Spatial Statistics with Image Analysis
Lecture 6

Johan Lindström

November 19, 2018

Lecture L06

Computer exercise 2
Estimation
Parameter Estimation
Reconstruction
Gaussian Markov random fields
Sparse Q
Requirements
Selecting q
Construction of Q
The q-structure
CAR and SAR models
CAR(1) models
Boundary effects
Precision matrix
SAR models
What’s left?

Computer exercise 2 — Estimation

![Covariance Cloud](image1)

![Covariance Cloud](image2)

![Binned Covariance Estimates](image3)

![Simulated Field](image4)

![LS-estimated Covariance](image5)

![Observations per Bin](image6)
Computer exercise 2 — Non-parametric (binned) estimate

- Let \( H_k = \{ (i, j) : i \neq j, \, kh \leq \|s_j - s_i\| < (k + 1)h \} \)
- The non-parametric or “binned” covariance estimator is
  
  \[ \hat{\sigma}^2 + \sigma^2_e = \frac{1}{n} \sum_i e_i^2 \]
  
  \[ \hat{r}(kh) = \frac{1}{m_k} \sum_{H_k} e_i e_j \]

Consider the two estimates

\[ \hat{r}(kh) = \frac{1}{m_k} \sum_{H_k} e_i e_j \quad \text{and} \quad \hat{r}((k+1)h) = \frac{1}{m_{k+1}} \sum_{H_{k+1}} e_i e_j \]

The sums for each estimate will contain terms like:

\[ e(s_i) e(s_i + kh) \quad \text{and} \quad e(s_i) e(s_i + (k+1)h) \]

- The covariance between \( e(s_i + kh) \) is \( e(s_i + (k+1)h) \) is \( r(h) \)
- The high covariance implies that the estimates

\[ \hat{r}(kh) \quad \text{and} \quad \hat{r}((k+1)h) \]

will be highly correlated.

Computer exercise 2 — Parameter Estimation

![Graphs showing estimation results](image-url)
Computer exercise 2 — Reconstruction

For a known mean value ($\beta$) the best linear unbiased predictor (BLUP) is

$$\hat{y}_0 = X_0 \beta + \Sigma_{0k} \Sigma_{kk}^{-1} (Y_k - X_k \beta),$$

with prediction uncertainty

$$V\left( y_0 \mid Y_k, \theta, \hat{\beta} \right) = \Sigma_{00} - \Sigma_{0k} \Sigma_{kk}^{-1} \Sigma_{k0}. $$

Note that:

- The reconstructed field is smooth (excludes the nugget compared to the simulated field).
- Minimum prediction variance is $\sigma^2.$
- Variance increases away from observed points and towards the edges.
- Often better to plot the standard deviation.

Gaussian Markov random fields (GMRF)

A Gaussian random field $X \sim N(\mu, \Sigma)$ that satisfies

$$p\left( x_i \mid \{x_j : j \neq i\} \right) = p\left( x_i \mid \{ x_j : j \in N_i \} \right)$$

is a Gaussian Markov random field.

$Q = \Sigma^{-1}$ is called the precision matrix and the model becomes

$$X \sim N(\mu, Q^{-1})$$

cf.

$$X \sim N(\mu, \Sigma)$$

With conditional expectation and variance

$$E(X_u | X_k) = \mu_u - Q_{uu}^{-1} Q_{uk} (X_k - \mu_k), \quad V(X_u | X_k) = Q_{uu}^{-1}.$$
The precision matrix is sparse

Elements in the precision matrix of a Gaussian Markov random field are non-zero only for neighbours and diagonal elements.

\[ j \notin \{i, N_i\} \iff Q_{ij} = 0. \]

Under mild conditions the Cholesky factorisation \( Q = R^T R \) will be sparse if \( Q \) is sparse.

For a stationary field, the elements of \( Q \) are determined via a simple function of \( u, v \),

\[ Q_{ij} = q(v - u) \]

Requirements on \( q \)

There are some requirements on the \( q \)-function to ensure that the resulting matrix is a valid precision matrix.

1. To ensure symmetry we require that

\[ q(v - u) = \begin{bmatrix} a & b & c \\ d & e & d \\ c & b & a \end{bmatrix}. \]

2. A sufficient condition for the resulting \( Q \) to be positive-definite is that

\[ e > 0, \]

\[ |e| > 2(|a| + |b| + |c| + |d|). \]

The resulting \( Q \)-matrix is said to be diagonally dominant

\[ |Q_{ii}| > \sum_{i \neq j} |Q_{ij}|. \]

Gaussian random field and GMRFs

Gaussian random fields:

+ Easy to define reasonable (stationary) dependence structures through the covariance function.
  
  − Hard computations due to a large, dense covariance matrix.

Gaussian Markov random fields:

− Difficult to construct reasonable dependence structures.
  
  + Fast computations due to a sparse precision matrix.

For larger neighbourhoods diagonally dominance excludes several interesting parameter sets.
Constructing GMRF models

- In order to use GMRFs instead of full covariance models, we need to construct useful, sparse, $Q$-matrices.
- A simple, traditional choice, is the conditional autoregressive, CAR(1) model, with local $q$-pattern

$$q = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 + K^2 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Here the conditional expectations is the weighted sum of the four nearest neighbors.

Interpreting $q$

For stationary fields the local $q$-pattern defines each column in $Q$:

$$q = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 + K^2 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Given the conditional expectation for a GMRF

$$E(X_u | X_k) = \mu_u - Q^{-1}_{uu} Q_{uk} (X_k - \mu_k),$$

$q$ can also be interpreted as the conditional weighting of nearby pixels

$$q = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 + K^2 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$X_{image} = \begin{bmatrix} \vdots & X_{i-1,j-1} & X_{i-1,j} & X_{i-1,j+1} & \vdots \\ \vdots & X_{i,j-1} & X_{i,j} & X_{i,j+1} & \vdots \\ \vdots & X_{i+1,j-1} & X_{i+1,j} & X_{i+1,j+1} & \vdots \end{bmatrix}$$

Assuming $\mu = 0$:

$$E(x_{ij} | \text{all other pixels}) = \frac{X_{i-1,j-1} + X_{i-1,j} + X_{i-1,j+1} + X_{i+1,j-1} + X_{i+1,j} + X_{i+1,j+1}}{4 + K^2}$$
CAR(1) models

In order to use GMRFs instead of full covariance models, we need to construct useful, sparse, $Q$-matrices.

### Conditional autoregressive models (CAR)

A **mean zero** CAR(1) model is defined by

$$x_i | \{x_j : j \in N_i \} \in \mathbb{N} \left( \frac{1}{\kappa^2 + \|N_i\|} \sum_{j \in N_i} x_j, \frac{1}{\kappa^2 + \|N_i\|} \right)$$

Remember

$$x_i | \{x_j : j \neq i \} \in \mathbb{N} \left( -\frac{1}{Q_{ii}} \sum_{j \neq i} Q_{ij} x_j, Q_{ii}^{-1} \right)$$

### CAR(1) models

For observations on a regular grid the resulting local $q$-pattern (with pixel $x_i$ marked in red) is

$$q = \tau \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 + \kappa^2 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

With special treatment of the edges

$$q_{\text{edge}} = \tau \begin{bmatrix} -1 & 0 \\ 3 + \kappa^2 & -1 \\ -1 & 0 \end{bmatrix}$$

$$q_{\text{corner}} = \tau \begin{bmatrix} 2 + \kappa^2 & -1 \\ -1 & 0 \end{bmatrix}$$

### Boundary effects

Three options for handling the boundary effects exist.

- **Dirichlet boundary condition:** $x(u) = 0$ on the boundary.
- **Neumann boundary condition:** $x'(u) = 0$ perpendicular to the boundary.
- **Torus:**
  By folding the image onto a torus we completely avoid edges, eliminating the need for boundary conditions. However it introduces strange dependencies . . .
Boundary effects

The three options correspond to different edge corrections

\[
q_{\text{Dirichlet}} = \begin{bmatrix}
-1 & 0 \\
4 + \kappa^2 & -1 \\
-1 & 0
\end{bmatrix}
\]

\[
q_{\text{Neumann}} = \begin{bmatrix}
-1 & 0 \\
3 + \kappa^2 & -1 \\
-1 & 0
\end{bmatrix}
\]

\[
q_{\text{Torus}} = \begin{bmatrix}
-1 & 0 & \cdots & 0 \\
4 + \kappa^2 & -1 & 0 & \cdots & 0 \\
-1 & 0 & \cdots & 0 & -1
\end{bmatrix}
\]

Boundary effects (cont.)

Dirichlet Neumann Torus

Precision matrix for a CAR(1)

Dividing the \(q\)-pattern into a component containing the parameter, \(\kappa^2\), and a 2nd order finite difference operator

\[
q = \tau \left( \kappa^2 \begin{bmatrix} 1 \\ \hline (I) \end{bmatrix} + \begin{bmatrix} -1 & 4 & -1 \\ \hline 4 & -1 & -1 \\ \approx -\Delta (G) \end{bmatrix} \right)
\]

We can write the full precision matrix as

\[
Q = \tau \left( \kappa^2 I + G \right)
\]

The resulting matrix is positive definite if \(\kappa^2 > 0\).
Precision matrix for a non-gridded data

When constructing a CAR(1) precision matrix for regional data we first need to determine the dependence between regions. This is commonly done by constructing an undirected graph.

1. Replace all regions with nodes (or vertices).
2. If two regions share a common border, connect the nodes with an edge.
3. The resulting graph defines a neighbourhood for the region.

Two nodes connected by an edge are said to be neighbours.

The CAR(1) precision is constructed as

\[
Q_{ij} = \begin{cases} 
  k^2 + \|N_i\| & \text{if } i = j, \\
  -1 & \text{if } j \in N_i, \\
  0 & \text{if } j \notin N_i.
\end{cases}
\]

Often written as

\[
Q = k^2 I + G
\]

Where \(G\) is a neighbourhood matrix.
**Precision matrix for a non-gridded data — Example**

\[ Q = \kappa^2 I + G \]

\[
G = \begin{bmatrix}
3 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\
0 & 4 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\
-1 & 0 & 3 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 4 & -1 & -1 & 0 & 0 & -1 \\
-1 & -1 & -1 & -1 & -1 & 10 & -1 & -1 & -1 & -1 \\
0 & -1 & 0 & 0 & -1 & -1 & 3 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & -1 & 0 & 2 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 2 & 0 \\
-1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 3
\end{bmatrix}
\]

**Simultaneous autoregressive models (SAR)**

A **mean zero** SAR(1) models is defined by

\[
x_i - \frac{1}{\kappa^2 + \|N_i\|} \sum_{j \in N_i} x_j \sim N(0, \sigma^2) \quad \forall i
\]

i.e. the difference between each pixel and its averaged neighbours should be **jointly (simultaneously) normal**.

Compare this to the **conditional** normality in the CAR model.

\[
x_i | \{x_j : j \in N_i\} \sim N\left(\frac{1}{\kappa^2 + \|N_i\|} \sum_{j \in N_i} x_j, \frac{1}{\tau (\kappa^2 + \|N_i\|)} \right)
\]

**Precision matrix for a SAR(1)**

We can write the SAR model as

\[
(\kappa^2 + \|N_i\|)x_i - \sum_{j \in N_i} x_j = (\kappa^2 + \|N_i\|) \varepsilon_i \quad \varepsilon_i \sim N\left(0, \sigma^2 \right)
\]

Since this has to be valid, simultaneously, for all \(i\) we express the model as a linear equation system

\[
(\kappa^2 I + G)x = e \quad e \sim N\left(0, \tau^{-1} I\right)
\]

Here \(\tau\) and \(\sigma^2\) are related by

\[
\tau^{-1} = (\kappa^2 + \|N_i\|)^2 \sigma^2
\]
### Precision matrix for a SAR(1)

Solving the equation system

\[(k^2 I + G)x = e \quad e \in N \left(0, \tau^{-1}I\right),\]

\[x = (k^2 I + G)^{-1}e.\]

and using that a linear combination of Gaussians is Gaussian

\[
\begin{align*}
E(x) &= E\left((k^2 I + G)^{-1}e\right) = 0 \\
V(x) &= V\left((k^2 I + G)^{-1}e\right) = (k^2 I + G)^{-1}\tau^{-1}I(k^2 I + G)^{-\top}
\end{align*}
\]

We obtain the precision matrix for the SAR(1) model

\[
Q_{SAR} = \tau\left(k^2 I + G\right)^\top\left(k^2 I + G\right) \propto Q_{CAR}^2
\]

### q-pattern for CAR and SAR models (ignoring \(\tau\))

**CAR(1):**

\[
q = k^2 \begin{bmatrix} 1 \\ (i) \end{bmatrix} + \begin{bmatrix} -1 & 4 & -1 \\ -1 & \approx -\Delta(G) \end{bmatrix}
\]

**SAR(1) = CAR(2):**

\[
q = k^4 \begin{bmatrix} 1 \\ (i) \end{bmatrix} + 2k^2 \begin{bmatrix} -1 & 4 & -1 \\ -1 & \approx -\Delta(G) \end{bmatrix} + \begin{bmatrix} 1 & 2 & -8 & 2 \\ 1 & -8 & 20 & -8 \\ 2 & -8 & 2 & 1 \end{bmatrix} \approx \Delta^2(G^\top G)
\]

### Parameter restrictions

For the CAR(1) process it is easy to see that \(k^2 > 0\) gives a valid (positive definite) precision matrix.

For the SAR process we have that \(Q_{SAR} \propto Q_{CAR}^2\) is positive definite if \(Q_{CAR}\) is positive definite, i.e. \(k^2 > 0\).

If we instead require \(Q_{SAR}\) to be **diagonally dominant**

\[
|Q_{ij}| > \sum_{i\neq j} |Q_{ij}|.
\]

\[
|k^4 + 8k^2 + 20| > 4 \cdot |2k^2| + 4 \cdot 1 + 4 \cdot 2 + 4 \cdot |20| + |8|.
\]

\[
|k^4| > 24
\]
Remaining Points

Friday:
- Bayesian Hierarchical models.
- Parameter estimation for GMRF:s.
- Non-Gaussian observations.

Monday:
- Home assignment 2.
- Non-Gaussian observations.
- Matérn-like GMRF:s.