

Chapter 4.2 - Lattice processes

Lecture notes

Odd Kolbjørnsen Geir Storvik

February 27, 2017

Last Time

- 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)

Today

- Lattice model and other models
- Neighborhood
- Clique
- Negpotential function
- Hammersley- Clifford theorem
- Auto spatial models
 - CAR model
 - Latent Gaussian process
 - Auto logistic model (Ising model)
 - auto Poisson model

Next time:

- How to construct CAR from scratch
- Examples of Models in INLA (disease mapping, etc)
- Examples of models outside INLA (Potts model)

Spatial processes

Smooth model

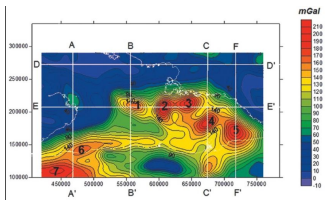
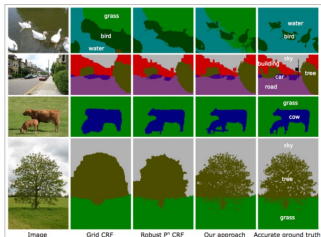
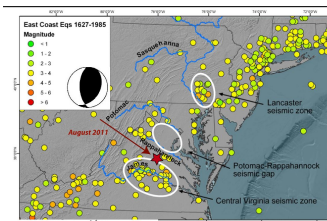


Figure 3. Complete Bouguer anomaly map for Southeastern Cuba island. Gravity anomaly highs are named: 1 Levingston, 2 La Guira, 3 Pilo, 4 El Salvador, 5 La Perrera, 6 Eje Magnético Sur, 7 New oceanic crust coming from the Cayman dispersion center. Capital letters and lines indicate the six 2D cross-sections made to the 3D density model.

Discrete Model (segmentation)



Event observations

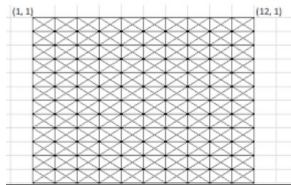
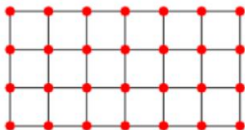
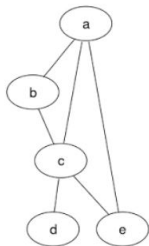


- Process model (What is modeled?)
 - Smooth phenomena (e.g. Gaussian models)
 - Discrete phenomena (e.g. Potts model)
 - Event models (e.g. Poisson process)
- Modeling approach (How is it modeled?)
 - Joint distribution (Continuous domain),
e.g. Gaussian random field with a given Covariance function
 - Sequential approach (Directed acyclic graph)
e.g. Gaussian AR(1) model
 - Conditional approach/ Lattice model (Undirected graph)
e.g. Gaussian Markov random fields on a lattice (Precision matrix)

Want to make a statistical model of the process which is 1) consistent, 2) models important features in a good way, 3) we can do inference in.

Lattice processes

- Aim: Construct model based on "local modeling" (Markov models)
- Neighbors are illustrated in undirected graphs



- Motivation:
 - Fast computations on large Gaussian models
 - Model non-Gaussian dependency in spatial models.
- Basic question. Does a joint distribution with the specified conditional distributions exist?

Making a joint distribution from the conditionals

- Last time : Besag's Lemma = First piece of the puzzle. Relation between the joint distribution and the conditional approach. (Assuming the joint distribution exist)

$$\frac{\Pr(\mathbf{y})}{\Pr(\mathbf{w})} = \prod_{i=1}^n \frac{\Pr(y_i | y_1, \dots, y_{i-1}, w_{i+1}, \dots, w_n)}{\Pr(w_i | y_1, \dots, y_{i-1}, w_{i+1}, \dots, w_n)}$$

- Now: Hammersley - Clifford theorem. On the graph (G)
Graph = Nodes (=Lattice) + Edges (=Neighbors)

$$\Pr(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{c(\boldsymbol{\theta})} \exp \{Q(\mathbf{y}; \boldsymbol{\theta})\} = \frac{1}{c(\boldsymbol{\theta})} \exp \left\{ \sum_{c \in C_G} \psi_c(y_c, \boldsymbol{\theta}) \right\}$$

with C_G being the set of all cliques, $Q(\mathbf{0}, \boldsymbol{\theta}) = 0$, and $c(\boldsymbol{\theta})$ is a normalizing constant.

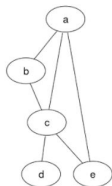
Neighbor/Neighborhood

On a countable set of indexes, $i=1,2,\dots$

(corresponding to the random variables on the lattice Y_i , $i = 1, 2, \dots$)

- 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 from each (node) to each of its neighbors.



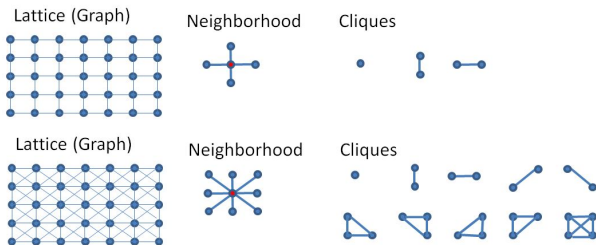
The maximum number of edges for a finite set of n nodes are $\frac{n(n-1)}{2}$.
often it is kn , with $k \ll \frac{n-1}{2}$

Cliques

- Clique: A set of locations (e.g. nodes, points in lattice) 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: $\{\mathbf{s}_1\}, \{\mathbf{s}_2\}, \{\mathbf{s}_3\}, \{\mathbf{s}_1, \mathbf{s}_2\}, \{\mathbf{s}_1, \mathbf{s}_3\}$

Cliques

- Clique: A set of locations (e.g. nodes, points in lattice) 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: $\{\mathbf{s}_1\}, \{\mathbf{s}_2\}, \{\mathbf{s}_3\}, \{\mathbf{s}_1, \mathbf{s}_2\}, \{\mathbf{s}_1, \mathbf{s}_3\}$
- Example: Lattice



Negpotential function

- Assume $P(\mathbf{0}) > 0$ (always possible by defining " $\mathbf{0}$ " appropriately)
- Define *Negpotential function*: $Q(\mathbf{y}) \equiv \log\{\Pr(\mathbf{y}) / \Pr(\mathbf{0})\}$

Negpotential function

- Assume $P(\mathbf{0}) > 0$ (always possible by defining " $\mathbf{0}$ " appropriately)
- Define *Negpotential function*: $Q(\mathbf{y}) \equiv \log\{\Pr(\mathbf{y}) / \Pr(\mathbf{0})\}$
- and expand it as:

$$Q(\mathbf{y}) = \sum_i y_i G_i(y_i) + \sum_{ij} y_i y_j G_{ij}(y_i, y_j) + \\ \sum_{ijk} y_i y_j y_k G_{ijk}(y_i, y_j, y_k) + \cdots + y_1 y_2 \cdots y_m G_{1,\dots,m}(y_1, \dots, y_m)$$

Terms in Negpotential function - expansion

The G -functions are defined/constructed through

$$\begin{aligned} \mathbf{y} &= (0, \dots, 0, y_i, 0, \dots, 0), y_i \neq 0 \\ &\Rightarrow G_i(y_i) = Q(0, \dots, 0, y_i, 0, \dots, 0) / y_i \\ &\Rightarrow Q(0, \dots, 0, y_i, 0, \dots, 0) = y_i G_i(y_i) \end{aligned} \quad (*)$$

$$\begin{aligned} \mathbf{y} &= (0, \dots, 0, y_i, 0, \dots, 0, y_j, 0, \dots, 0), y_i, y_j \neq 0 \\ &\Rightarrow G_{i,j}(y_i, y_j) = [Q(0, \dots, 0, y_i, 0, \dots, 0, y_j, 0, \dots, 0) - y_i G_i(y_i) - y_j G_j(y_j)] / (y_i y_j) \\ &\Rightarrow Q(0, \dots, 0, y_i, 0, \dots, 0, y_j, 0, \dots, 0) = y_i G_i(y_i) + y_j G_j(y_j) + y_i y_j G_{i,j}(y_i, y_j) \end{aligned}$$

⋮

Hammersley-Clifford theorem

$$Q(\mathbf{y}) = \sum_i y_i G_i(y_i) + \sum_{ij} y_i y_j G_{ij}(y_i, y_j) + \\ \sum_{ijk} y_i y_j y_k G_{ijk}(y_i, y_j, y_k) + \cdots + y_1 y_2 \cdots y_m G_{1, \dots, m}(y_1, \dots, y_m)$$

Hammersley-Clifford theorem

$$Q(\mathbf{y}) = \sum_i y_i G_i(y_i) + \sum_{ij} y_i y_j G_{ij}(y_i, y_j) + \sum_{ijk} y_i y_j y_k G_{ijk}(y_i, y_j, y_k) + \cdots + y_1 y_2 \cdots y_m G_{1, \dots, m}(y_1, \dots, y_m)$$

Theorem: If $\{i_1, i_2, \dots, i_k\}$ is not a clique, then $G_{i_1, i_2, \dots, i_k}(\cdot) \equiv 0$

Hammersley-Clifford theorem

$$Q(\mathbf{y}) = \sum_i y_i G_i(y_i) + \sum_{ij} y_i y_j G_{ij}(y_i, y_j) + \sum_{ijk} y_i y_j y_k G_{ijk}(y_i, y_j, y_k) + \cdots + y_1 y_2 \cdots y_m G_{1, \dots, m}(y_1, \dots, y_m)$$

Theorem: If $\{i_1, i_2, \dots, i_k\}$ is not a clique, then $G_{i_1, i_2, \dots, i_k}(\cdot) \equiv 0$

The hard thing here is to get all the definitions straight...

Hammersley-Clifford theorem proof

If a set is not a clique: Then there exist (at least) two indices not being neighbors in this set. Assume these are $i_1 = i$ and $i_2 = l$.

Define $\mathbf{y}_i = (y_1, \dots, y_{i-1}, 0, y_{i+1}, \dots, y_m)$.

$$\frac{\Pr(\mathbf{y})}{\Pr(\mathbf{y}_i)} = \frac{\Pr(y_i | \mathbf{y}_{-i})}{\Pr(0 | \mathbf{y}_{-i})} \quad \text{do not depend on } y_l$$

$$\frac{\Pr(\mathbf{y})}{\Pr(\mathbf{y}_i)} = \exp(Q(\mathbf{y}) - Q(\mathbf{y}_i))$$

so $Q(\mathbf{y}) - Q(\mathbf{y}_i)$ do not depend on y_l , since this was the way we constructed it by the way we picked indices

Proof (cont)

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(\mathbf{y}_i) + \sum_{j \neq i} y_i y_j G_{i,j}(\mathbf{y}_i, y_j) + \sum_{j, k \neq i} y_i y_j y_k G_{i,j,k}(\mathbf{y}_i, y_j, y_k) + \dots$$

- The sum only include terms where y_i is included since the others cancel out in the subtraction

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(y_i) + \sum_{j \neq i} y_i y_j G_{i,j}(y_i, y_j) + \sum_{j, k \neq i} y_i y_j y_k G_{i,j,k}(y_i, y_j, y_k) + \dots$$

- The sum only include terms where y_i is included since the others cancel out in the subtraction
- Putting $y_j = 0, j \neq i, l$ we get:
 - $y_i y_l G_{i,l}(y_i, y_l) = 0$ since this is only term depending on y_l

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(y_i) + \sum_{j \neq i} y_i y_j G_{i,j}(y_i, y_j) + \sum_{j, k \neq i} y_i y_j y_k G_{i,j,k}(y_i, y_j, y_k) + \dots$$

- The sum only include terms where y_i is included since the others cancel out in the subtraction
- Putting $y_j = 0, j \neq i, l$ we get:
 - $y_i y_l G_{i,l}(y_i, y_l) = 0$ since this is only term depending on y_l
 - For $y_l = 0, y_i y_l G_{i,l}(y_i, y_l) = 0$ for any y_i

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(y_i) + \sum_{j \neq i} y_i y_j G_{i,j}(y_i, y_j) + \sum_{j, k \neq i} y_i y_j y_k G_{i,j,k}(y_i, y_j, y_k) + \dots$$

- The sum only include terms where y_i is included since the others cancel out in the subtraction
- Putting $y_j = 0, j \neq i, l$ we get:
 - $y_i y_l G_{i,l}(y_i, y_l) = 0$ since this is only term depending on y_l
 - For $y_l = 0, y_i y_l G_{i,l}(y_i, y_l) = 0$ for any y_i
 - Imply $y_i y_l G_{i,l}(y_i, y_l) = 0$ also for $y_l \neq 0$ and any y_i (since $l \notin \mathcal{N}(i)$)

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(\mathbf{y}_i) + \sum_{j \neq i} y_i y_j G_{i,j}(y_i, y_j) + \sum_{j,k \neq i} y_i y_j y_k G_{i,j,k}(y_i, y_j, y_k) + \dots$$

- The sum only include terms where y_i is included since the others cancel out in the subtraction
- Putting $y_j = 0, j \neq i, l$ we get:
 - $y_i y_l G_{i,l}(y_i, y_l) = 0$ since this is only term depending on y_l
 - For $y_l = 0, y_i y_l G_{i,l}(y_i, y_l) = 0$ for any y_i
 - Imply $y_i y_l G_{i,l}(y_i, y_l) = 0$ also for $y_l \neq 0$ and any y_i (since $l \notin \mathcal{N}(i)$)
 - Imply $G_{i,l}(y_i, y_l) = 0$

$$Q(\mathbf{y}) - Q(\mathbf{y}_i) = y_i G_i(y_i) + \sum_{j \neq i} y_i y_j G_{i,j}(y_i, y_j) + \\ \sum_{j,k \neq i} y_i y_j y_k G_{i,j,k}(y_i, y_j, y_k) + \dots$$

- Showed for all second order effects
- Similarly for triplets: Putting $y_j = 0, j \neq i, k, l$
- Using that $G_{i,l}(y_i, y_l) = 0$ for the non neighboring pair.
- ...
- implies $G_{i,k,l}(y_i, y_k, y_l) = 0$
- Iterating argument expanding to higher order 4th, 5th,... using at all of orders below are zero.
- Imply all G 's containing i, l are zero.

Existence of simultaneous distribution

Besag's lemma, Hammersley-Clifford: Results assuming simultaneous distribution exist.

Assume conditional distributions and

- 1 $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_m$
- 2 The G-functions are invariant under permutation
- 3 $\sum_{\mathbf{t} \in \Omega} \exp\{Q(\mathbf{t})\} < \infty$

Then simultaneous distribution exist.

Condition 1 can be relaxed, see eq (4.116) in the book

- Auto-spatial models: Only second order terms

$$Q(\mathbf{y}) = \sum_i y_i G_i(y_i) + \sum_{ij} y_i y_j G_{ij}(y_i, y_j)$$

- The CAR model
- Latent Gaussian processes
- auto Poisson model

Note: $\{i, j, k\}$ might be a clique but still $G_{ijk}(y_i, y_j, y_k) = 0$.

CAR model, $\mu = 0$

$$\begin{aligned}\log P(\mathbf{y}) &= -\frac{m}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} \\ &= -\frac{m}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{Q}| - \frac{1}{2} \mathbf{y}^T \mathbf{Q} \mathbf{y} \\ &= -\frac{m}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{Q}| - \frac{1}{2} \sum_{ij} y_i Q_{ij} y_j \\ &= -\frac{m}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{Q}| - \frac{1}{2} \sum_i y_i^2 Q_{ii} - \frac{1}{2} \sum_i \sum_{j \in \mathcal{N}_i} y_i Q_{ij} y_j\end{aligned}$$

giving

$$\begin{aligned}\log P(\mathbf{0}) &= -\frac{m}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{Q}| \\ Q(\mathbf{y}) &= -\frac{1}{2} \sum_i y_i^2 Q_{ii} - \frac{1}{2} \sum_i \sum_{j \in \mathcal{N}_i} y_i Q_{ij} y_j \\ G_i(y_i) &= -\frac{1}{2} y_i Q_{ii} \quad G_{ij}(y_i, y_j) = -\frac{1}{2} Q_{ij}, \quad G_{ijk}(y_i, y_j, y_k) = 0\end{aligned}$$

CAR model, $\mu = 0$

$$P(\mathbf{y}) = P(\mathbf{0})e^{\mathbf{Q}(\mathbf{y})}$$

$$\begin{aligned} \mathbf{Q}(\mathbf{y}) &= -\frac{1}{2} \sum_i y_i^2 Q_{ii} - \frac{1}{2} \sum_{i \neq j} y_i Q_{ij} y_j \\ &= -\frac{1}{2} \sum_i y_i^2 Q_{ii} - \frac{1}{2} \sum_i \sum_{j \in \mathcal{N}_i} y_i Q_{ij} y_j \end{aligned}$$

Assume $Q_{ij} = 0$. Then

- $Q_{ji} = 0$
- $i \notin \mathcal{N}_j$ and $j \notin \mathcal{N}_i$
- If \mathcal{N}_i are small for all i , then \mathbf{Q} is **sparse**

Latent Gaussian processes

Assume

$$\mathbf{y} \sim \text{MVN}(\mathbf{0}, \mathbf{Q}^{-1})$$

$$z_i | \mathbf{y} \sim f(z_i | y_i),$$

\mathbf{Q} sparse

Cond. independent

Then

$$p(\mathbf{y} | \mathbf{z}) \propto p(\mathbf{y}) p(\mathbf{z} | \mathbf{y}) = p(\mathbf{y}) \prod_i p(z_i | y_i)$$

$$\log p(\mathbf{y} | \mathbf{z}) = \text{Const} - \frac{m}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{Q}| - \frac{1}{2} \sum_{ij} y_i \mathbf{Q}_{ij} y_j + \sum_i \log f(z_i | y_i)$$

$$Q(\mathbf{y}) = \log p(\mathbf{y} | \mathbf{z}) - \log p(\mathbf{0} | \mathbf{z})$$

$$= -\frac{1}{2} \sum_{ij} y_i \mathbf{Q}_{ij} y_j + \sum_i \log f(z_i | y_i) - \sum_i \log f(z_i | 0)$$

$$= -\frac{1}{2} \sum_i y_i^2 \mathbf{Q}_{ii} + \sum_i \log f(z_i | y_i) - \sum_i \log f(z_i | 0) - \frac{1}{2} \sum_{ij, i \neq j} y_i \mathbf{Q}_{ij} y_j$$

$$y_i G_i(y_i) = -\frac{1}{2} y_i \mathbf{Q}_{ii} y_i + \log f(z_i | y_i) - \log f(z_i | 0)$$

$$G_{ij}(y_i) = -\frac{1}{2} \mathbf{Q}_{ij}$$

Finding G functions and simultaneous distribution

Assume simultaneous distribution exist

For MRF

$$\begin{aligned} P(y_i | \mathbf{y}_{-i}) &= \frac{P(\mathbf{y})}{P(\mathbf{y}_{-i})} \propto P(\mathbf{y}) \propto \exp(Q(\mathbf{y})) \\ &\propto \exp\{y_i G_i(y_i) + y_i \sum_j y_j G_{ij}(y_i, y_j) + \dots\} \end{aligned}$$

Assume conditional distributions are specified.

Put $y_j = 0, j \neq i$ for finding $G_i(y_i)$

Put $y_k = 0, k \neq i, j$ for finding $G_{ij}(y_i, y_j)$

etc

$$P(\mathbf{y}) \propto \exp(Q(\mathbf{y}))$$

For auto-logistic, $y_i \in \{0, 1\}$

$$\Pr(y_i | \mathbf{y}_{-i}) = \frac{\exp\{\alpha_i y_i + \sum_{j \neq i} \theta_{ij} y_i y_j\}}{1 + \exp\{\alpha_i + \sum_{j \neq i} \theta_{ij} y_j\}}$$
$$\propto \exp\{\alpha_i y_i + \sum_{j \neq i} \theta_{ij} y_i y_j\}$$

$$G_i(y_i) = \alpha_i$$

$$G_{ij}(y_i, y_j) = \theta_{ij}$$

$$Q(\mathbf{y}) = \sum_i \alpha_i y_i + \sum_{i,j, i \neq j} y_i y_j \theta_{ij}$$

$$P(\mathbf{y}) \propto \exp\left\{ \sum_i \alpha_i y_i + \sum_{i,j, i \neq j} y_i y_j \theta_{ij} \right\}$$

On regular lattice: *Ising model*

- $i = (u, v), \mathcal{N}_i = \{(u, v - 1), (u, v + 1), (u - 1, v), (u + 1, v)\}$
- $\theta_{ij} = 0$ for non-neighbors

For auto-Poisson (misprint in eq (4.114) in the book)

$$\begin{aligned} P(y_i|y_{-i}) &= \text{Poisson}(\exp\{\alpha_i + \sum_{j=1}^m \theta_{ij}y_j\}) && \theta_{ij} = 0, j \notin \mathcal{N}_i \\ &= \text{Poisson}(\exp\{\alpha_i + \sum_{j \in \mathcal{N}_i} \theta_{ij}y_j\}) \\ &= \exp\{y_i(\alpha_i + \sum_{j \in \mathcal{N}_i} \theta_{ij}y_j) - \log y_i! - e^{\alpha_i + \sum_{j \in \mathcal{N}_i} \theta_{ij}y_j}\} \\ &\propto \exp\{y_i\alpha_i - \log y_i! + y_i \sum_{j \in \mathcal{N}_i} \theta_{ij}y_j\} \end{aligned}$$

$$G_i(y_i) = \alpha_i - y_i^{-1} \log y_i!$$

$$G_{ij}(y_i, y_j) = \theta_{ij}$$

$$Q(\mathbf{y}) = \sum_i (\alpha_i y_i - \log y_i!) + \sum_{i,j} y_i y_j \theta_{ij}$$

Auto models - general

Exponential family:

$$f_Y(y|\theta) = \exp\{(y\tilde{\theta} - a(\theta))/\phi + \log c(y; \phi)\}$$

Auto models - general

Exponential family:

$$f_Y(y|\theta) = \exp\{(y\tilde{\theta} - a(\theta))/\phi + \log c(y; \phi)\}$$

Auto models (a bit more general in the book)

$$[Y_i | \mathbf{Y}_{-i}] \equiv \exp(A_i(\mathbf{Y}_{-i})(Y_i - B_i) + C_i(Y_i) + D_i(\mathbf{Y}_{-i}))$$

$$\tilde{\theta} = A_i(\mathbf{Y}_{-i}) = \alpha_i + \sum_{j \neq i} \theta_{ij}(Y_j - B_j)$$

$$Q(\mathbf{y}) = \sum_i [\alpha_i(y_i - B_i) + C_i(y_i)] + \sum_{i < j} \theta_{ij}(y_i - B_i)(y_j - B_j)$$

- auto Gaussian (CAR)
- auto gamma model
- Winsorized auto Poisson model
- auto beta model

Auto models - general

Exponential family:

$$f_Y(y|\theta) = \exp\{(y\tilde{\theta} - a(\theta))/\phi + \log c(y; \phi)\}$$

Auto models (in the book)

$$[Y_i | \mathbf{Y}_{-i}] \equiv \exp(A_i(\mathbf{Y}_{-i})(B_i(Y_i)) + C_i(Y_i) + D_i(\mathbf{Y}_{-i}))$$

$$\tilde{\theta} = A_i(\mathbf{Y}_{-i}) = \alpha_i + \sum_{j \neq i} \theta_{ij} (B_j(Y_j) - B_j(0))$$

$$Q(\mathbf{y}) = \sum_i [\alpha_i (B_i(y_i) - B_i(0)) + C_i(y_i)] \\ + \sum_{i < j} \theta_{ij} (B_i(y_i) - B_i(0))(B_j(y_j) - B_j(0))$$

- auto Gaussian (CAR)
- auto gamma model
- Winsorized auto Poisson model (truncated at $y_i > t_i$)
- auto beta model