

Handout 4: Aerial unit data / spatial data on lattices

Lecturer & author: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

Aim. To introduce Aerial unit data modeling: the basic building models.**Reading list & references:**

- [1] Cressie, N. (2015; Part II). Statistics for spatial data. John Wiley & Sons.
- [2] Gaetan, C., & Guyon, X. (2010; Ch 3). Spatial statistics and modeling (Vol. 90). New York: Springer.

Specialized reading.

- [3] Kent, J. T., & Mardia, K. V. (2022). Spatial analysis (Vol. 72). John Wiley & Sons. (on Spatial analysis)

Part 1. Basic stochastic models & related concepts for model building

Note 1. Recall from Section 2.2 of “Handout 1: Types of spatial data” that modeling aerial unit / lattice data types involves the use of random field models with a discrete index set. Such data are collected over areal units such as pixels, census districts or tomographic bins. Often, there is a natural adjacency relation or neighborhood structure.

Note 2. This means we need to introduce new basic building models able to acceptably represent the characteristics of the underline data generating mechanisms. These as the “Discrete Random Fields”.

1. DISCRETE RANDOM FIELDS

Note 3. We re-introduce the definition of the random field adjusting it to the aerial unit data framework.

Definition 4. A random field $Z = (Z_s; s \in \mathcal{S})$ on a set of indexes \mathcal{S} taking values in $\mathcal{Z}^{\mathcal{S}}$ is a family of random variables $\{Z_s := Z_s(\omega); s \in \mathcal{S}, \omega \in \Omega\}$ where each $Z_s(\omega)$ is defined on the same probability space $(\Omega, \mathfrak{F}, \text{pr})$ and taking values in \mathcal{Z} .

Note 5. In aerial unite data modeling, the (spatial) set of sites \mathcal{S} , at which the process is defined, is discrete, it can be finite or infinite (e.g. $\mathcal{S} \subseteq \mathbb{Z}^d$), regular (e.g. pixels of an image) or irregular (states of a country).

Note 6. The general state space \mathcal{Z} of the random field can be quantitative, qualitative or mixed. E.g., $\mathcal{Z} = \mathbb{R}^+$ in a Gamma random field, $\mathcal{Z} = \mathbb{N}$ in a Poisson random field, $\mathcal{Z} = \{0, 1\}$ in a binary random field.

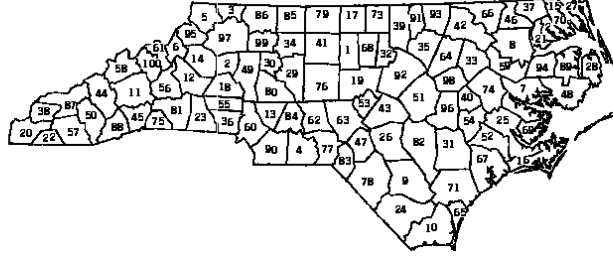


FIGURE 1.1. Lattice of spatial sites for North Carolina SIDS data. Each county is a site. Each site is coded according to its alphabetical order. The collection of sites is the lattice of sites.

Note 7. If \mathcal{Z} is finite or countably infinite, the (joint)distribution of Z has a PMF

$$\text{pr}_Z(z) = \text{pr}(Z = z) = \text{pr}(\{Z_s = z_s; s \in \mathcal{S}\}), \quad \forall z_s \in \mathcal{Z}^{\mathcal{S}}$$

otherwise if $\mathcal{Z} \subseteq \mathbb{R}^d$ and Z continuous we will use the joint PDF.

Definition 8. The discrete set of sites $\mathcal{S} = \{s_i; i = 1, \dots, n\}$ is often called lattice of sites.

Notation 9. We will more often use the notation Z_s instead of $Z(s)$ or Z_i instead of $Z(s_i)$. Hence, since $\mathcal{S} = \{s_i; i = 1, \dots, n\}$, we can consider a more convenient notation

$$Z = (Z_s; s \in \mathcal{S})^\top = (Z_i = Z(s_i); i = 1, \dots, n)^\top.$$

Notation 10. The notation $i \sim j$ between two sites $i, j \in \mathcal{S}$ means that “sites i and j are adjacent”.

Example 11. Recall the North Carolina SIDS data Ex 24 in Handout 1. Fig 1.1 presents the sites and the lattice of sites. Each county is a site. Each site is coded according to its alphabetical order. The collection of sites is the lattice of sites coded according to alphabetical order of the county name. One may define the “adjacency between sites $i \sim j$ ” as the counties that share common borders. Then for site $i = 43$, $i \sim j$ involves any $j \in \{63, 53, 19, 92, 51, 82, 26, 47\}$ in Fig 1.1.

Example 12. (Logistic/Ising model) Consider Z_i denotes a characteristic presence coded as 1 or absence coded as 0 on a region labeled by $i \in \mathcal{S}$. Then $\mathcal{Z} = \{0, 1\}$. The Ising model is defined by the (joint) PMF

$$(1.1) \quad \text{pr}_Z(z) \propto \exp \left(\alpha \sum_{i \in \mathcal{S}} z_i + \beta \sum_{\{i, j\}: i \sim j} z_i z_j \right), \quad \forall z \in \mathcal{Z}^{\mathcal{S}}$$

E.g., it can model a black & white image noisy image, where \mathcal{S} denotes the labels of the image pixels, and Z_i denotes the presence of a black pixel ($Z_i = 1$) or its absence ($Z_i = 0$). Under Ising model, the characteristic is observed with probability $\text{pr}_{Z_i}(z_i = 1) = \frac{\exp(\alpha)}{1 + \exp(\alpha)}$

when $\beta = 0$. The characteristic's present is encouraged in neiboghring sites when $\beta > 0$, and discouraged when $\beta < 0$.

Notation 13. We use notation, for $\mathcal{A} \subset \mathcal{S}$

$$\text{pr}_{\mathcal{A}}(z_{\mathcal{A}}|z_{\mathcal{S}\setminus\mathcal{A}}) = \text{pr}(Z_{\mathcal{A}} = z_{\mathcal{A}}|Z_{\mathcal{S}\setminus\mathcal{A}} = z_{\mathcal{S}\setminus\mathcal{A}})$$

Definition 14. Local characteristics of a random field Z on \mathcal{S} with values in \mathcal{Z} are the conditionals

$$\text{pr}_i(z_i|z_{\mathcal{S}-i}) = \text{pr}_{\{i\}}(z_{\{i\}}|z_{\mathcal{S}\setminus\{i\}}), \quad i \in \mathcal{S}, z \in \mathcal{Z}$$

Example 15. The characteristics of the Ising model in (1.1) are

$$\text{pr}_i(z_i = 1|z_{\mathcal{S}-i}) = \frac{\exp\left(\alpha + \beta \sum_{\{i,j\}: i \sim j} z_j\right)}{1 + \exp\left(\alpha + \beta \sum_{\{i,j\}: i \sim j} z_j\right)}$$

2. COMPATIBILITY OF CONDITIONAL DISTRIBUTIONS

Note 16. Essentially, we attempt to answer the following question. Under what conditions a parameterized family $\{\pi_i(z_i|z_{\mathcal{S}-i}); i \in \mathcal{S}\}$ of distributions on \mathcal{S} conditioned on $z_{\mathcal{S}-i}$ can represent conditional distributions of a joint distribution $\text{pr}_Z(\cdot)$?

Note 17. To answer the above we need to be able to specify partially or wholly the joint and conditional distribution of pr_Z . However, an arbitrary chosen set of conditional distributions $\{\pi_i(\cdot|\cdot)\}$ is not generally compatible, and hence we need to impose conditions.

Proposition 18. ¹(Compatibility condition) Let F be a joint distribution with $dF(x, y) = f(x, y) d(x, y)$ on $\mathcal{S}_x \times \mathcal{S}_y$. Let candidate condition distributions

$$G \text{ with } dG(x|y) = g(x|y) dx, \text{ on } x \in \mathcal{S}_x$$

$$Q \text{ with } dQ(y|x) = q(y|x) dy, \text{ on } y \in \mathcal{S}_y$$

and let $N_g = \{(x, y) : g(x|y) > 0\}$ and $N_q = \{(x, y) : q(y|x) > 0\}$. A distribution F with conditionals exists iff

$$(1) \quad N_g = N_q = N$$

$$(2) \quad \text{there exist functions } u \text{ and } v \text{ where } g(x|y)/q(y|x) = u(x)v(y) \text{ for all } (x, y) \in N \text{ and } \int u(x) dG(x|y) < \infty$$

1

Arnold, B. C., & Press, S. J. (1989). Compatible conditional distributions. Journal of the American Statistical Association, 84(405), 152-156.

Note 19. Essentially the above conditions guarantee that

$$\textcolor{red}{k}(y) g(x|y) = f(x, y) = \textcolor{red}{h}(x) q(y|x)$$

where k, g, h, q are densities.

Example 20. The conditionals $x|y \sim N(a + by, \sigma^2 + \tau^2 y^2)$ and $y|x \sim N(c + dx, \tilde{\sigma}^2 + \tilde{\tau}^2 x^2)$ are compatible if $\tau^2 = \tilde{\tau}^2 = 0$ and $d/\tilde{\sigma}^2 = b/\sigma^2$.

Solution. See Ex 24 in the Exercise sheet.

Note 21. Proposition 18 can be extended to more dimensions.

Note 22. The following theorem shows that local characteristics can determine the entire distribution in certain cases.

Theorem 23. (*Besag's factorization theorem; Brook's Lemma*) Let Z be a \mathcal{Z} valued random field taking values in $\mathcal{Z}^{\mathcal{S}}$ where $\mathcal{S} = \{1, \dots, n\}$ with $n \in \mathbb{N}$, and such as $\text{pr}_Z(z) > 0, \forall z \in \mathcal{Z}^{\mathcal{S}}$. Then for all

$$(2.1) \quad \frac{\text{pr}_Z(z)}{\text{pr}_Z(z^*)} = \prod_{i=1}^n \frac{\text{pr}_i(z_i|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^*|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}, \quad \forall z, z^* \in \mathcal{Z}^{\mathcal{S}}$$

Proof. I will show that

$$\text{pr}_Z(z) = \prod_{i=1}^n \frac{\text{pr}_i(z_i|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^*|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)} \text{pr}_Z(z^*)$$

It is

$$\text{pr}_Z(z_1, \dots, z_n) = \frac{\text{pr}_n(z_n|z_1, \dots, z_{n-2}, z_{n-1})}{\text{pr}_n(z_n^*|z_1, \dots, z_{n-2}, z_{n-1})} \text{pr}_Z(z_1, \dots, z_{n-1}, z_n^*)$$

Let proposition P_j be

$$\text{pr}_Z(z) = \prod_{i=n-j}^n \frac{\text{pr}_i(z_i|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^*|z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)} \text{pr}_Z(z_1, \dots, z_{n-j-1}, z_{n-j}^*, \dots, z_n^*)$$

Proposition P_0 is true

$$(2.2) \quad \text{pr}_Z(z) = \frac{\text{pr}_n(z_n|z_1, \dots, z_{n-1})}{\text{pr}_n(z_n^*|z_1, \dots, z_{n-1})} \text{pr}_Z(z_1, \dots, z_{n-1}, z_n^*)$$

Proposition P_1 is true

$$\text{pr}_Z(z_1, \dots, z_{n-1}, z_n^*) = \frac{\text{pr}_{n-1}(z_{n-1}|z_1, \dots, z_{n-2}, z_n^*)}{\text{pr}_{n-1}(z_{n-1}^*|z_1, \dots, z_{n-2}, z_n^*)} \text{pr}_Z(z_1, \dots, z_{n-2}, z_{n-1}^*, z_n^*)$$

Assume that P_j is true. Then proposition P_{j+1} is true as well, because

$$\begin{aligned}
\text{pr}_Z(z) &= \prod_{i=n-j}^n \frac{\text{pr}_i(z_i | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^* | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)} \text{pr}_Z(z_1, \dots, z_{n-j-1}, z_{n-j}^*, \dots, z_n^*) \\
&= \prod_{i=n-j}^n \frac{\text{pr}_i(z_i | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^* | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)} \\
&\quad \times \frac{\text{pr}_{n-j-1}(z_{n-j-1} | z_1, \dots, z_{n-j-2}, z_{n-j}^*, \dots, z_n^*)}{\text{pr}_{n-j-1}(z_{n-j-1}^* | z_1, \dots, z_{n-j-2}, z_{n-j}^*, \dots, z_n^*)} \text{pr}_Z(z_1, \dots, z_{n-j-2}, z_{n-j-1}^*, \dots, z_n^*) \\
&= \prod_{i=n-(j+1)}^n \frac{\text{pr}_i(z_i | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)}{\text{pr}_i(z_i^* | z_1, \dots, z_{i-1}, z_{i+1}^*, \dots, z_n^*)} \text{pr}_Z(z_1, \dots, z_{n-(j+1)-1}, z_{n-(j+1)}^*, \dots, z_n^*)
\end{aligned}$$

Then (2.1) is correct according to the induction principle. \square

Note 24. The theorem shows that the joint $\text{pr}_Z(\cdot)$ can be constructed from its conditionals $\{\text{pr}_i(\cdot|\cdot)\}$ if distributions $\{\text{pr}_i(\cdot|\cdot)\}$ are compatible for $\text{pr}_Z(\cdot)$, as this reconstruction has to be invariant wrt the coordinate permutation $\{1, \dots, n\}$ and the reference state z^* — these invariances correspond to the conditions in Proposition 18.

3. GAUSSIAN AUTOREGRESSIVE MODELS

Modeling
snapshot

Definition 25. Adjacency matrix N is called a matrix N with $[N]_{i,j} = 1$ ($i \sim j$) for some symmetric neighborhood relation \sim on \mathcal{S} . It aims at spatially connecting unites i and j .

Definition 26. Proximity matrix W is called a matrix W which aims at spatially connecting unites i and j in some fashion for some symmetric neighbourhood relation \sim on \mathcal{S} . Usually $[W]_{i,i} = 0$

3.1. Conditional autoregressive models (CAR).

Definition 27. Assume a random field $Z = (Z_s; s \in \mathcal{S})$ on a set of indexes \mathcal{S} with values in \mathcal{Z} . We say that Z follows a conditional autoregressive model (CAR) if the distribution of each element Z_s of the random field Z is specified conditionally on the values at the neighboring sites of s .

3.1.1. Gaussian CAR.

Definition 28. Gaussian CAR assumes that the local characteristics $\{\text{pr}_i(z_i | z_{\mathcal{S}-i})\}$ are Gaussian distributions

$$(3.1) \quad Z_i | z_{\mathcal{S}-i} \sim N \left(\mu_i + \sum_{j \neq i} b_{i,j} (Z_j - \mu_j), \kappa_i \right)$$

with mean $E(Z_i | Z_{\mathcal{S}-i}) = \mu_i + \sum_{j \neq i} b_{i,j} (Z_j - \mu_j)$ and variance $\text{Var}(Z_i | Z_{\mathcal{S}-i}) = \kappa_i$ for $i \in \mathcal{S}$.

Proposition 29. Let $K = \text{diag}(\{\kappa_i\})$ with $\kappa_i > 0$, matrix B with $B_{i,i} := [B]_{i,i} = 0$, and real vector μ with suitable dimensions. If Z follows a Gaussian CAR (Def 28), $I - B$ is non-singular, and $(I - B)^{-1} K > 0$, then the joint distribution of Z is

$$(3.2) \quad Z \sim N(\mu, (I - B)^{-1} K).$$

Proof. Without loss of generality, consider zero mean $\mu = 0$ (or equivalently set $Z := Z - \mu$). The full conditionals $Z_i | z_{\mathcal{S}-i}$ in (3.1) are compatible with the joint distribution $\text{pr}_Z(z)$. By using Besag's factorization theorem with reference state $z^* = 0$ we get

$$\begin{aligned} \text{pr}_Z(z) &= \prod_{i=1}^n \frac{\text{pr}_i(z_i | z_1, \dots, z_{i-1}, z_{i+1}^* = 0, \dots, z_n^* = 0)}{\text{pr}_i(z_i^* = 0 | z_1, \dots, z_{i-1}, z_{i+1}^* = 0, \dots, z_n^* = 0)} \text{pr}_Z(z^* = 0) \\ &= \prod_{i=1}^n \frac{N(z_i | \sum_{j < i} b_{i,j} z_j + 0, \kappa_i)}{N(0 | \sum_{j < i} b_{i,j} z_j + 0, \kappa_i)} \text{pr}_Z(z^* = 0) \\ &\propto \prod_{i=1}^n \exp \left(-\frac{1}{2\kappa_i} \left(z_i - \sum_{j < i} b_{i,j} z_j \right)^2 + \frac{1}{2\kappa_i} \left(0 - \sum_{j < i} b_{i,j} z_j \right)^2 \right) \\ &= \prod_{i=1}^n \exp \left(-\frac{1}{2\kappa_i} \left(z_i^2 - 2z_i \sum_{j < i} b_{i,j} z_j \right) \right) \text{pr}_Z(z^* = 0) \\ &= \exp \left(-\sum_i \frac{z_i^2}{2\kappa_i} + \frac{1}{2} \sum_i \sum_{j < i} \frac{b_{i,j}}{\kappa_i} z_i z_j \right) \text{pr}_Z(z^* = 0) \\ &= \exp \left(-\frac{1}{2} z^\top K^{-1} z + \frac{1}{2} z^\top K^{-1} B z \right) \text{pr}_Z(z^* = 0) = \exp \left(-\frac{1}{2} z^\top [K^{-1} (I - B)] z \right) \text{pr}_Z(z^* = 0) \\ (3.3) \quad &= N(z | 0, (I - B)^{-1} K) \end{aligned}$$

Recovering the mean from (3.3), it is

$$\text{pr}_Z(z) = N(z - \mu | 0, (I - B)^{-1} K) = N(z | \mu, (I - B)^{-1} K)$$

□

Note 30. When CAR is used for modeling, B is often specified to be sparse either due to some natural problem specific property, or for our computational convenience as it may allow the use of sparse solvers. To achieve this, one way is to specify $B = \phi N$ where $\phi > 0$ and N is an adjacency matrix; that is $[B]_{i,j} = \phi 1(i \sim j) 1(i \neq j)$ will be non-zero only for adjacent pairs i and j .

Note 31. The system in (3.2) can be rewritten as

$$(3.4) \quad Z = \mu + B(Z - \mu) + E \iff E = (I - B)(Z - \mu)$$

by setting $E = (I - B)(Z - \mu)$. The distribution of Z in (3.2) induces a distribution on E as $E \sim N(0, K(I - B)^\top)$ because

$$E(E) = E((I - B)(Z - \mu)) = (I - B)E(Z - \mu) = 0$$

$$\text{Var}(E) = \text{Var}((I - B)Z) = (I - B)\text{Var}(Z)(I - B)^\top = (I - B)(I - B)^{-1}K(I - B)^\top$$

3.2. Simultaneous Autoregressive (SAR) models.

3.2.1. Gaussian SAR.

Note 32. CAR sets the AR relation, and specifies the distribution on Z which induces the distribution on E ; see 3.4. SAR does the reverse; sets the same AR relation but it specifies the distribution on E which induces the distribution on Z —this is more might be more intuitive (?).

Definition 33. Consider discrete set of sites $\mathcal{S} = \{s_i; i = 1, \dots, n\}$. Consider a random field $Z = (Z_s; s \in \mathcal{S})^\top = (Z_i = Z(s_i); i = 1, \dots, n)^\top$ on the discrete set of indexes \mathcal{S} with values in \mathcal{Z} . Define

$$Z = \mu + \tilde{B}(Z - \mu) + E \iff E = (I - \tilde{B})(Z - \mu)$$

Assume that matrix \tilde{B} is such that $(I - \tilde{B})^{-1}$ exists, and $[\tilde{B}]_{i,i} = 0$. Assume that $E = (E_i; i = 1, \dots, n)$ is an n -dimensional Gaussian random vector $E \sim N_n(0, \Lambda)$ with $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ whose elements are indexed by \mathcal{S} . Then we say that Z follows a Gaussian Simultaneous Autoregressive (SAR) model.

Proposition 34. *The joint distribution of Z following the SAR model in Def 33 is*

$$(3.5) \quad Z \sim N\left(\mu, (I - \tilde{B})^{-1} \Lambda (I - \tilde{B}^\top)^{-1}\right)$$

Proof. Z is a linear combination of Gaussians, hence it follows a Gaussian distribution. Its mean and variance are

$$E(Z) = E\left((I - \tilde{B})^{-1} E + \mu\right) = \mu,$$

$$\text{Var}(Z) = \text{Var}\left((I - \tilde{B})^{-1} E + \mu\right) = (I - \tilde{B})^{-1} \text{Var}(E) (I - \tilde{B}^\top)^{-1} = (I - \tilde{B})^{-1} \Lambda (I - \tilde{B}^\top)^{-1}$$

□

3.3. CAR vs SAR.

Remark 35. From (3.2) and (3.5), CAR and SAR are equivalent iff

$$\underbrace{(I - B)^{-1} K}_{\text{CAR}} = \underbrace{(I - \tilde{B})^{-1} \Lambda (I - \tilde{B}^\top)^{-1}}_{\text{SAR}}$$

Proposition 36. *Any SAR can be written as a CAR model.*

Proof. Let Λ be $n \times n$ positive diagonal matrix. Let \tilde{B} be $n \times n$ positive matrix where $I - \tilde{B}$ is non-singular and $\tilde{B}_{i,i} := [\tilde{B}]_{i,i} = 0$. Then $(I - \tilde{B})^{-1} \Lambda (I - \tilde{B}^\top)^{-1}$ is well defined and I need to solve wrt B and $K = \text{diag}(\kappa_1, \dots, \kappa_n)$

$$\begin{aligned} (I - B)^{-1} K &= (I - \tilde{B})^{-1} \Lambda (I - \tilde{B}^\top)^{-1} \Leftrightarrow \\ K^{-1} (I - B) &= (I - \tilde{B}^\top) \Lambda^{-1} (I - \tilde{B}) \Leftrightarrow \\ K^{-1} - K^{-1} B &= \Lambda^{-1} - \tilde{B}^\top \Lambda^{-1} - \Lambda^{-1} \tilde{B} + \tilde{B}^\top \Lambda^{-1} \tilde{B} \end{aligned}$$

If I focus of the diagonal part and set $B_{i,i} := [B]_{i,i} = 0$

$$[K^{-1}]_{i,i} - \cancel{[K^{-1}B]_{i,i}} \stackrel{=0}{=} [\Lambda^{-1}]_{i,i} - \cancel{[\tilde{B}^\top \Lambda^{-1}]_{i,i}} \stackrel{=0}{=} \cancel{[\Lambda^{-1} \tilde{B}]_{i,i}} \stackrel{=0}{=} [\tilde{B}^\top \Lambda^{-1} \tilde{B}]_{i,i}$$

so

$$\kappa_i = \left(\frac{1}{\lambda_i} + \sum_{j=1}^n \frac{\tilde{B}_{j,i}^2}{\lambda_j} \right)^{-1} > 0, \quad \forall i = 1, \dots, n$$

and hence I can solve with respect to K and B in a manner that they satisfy the assumptions of CAR. \square

Remark 37. The converse of Prop 36 is not true.

Proposition 38. *Any positive-definite covariance matrix Σ can be expressed as the covariance matrix of a CAR model $\Sigma_{\text{CAR}} = (I - B)^{-1} K$, for a unique pair of matrices B and K where $(I - B)$ is non-singular and K is diagonal.*

Proof. Express

$$\Sigma^{-1} = D - R$$

for

$$[D]_{i,j} = \begin{cases} [\Sigma^{-1}]_{i,i} & i = j \\ 0 & i \neq j \end{cases}, \text{ and } [R]_{i,j} = \begin{cases} 0 & i = j \\ -[\Sigma^{-1}]_{i,j} & i \neq j \end{cases}$$

then

$$\Sigma = (D - R)^{-1} = (D (I - D^{-1} R))^{-1} = (I - D^{-1} R)^{-1} D^{-1}$$

Now define $B = D^{-1}R$ and $K = D^{-1}$, and you get $\Sigma = \Sigma_{\text{CAR}}$. Now regarding the uniqueness, assume there is another pair of \mathring{B} , and \mathring{K} such that $\Sigma_{\text{CAR}} = \left(I - \mathring{B}\right)^{-1} \mathring{K}$. Then

$$\text{diag}(\Sigma^{-1}) = \text{diag}(\Sigma_{\text{CAR}}^{-1}) = \text{diag}\left(\mathring{K}^{-1} \left(I - \mathring{B}\right)\right) = \text{diag}\left(\mathring{K}^{-1}\right)$$

and similarly $\text{diag}(\Sigma^{-1}) = \text{diag}(K^{-1})$. Hence it has to be $\mathring{K} = K$ because both are diagonal matrices. Then it is

$$\left(I - \mathring{B}\right)^{-1} \mathring{K} = (I - B)^{-1} K \stackrel{\mathring{K}=K}{\Longleftrightarrow} \mathring{B} = B.$$

So the representation is unique. \square

Proposition 39. *Any positive-definite covariance matrix Σ can be expressed as the covariance matrix of a SAR model $\Sigma_{\text{SAR}} = \left(I - \tilde{B}\right)^{-1} \Lambda \left(I - \tilde{B}^\top\right)^{-1}$ for a (non-unique) pair of matrices \tilde{B} and Λ where $\left(I - \tilde{B}\right)$ is non-singular, $\left[\tilde{B}\right]_{i,i} = 0$, and Λ is diagonal.*

Proof. Express

$$\Sigma^{-1} = LL^\top$$

where L is a lower triangular matrix with $[L]_{i,i} > 0$. Such matrix decomposition can be done by Cholesky decomposition, square-matrix decomposition, etc... and hence it is not always unique. Then

$$\Sigma = (LL^\top)^{-1} = L^{-\top} L^{-1}$$

Now express, $L = D - C$ for

$$[D]_{i,j} = \begin{cases} [L]_{i,i} & i = j \\ 0 & i \neq j \end{cases}, \text{ and } [C]_{i,j} = \begin{cases} 0 & i = j \\ -[L]_{i,j} & i \neq j \end{cases}$$

then

$$\begin{aligned} \Sigma &= (D - C)^{-\top} (D - C)^{-1} = (I - D^{-1}C)^{-\top} D^{-\top} D^{-1} (I - D^{-1}C)^{-1} \\ &= (I - C^\top D^{-\top})^{-1} D^{-\top} D^{-1} \left(I - (C^\top D^{-\top})^\top\right)^{-1} \end{aligned}$$

Set $\tilde{B} = C^\top D^{-\top}$ and $\Lambda = D^{-\top} D^{-1}$ and you get $\Sigma_{\text{SAR}} = \Sigma$ for non-unique pairs of \tilde{B} and Λ . \square

Proposition 40. *Any SAR model can be written as a unique CAR model.*

Proof. SAR and CAR are both Gaussian's with the same mean. SAR's variance matrix is positive definite, and hence it can be written in a unique manner as a CAR's variance matrix by Prop 38. \square

Proposition 41. *Any CAR model can be written as a non-unique SAR model.*

Proof. SAR and CAR are both Gaussian's with the same mean. CAR's variance matrix is positive definite, and hence it can be written in a non-unique manner as a SAR's variance matrix by Prop 39. \square

Example 42. Show that

- (1) Z_i and E_j are independent for $i \neq j$ in Gaussian CAR
- (2) Z_i and E_j are not necessarily independent for $i \neq j$ in Gaussian SAR

Solution.

- (1) For Gaussian CAR,

$$\text{Cov}(E, Z) = \text{Cov}((I - B)Z, Z) = (I - B)\text{Var}(Z) = (I - B)(I - B)^{-1}K = K$$

which is a diagonal; hence Z_i and E_j are independent for $i \neq j$.

- (2) For Gaussian SAR,

$$\text{Cov}(Z, E) = \text{Cov}\left(\left(I - \tilde{B}\right)^{-1}E, E\right) = \left(I - \tilde{B}\right)^{-1}\text{Var}(E) = \left(I - \tilde{B}\right)^{-1}\Lambda$$

which is not a diagonal matrix in general; hence Z_i and E_j may be dependent for $i \neq j$.

4. RELATED RANDOM FIELD STRUCTURES

Note 43. We introduce general modeling structures of basic models which are computationally convenient, yet reasonable for use in spatial statistics models. Convenient because they aim to break a high-dimensional problem into smaller ones using conditional independence, and reasonable because they allow representation of spatial dependence as well. We introduce the Gibbs Random Fields and the Markov Random Fields. The Ising model, CAR, and SAR are just particular cases of such models.