Chapter 4.2 - Lattice processes
Lecture notes

Odd Kolbjørnsen Geir Storvik

February 20, 2017

Last time

- Software
  - Spatial statistics
  - R package
  - INLA
- Multivariate processes
- Directional empirical variogram (exercise)

Today

- Precision matrices
- Lattice process
- AR(1) as lattice process
- MRF – Markov random field
- Neighborhood
- Clique
- Undirected graph (MRF)
- Gaussian CAR – Conditional auto regressive (intro)
- Besag's lemma (conditional vs joint distribution)

Next time

- From where we stop today
- Negpotential function
- Hammersley- Clifford theorem
- Gaussian CAR

Conditional distribution and precision matrices

Assume

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Then if $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$\mathbf{x}_1 | \mathbf{x}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$$

Conditional distribution and precision matrices

Assume

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Then if $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$\mathbf{x}_1 | \mathbf{x}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$$

Define $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$, *precision matrix*. From linear algebra:

$$\mathbf{Q} = \begin{pmatrix} [\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}]^{-1} & -\mathbf{Q}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{Q}_{11} & [\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}]^{-1} \end{pmatrix}$$

Conditional distribution and precision matrices

Assume

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Then if $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$\mathbf{x}_1|\mathbf{x}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$$

Define $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$, *precision matrix*. From linear algebra:

$$\mathbf{Q} = \begin{pmatrix} [\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}]^{-1} & -\mathbf{Q}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{Q}_{11} & [\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}]^{-1} \end{pmatrix}$$

Giving

$$\begin{aligned} \text{var}[\mathbf{x}_1|\mathbf{x}_2] =& \mathbf{Q}_{11}^{-1} \\ E[\mathbf{x}_1|\mathbf{x}_2] =& \boldsymbol{\mu}_1 - \boldsymbol{Q}_{11}^{-1}\mathbf{Q}_{12}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

Conditional distribution and precision matrices

Assume

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Then if $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$\mathbf{x}_1 | \mathbf{x}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$$

Define $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$, *precision matrix*. From linear algebra:

$$\mathbf{Q} = \begin{pmatrix} [\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}]^{-1} & -\mathbf{Q}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{Q}_{11} & [\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}]^{-1} \end{pmatrix}$$

Giving

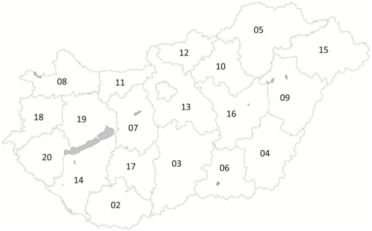
$$\begin{aligned} \text{var}[\mathbf{x}_1|\mathbf{x}_2] =& \mathbf{Q}_{11}^{-1} \\ E[\mathbf{x}_1|\mathbf{x}_2] =& \boldsymbol{\mu}_1 - \boldsymbol{Q}_{11}^{-1}\mathbf{Q}_{12}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

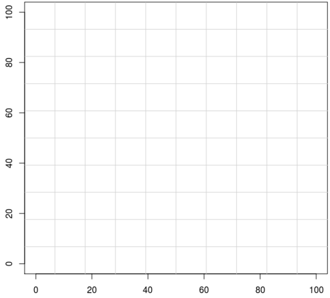
$\mathbf{x}_1 = Y(s), \mathbf{x}_2 = \mathbf{Y}(-s)$:

$$\begin{aligned} \text{var}[Y(s)|\mathbf{Y}(-s)] =& Q_{ss}^{-1} \\ E[Y(s)|\mathbf{Y}(-s)] =& \mu(s) - Q_{ss}^{-1}\mathbf{Q}_{s,-s}(\mathbf{Y}(-s) - \boldsymbol{\mu}(-s)) \end{aligned}$$

Lattice processes

- $\{Y(\mathbf{s}) : \mathbf{s} \in D_s\}$, D_s finite (or countable)
- Typically $\mathbf{s}$ now correspond to an area





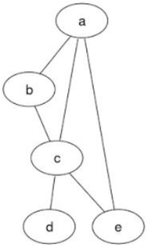
- important feature Neighboring "elements"

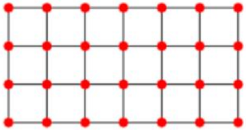
Lattice processes

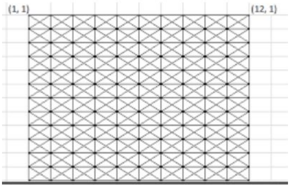
- Let

$$W_{ij} = \begin{cases} 1 & \text{if } \mathbf{s}_i \text{ and } \mathbf{s}_j \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases}$$

- Aim: Construct model based on $\{W_{ij}\}$. (Markov models)
- Neighbors are illustrated in undirected graphs







Markov model in one dimension

$$
\begin{aligned}
Y(1) \sim& N(0, \sigma^2/(1-\phi^2)) \\
Y(s) =& \phi Y(s-1) + \varepsilon(s), \qquad\qquad \varepsilon(s) \sim N(0, \sigma^2)
\end{aligned}
$$

We have

$$
\text{var}(\mathbf{Y}) = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \cdots & \phi^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \cdots & 1 \end{pmatrix}
$$

One-dim AR-process

Can show

$$[\text{var}(\mathbf{Y})]^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1 \end{pmatrix}$$

One-dim AR-process

Can show

$$[\text{var}(\mathbf{Y})]^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1 \end{pmatrix}$$

$$\begin{aligned}
\text{var}[Y(s)|\mathbf{Y}_{-s}] =& Q_{ss}^{-1} = \frac{\sigma^2}{1+\phi^2}, i = 2, \dots, n-1 \\
E[Y(s)|\mathbf{Y}_{-s}] =& \mu(s) - Q_{ss}^{-1}\mathbf{Q}_{s,-s}[\mathbf{Y}(-s) - \boldsymbol{\mu}(-s)] \\
=& -\frac{\sigma^2}{1+\phi^2}[0, 0, \dots, 0, \frac{-\phi}{\sigma^2}, \frac{-\phi}{\sigma^2}, 0, \dots, 0]\begin{pmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(s-2) \\ Y(s-1) \\ Y(s+1) \\ Y(s+2) \\ \vdots \\ Y(m) \end{pmatrix} \\
=& \frac{\phi}{1+\phi^2}(Y(s-1) + Y(s+1)),
\end{aligned}$$

One-dim AR-process

Can show

$$[\text{var}(\mathbf{Y})]^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1 \end{pmatrix}$$

$$\begin{aligned}
\text{var}[Y(s)|\mathbf{Y}_{-s}] =& Q_{ss}^{-1} = \frac{\sigma^2}{1+\phi^2}, i = 2, \ldots, n-1 \\
E[Y(s)|\mathbf{Y}_{-s}] =& \mu(s) - Q_{ss}^{-1}\mathbf{Q}_{s,-s}[\mathbf{Y}(-s) - \boldsymbol{\mu}(-s)] \\
=& -\frac{\sigma^2}{1+\phi^2}[0, 0, \ldots, 0, \frac{-\phi}{\sigma^2}, \frac{-\phi}{\sigma^2}, 0, \ldots, 0]\begin{pmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(s-2) \\ Y(s-1) \\ Y(s+1) \\ Y(s+2) \\ \vdots \\ Y(m) \end{pmatrix} \\
=& \frac{\phi}{1+\phi^2}(Y(s-1) + Y(s+1)),
\end{aligned}$$

Note: Distribution only depend on *neighbors* $\{s-1, s+1\}$.

AR(1) - properties

- $Y(s) = \phi Y(s-1) + \varepsilon(s)$
- Conditional distribution only depend on *neighbors* $\{s-1, s+1\}$:

$$
\begin{aligned}
E[Y(s)|\mathbf{Y}_{-s}] =& \frac{\phi}{1+\phi^2}(Y(s-1) + Y(s+1)) \\
\text{var}[Y(s)|\mathbf{Y}_{-s}] =& \frac{\sigma^2}{1+\phi^2}, i = 2, ..., n-1
\end{aligned}
$$

AR(1) - properties

- $Y(s) = \phi Y(s-1) + \varepsilon(s)$
- Conditional distribution only depend on *neighbors* $\{s-1, s+1\}$:

$$
\begin{aligned}
E[Y(s)|\mathbf{Y}_{-s}] =& \frac{\phi}{1+\phi^2}(Y(s-1) + Y(s+1)) \\
\text{var}[Y(s)|\mathbf{Y}_{-s}] =& \frac{\sigma^2}{1+\phi^2}, i = 2, ..., n-1
\end{aligned}
$$

- Precision matrix is *sparse*

Various attempts to generalize to space.

Non-Gaussian AR

$$
\begin{aligned}
p(\mathbf{Y}) =& p(Y_1)\prod_{s=2}^{n} p(Y_s|Y_{s-1})\\
=& \prod_{s=2}^{n} P_s(Y_s, Y_{s-1}) \;\; (\text{ assuming } > 0))\\
=& \exp\left\{\sum_{s=2}^{n} \log P_s(Y_s, Y_{s-1})\right\}\\
=& \exp\left\{\sum_{s=2}^{n} \psi_s(Y_s, Y_{s-1})\right\}
\end{aligned}
$$

Non-Gaussian AR

$$\begin{aligned}
p(\mathbf{Y}) =& p(Y_1)\prod_{s=2}^{n} p(Y_s|Y_{s-1})\\
=&\prod_{s=2}^{n} P_s(Y_s, Y_{s-1}) \ \ (\text{ assuming } \ > 0))\\
=&\exp\left\{\sum_{s=2}^{n}\log P_s(Y_s, Y_{s-1})\right\}\\
=&\exp\left\{\sum_{s=2}^{n}\psi_s(Y_s, Y_{s-1})\right\}
\end{aligned}$$

Conditional distribution only depend on *neighbors* $\{s-1, s+1\}$:

$$\Rightarrow p(Y_s|\mathbf{Y}_{-s}) \propto \exp\left\{\sum_{i=s}^{s+1}\psi_s(Y_s, Y_{s-1})\right\}$$

Note: a function on this form does not automatically give a legal distribution

From 1D to nD

Learning from 1D

- The Markov model can be defined according to an undirected graph (just using neighbors)
- Not all constellations are legal
- Edge effects must be accounted for $p(Y_1|Y_2)$ and $p(Y_n|Y_{n-1})$
- The Gaussian model provides a sparse precision matrix

An approach in nD (e.g. 2D, 3D,...) need

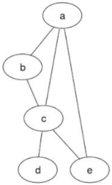
- A definition of neighbors
- Legal joint definitions
- resolve edge effects (often treated ad hoc)

Neighborhood

On a countable set of indexes, i=1,2,...
(corresponding to the random variables on the lattice $Y_i,\ i = 1, 2, \ldots$)

- i is not a neighbor to itself $i \notin \mathcal{N}(i)$
- if j is a neighbor of i then i is a neighbor of j , $j \in \mathcal{N}(i) \Leftrightarrow i \in \mathcal{N}(j)$

A neighborhood system defines an undirected graph, where there is a line form each (node) to each of its neighbors.



The maximum number of edges for a finite set of n nodes are $\frac{n(n-1)}{2}$.
often it is kn, with $k << \frac{n-1}{2}$

Cliques (to be used later, I will remind you)

- Neighbors: $[Y(\mathbf{s}_i)|\mathbf{Y}_{-i}] = [Y(\mathbf{s}_i)|\mathbf{Y}(\mathcal{N}(\mathbf{s}_i))]$.
- Clique: A set of locations that consists either of a single site or a set of sites that are *all* neighbors to each other
- Example: $n = 3, \mathcal{N}(\mathbf{s}_1) = \{\mathbf{s}_2, \mathbf{s}_3\}, \mathcal{N}(\mathbf{s}_2) = \{\mathbf{s}_1\}, \mathcal{N}(\mathbf{s}_3) = \{\mathbf{s}_1\}$,
  Cliques: $\{\mathbf{s}_1\}, \{\mathbf{s}_2\}, \{\mathbf{s}_3\}, \{\mathbf{s}_1, \mathbf{s}_2\}, \{\mathbf{s}_1, \mathbf{s}_3\}$

Gaussian CAR models in space, sec 4.2.1

- Define $\mathcal{N}(\mathbf{s}_i) = \{j; W_{ij} = 1\}$, neighbor set of site i.
- Assume conditional Gaussian distributions with

$$E[Y(s_i)|\mathbf{Y}_{-i}] = \sum_{j \in \mathcal{N}(\mathbf{s}_i)} c_{ij} Y(\mathbf{s}_j)$$
$$\text{var}[Y(s_i)|\mathbf{Y}_{-i}] = \sigma_i^2$$

- Note: Not obvious that this gives a *legal* simultaneous distribution. But assume it does what does that imply? (we know that this is ok for some selections of c_{ij})

If CAR legal

$$E[Y(s_i)|\mathbf{Y}_{-i}] = \sum_j c_{ij} Y(\mathbf{s}_j) \qquad \text{var}[Y(s_i)|\mathbf{Y}_{-i}] = \sigma_i^2$$

$(c_{ij} = 0, j \notin \mathcal{N}(\mathbf{s}_i))$

If CAR legal

$$E[Y(s_i)|\mathbf{Y}_{-i}] = \sum_j c_{ij} Y(\mathbf{s}_j) \qquad \text{var}[Y(s_i)|\mathbf{Y}_{-i}] = \sigma_i^2$$

$(c_{ij} = 0, j \notin \mathcal{N}(\mathbf{s}_i))$
Also

$$E[Y(s_i)|\mathbf{Y}(-s_i)] = -Q_{ii}^{-1}\mathbf{Q}_{i,-i}\mathbf{Y}(-s_i) \quad \text{var}[Y(s_i)|\mathbf{Y}(-s_i)] = Q_{ii}^{-1}$$

If CAR legal

$$E[Y(s_i)|\mathbf{Y}_{-i}] = \sum_j c_{ij} Y(\mathbf{s}_j) \qquad \text{var}[Y(s_i)|\mathbf{Y}_{-i}] = \sigma_i^2$$

$(c_{ij} = 0, j \notin \mathcal{N}(\mathbf{s}_i))$
Also

$$E[Y(s_i)|\mathbf{Y}(-s_i)] = -\, Q_{ii}^{-1}\mathbf{Q}_{i,-i}\mathbf{Y}(-s_i) \quad \text{var}[Y(s_i)|\mathbf{Y}(-s_i)] = Q_{ii}^{-1}$$

Implies (if legal model)

$$\begin{aligned} Q_{ii} =& \sigma_i^{-2} \\ -Q_{ii}^{-1}\mathbf{Q}_{i,-i}\mathbf{Y}(-s_i) =& \sum_{j \in \mathcal{N}(\mathbf{s}_i)} c_{ij} Y(\mathbf{s}_j) && j \neq i \\ -Q_{ii}^{-1} Q_{ij} =& c_{ij} && j \neq i \\ Q_{ij} =& -\, \sigma_i^{-2} c_{ij}, && j \neq i \\ \mathbf{Q} =& \mathbf{M}^{-1}[\mathbf{I} - \mathbf{C}] && \mathbf{M} = \text{diag}\{\sigma_i^2\} \end{aligned}$$

If CAR legal

$$E[Y(s_i)|\mathbf{Y}_{-i}] = \sum_j c_{ij} Y(\mathbf{s}_j) \qquad \text{var}[Y(s_i)|\mathbf{Y}_{-i}] = \sigma_i^2$$

$(c_{ij} = 0, j \notin \mathcal{N}(\mathbf{s}_i))$
Also

$$E[Y(s_i)|\mathbf{Y}(-s_i)] = -Q_{ii}^{-1}\mathbf{Q}_{i,-i}\mathbf{Y}(-s_i) \quad \text{var}[Y(s_i)|\mathbf{Y}(-s_i)] = Q_{ii}^{-1}$$

Implies (if legal model)

$$\begin{aligned} Q_{ii} =& \sigma_i^{-2} & \\ -Q_{ii}^{-1}\mathbf{Q}_{i,-i}\mathbf{Y}(-s_i) =& \sum_{j \in \mathcal{N}(\mathbf{s}_i)} c_{ij} Y(\mathbf{s}_j) & j \neq i \\ -Q_{ii}^{-1} Q_{ij} =& c_{ij} & j \neq i \\ Q_{ij} =& -\sigma_i^{-2} c_{ij}, & j \neq i \\ \mathbf{Q} =& \mathbf{M}^{-1}[\mathbf{I} - \mathbf{C}] & \mathbf{M} = \text{diag}\{\sigma_i^2\} \end{aligned}$$

Conditions on CAR models

- If legal, $\mathbf{Q} = \mathbf{M}^{-1}[\mathbf{I} - \mathbf{C}]$
- If $\mathbf{M}^{-1}(\mathbf{I} - \mathbf{C})$ is symmetric and positive definite, then
$$\mathbf{Y} \sim \text{MVN}(\mathbf{0}, (\mathbf{I} - \mathbf{C})^{-1}\mathbf{M})$$
- $\mathbf{Q} = \mathbf{M}^{-1}(\mathbf{I} - \mathbf{C})$ is sparse if $\mathbf{C}$ is sparse!
- $\mathbf{M}$ always positive definite and symmetric
- $j \in \mathcal{N}(\mathbf{s}_i) \Leftrightarrow i \in \mathcal{N}(\mathbf{s}_j)$ imply $c_{ij} = 0 \Leftrightarrow c_{ji} = 0$
- Need $\sigma_i^{-2} c_{ij} = \sigma_j^{-2} c_{ji}$ for neighbors
- Also, the c_{ij} can not be too large for getting positive definiteness.

CAR - extensions

- $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta}$, $\boldsymbol{\delta}$ a CAR process
- Hierarchical structure,

$$Z(\mathbf{s}_i)|Y(\mathbf{s}_i) \overset{ind}{\sim} \text{Poisson}(\exp(Y\{\mathbf{s}_i)\})$$

CAR versus Geostatistical modelling

Data on grid: Two possible models

- Geostatistical modelling with distances between centerpoints
- CAR model where neighbors if sharing border

Computational benefit:

- $\mathbf{Q}$ for CAR is sparse, much faster computation
  (Example R code `sparseMatrix2D.R` on webpage)

Markov Random Field (MRF, sec 4.2.2)

- CAR model:
  - Gaussian process
  - $Y(\mathbf{s}_i)$ only depend on $\{Y(\mathbf{s}_j); j \in \mathcal{N}(\mathbf{s}_i)\}$
  - Called a Markov-type model

Markov Random Field (MRF, sec 4.2.2)

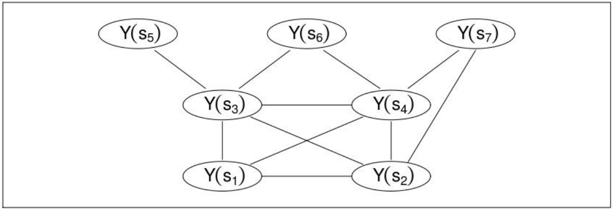
- CAR model:
  - Gaussian process
  - $Y(\mathbf{s}_i)$ only depend on $\{Y(\mathbf{s}_j); j \in \mathcal{N}(\mathbf{s}_i)\}$
  - Called a Markov-type model
- General Markov Random Field models:

$$
\begin{aligned}
[Y(\mathbf{s}_i)|\mathbf{Y}_{-i}] =& [Y(\mathbf{s}_i)|\mathbf{Y}(\mathcal{N}(\mathbf{s}_i))] \\
\mathbf{Y}(\mathcal{N}(\mathbf{s}_i)) =& \{Y(\mathbf{s}_j); j \in \mathcal{N}(\mathbf{s}_i)\}
\end{aligned}
$$

- Example: Auto Poisson model ($\theta_{ij} = \theta_{ji}$)

$$
[Y(\mathbf{s}_i)|\mathbf{Y}(\mathcal{N}(\mathbf{s}_i))] = \text{Poisson}\left(\exp\left\{\alpha_i + \sum_{j \in \mathcal{N}(\mathbf{s}_i)} \theta_{ij} Y(\mathbf{s}_j)\right\}\right)
$$

Displaying models by graphs



- *Undirected graph*
- Edge between $\mathbf{s}_i$ and $\mathbf{s}_j$ means that $\mathbf{s}_j \in \mathcal{N}(\mathbf{s}_i)$ (and $\mathbf{s}_i \in \mathcal{N}(\mathbf{s}_j)$)
- Shows conditional *independence*

$$[Y(\mathbf{s}_4)|\mathbf{Y}_{-s_4}] = [Y(\mathbf{s}_4)|Y(\mathbf{s}_1), Y(\mathbf{s}_2), Y(\mathbf{s}_3), Y(\mathbf{s}_6), Y(\mathbf{s}_7))]$$

- Do not show *type* and *strength* of dependence

Markov Random fields (MRF)

- Called an MRF if model *defined* through

$$
\begin{aligned}
[Y(\mathbf{s}_j)|\mathbf{Y}_{-i}] =& [Y(\mathbf{s}_j)|\mathbf{Y}(\mathcal{N}(\mathbf{s}_i))] \\
\mathbf{Y}(\mathcal{N}(\mathbf{s}_i)) =& \{Y(\mathbf{s}_j); j \in \mathcal{N}(\mathbf{s}_i)\}
\end{aligned}
$$

- Specifications needs to be made consistently so that a *simultaneous* model exist
- Much used in e.g.
  - image analysis (regular grids)
  - disease mapping (irregular grids, counties)

Requirements

- $\mathbf{y} \equiv (y(\mathbf{s}_1), ..., y(\mathbf{s}_n))' \in \mathcal{R}^n$
- $\Omega \equiv \{\mathbf{y} : \Pr(\mathbf{y}) > 0\}$, $\Omega_i \equiv \{y(\mathbf{s}_i) : \Pr(y(\mathbf{s}_i)) > 0\}$

Requirements

- $\mathbf{y} \equiv (y(\mathbf{s}_1), ..., y(\mathbf{s}_n))' \in \mathcal{R}^n$
- $\Omega \equiv \{\mathbf{y} : \Pr(\mathbf{y}) > 0\}$, $\Omega_i \equiv \{y(\mathbf{s}_i) : \Pr(y(\mathbf{s}_i)) > 0\}$
- Positivity condition: $\Omega = \Omega_1 \times \cdots \times \Omega_n$.
  (Not essential, but simplifies results)

Requirements

- $\mathbf{y} \equiv (y(\mathbf{s}_1), ..., y(\mathbf{s}_n))' \in \mathcal{R}^n$
- $\Omega \equiv \{\mathbf{y} : \Pr(\mathbf{y}) > 0\}$, $\Omega_i \equiv \{y(\mathbf{s}_i) : \Pr(y(\mathbf{s}_i)) > 0\}$
- Positivity condition: $\Omega = \Omega_1 \times \cdots \times \Omega_n$.
  (Not essential, but simplifies results)
- Besags Lemma (Brooks Lemma):

$$\frac{\Pr(\mathbf{y})}{\Pr(\mathbf{w})} = \prod_{i=1}^{n} \frac{\Pr(y(\mathbf{s}_i)|y(\mathbf{s}_1), ..., y(\mathbf{s}_{i-1}), w(\mathbf{s}_{i+1}), ..., w(\mathbf{s}_n))}{\Pr(w(\mathbf{s}_i)|y(\mathbf{s}_1), ..., y(\mathbf{s}_{i-1}), w(\mathbf{s}_{i+1}), ..., w(\mathbf{s}_n))}$$

  Gives relationship between conditional and simultaneous distributions (if simultaneous distribution exist)
- Typically $\mathbf{w}$ fixed (e.g $\mathbf{w} = \mathbf{0}$), gives

$$\Pr(\mathbf{y}) \propto \prod_{i=1}^{n} \frac{\Pr(y(\mathbf{s}_i)|y(\mathbf{s}_1), ..., y(\mathbf{s}_{i-1}), w(\mathbf{s}_{i+1}), ..., w(\mathbf{s}_n))}{\Pr(w(\mathbf{s}_i)|y(\mathbf{s}_1), ..., y(\mathbf{s}_{i-1}), w(\mathbf{s}_{i+1}), ..., w(\mathbf{s}_n))}$$

Proof Besag's lemma

Note (using $y_i = y(\mathbf{s}_i)$ for simplification)

$$\frac{\Pr(y_m|y_1,...,y_{m-1})}{\Pr(w_m|y_1,...,y_{m-1})} = \frac{\frac{\Pr(y_1,...,y_{m-1},y_m)}{\Pr(y_1,...,y_{m-1})}}{\frac{\Pr(y_1,...,y_{m-1},w_m)}{\Pr(y_1,...,y_{m-1})}} = \frac{\Pr(y_1,...,y_{m-1},y_m)}{\Pr(y_1,...,y_{m-1},w_m)}$$

so

$$\Pr(y_1,...,y_m) = \frac{\Pr(y_m|y_1,...,y_{m-1})}{\Pr(w_m|y_1,...,y_{m-1})}\Pr(y_1,...,y_{m-1},w_m)$$

repeating argument

$$= \frac{\Pr(y_m|y_1,...,y_{m-1})}{\Pr(w_m|y_1,...,y_{m-1})}\frac{\Pr(y_{m-1}|y_1,...,y_{m-2},w_m)}{\Pr(w_{m-1}|y_1,...,y_{m-2},w_m)}\Pr(y_1,...,y_{m-2},w_{m-1},w_m)$$

$$\vdots$$

$$\prod_{i=1}^{n}\frac{\Pr(y_i|y_1,...,y_{i-1},w_{i+1},...,w_n)}{\Pr(w_i|y_1,...,y_{i-1},w_{i+1},...,w_n)}\Pr(w_1,...,w_{m-2},w_{m-1},w_m)$$