|\mathrm{y}_{\partial_i})} = \exp[Q(\mathrm{y}) - Q(\mathrm{y}_i)]$$

$$= \exp[\alpha_i y_i + \textstyle\sum_{j=1}^{n} \theta_{ij} y_i y_j]$$

where

$$0_i \equiv \text{“}y_i = 0\text{”}$$

$$\mathrm{y}_i \equiv (y_1, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_n)'$$

Then

$$\frac{p(1_i|\mathrm{y}_{\partial_i})}{p(0_i|\mathrm{y}_{\partial_i})} = \exp[\alpha_i + \textstyle\sum_{j=1}^{n} \theta_{ij} y_j]$$

where

$$1_i \equiv \text{“}y_i = 1\text{”}$$

Noticing $p(1_i|\mathrm{y}_{\partial_i}) + p(0_i|\mathrm{y}_{\partial_i}) = 1$, we can get

$$p(0_i|\mathrm{y}_{\partial_i}) = \left\{ 1 + \exp[\alpha_i + \sum_{j=1}^{n} \theta_{ij} y_j] \right\}^{-1}$$

Therefore,

$$p(y_i|\mathrm{y}_{\partial_i}) = \frac{\exp[\alpha_i y_i + \sum_{j=1}^{n} \theta_{ij} y_i y_j]}{1 + \exp[\alpha_i + \sum_{j=1}^{n} \theta_{ij} y_j]}$$

Note that for binary data and assuming pairwise-only dependence, exponential-family conditional distributions are a necessity, not an assumption

We thus have:

$$logit(Prob(Y_i = 1 | Y_j, j \neq i)) = \alpha_i + \sum_{j=1}^{n} \theta_{ij} Y_j$$

From Besag (1974), we need: $\theta_{ij} = \theta_{ji}$, $\theta_{ii} = 0$ and $\theta_{ik} = 0$ for $k \notin \partial_i$

The autologistic model is also called the Ising model.

- The Ising model can be generalized for categorical data. The resulting model is called the Potts model.
- There is auto-Poisson model but it has some undesirable feature, negative spatial dependence, and thus is not commonly used.

Change of Support

- In some cased we have data at different spatial resolutions
  - continuous and discrete
  - discrete but different shape/size
- Change of support: We consider these different resolutions. The fundamental relationship is

$$Y(B) = \frac{1}{|B|} \int_{s \in B} Y(s) ds$$

- In practice
  - Continuous and discrete
  - Discrete and discrete
- In some methods/applications, the spatial domain is discretized, e.g., BAUs in data fusion with FRK, the mesh in SPDE, and Meshed GP.

Reading assignments:

- Lindgren, F. and Rue, H. (2013), Bayesian Spatial and Spatio-temporal Modelling with R-INLA, *J. Stat. Softw.*
- **[To be presented]** Jay M. Ver Hoef, Ephraim M. Hanks, Mevin B. Hooten (2018) On the relationship between conditional (CAR) and simultaneous (SAR) autoregressive models, *Spatial Statistics*, 25, 68-85.
- Zhang, L., Baladandayuthapani, V., Zhu, H, Baggerly, K. A., Majewski, T., Czerniak, B. A. and Morris, J. S. (2016). Functional CAR models for large spatially correlated functional datasets, *Journal of the American Statistical Association (Theory and Methods)*, 111, 772-786.

Summary

- Lattice data
- Change of support

Preview:

- Point process