Extending conditional-autoregressive models for space-time disease mapping

Alastair Rushworth

ISBA World Meeting, Cancún
17th July 2014
Acknowledgements

Work in this talk part of a larger collaboration:

University of Glasgow
{ Duncan Lee
Richard Mitchell
}

University of Southampton
{ Sujit Sahu
Sabyasachi Mukhopadhyay
}

UK Met Office
{ Paul Agnew
Christophe Sarrane
}

Funded by the EPSRC
{ EPSRC
Engineering and Physical Sciences Research Council
}
Context: England LHAs - Respiratory admission SIR

- Date = 12−2004
- Newcastle
- Birmingham
- Manchester
- London
- 123 months
- 323 Local Health Authorities (LHAs)
- January 2001 to December 2011
- 42636 data points in total
Space-time disease mapping: why do it?

- Understanding if public health policies have an effect over time

- Understanding health inequality (spatial disease risk)

- Boundary (spatial discontinuity) detection

- Clustering of regions with similar (eg. temporal) risk attributes
Modelling disease counts in space

Typical spatial model might look like

\[ Y_i | E_i, R_i \sim \text{Poisson}(E_i R_i) \]

\[ \ln(R_i) = \beta_0 + \sum_{j=1}^{p} x_{ij} \beta_j + \phi_i, \quad i = 1, \ldots, N \]

Where

\[ Y_i \quad \text{disease counts} \]

\[ E_i \quad \text{expected cases} \]

\[ R_i \quad \text{disease risk} \]

\[ \sum_{j=1}^{p} x_{ij} \beta_j \quad \text{covariates} \]

\[ \pi(\phi) \propto \exp \left[ -\frac{1}{2\sigma^2} \left( \sum_{i \sim j} w_{ij} (\phi_i - \phi_j)^2 \right) \right] \]
Localised smoothing

Different strategies achieving this:

Treat $\phi$’s prior variance, $\sigma$, as spatially varying

- Brewer and Nolan (2007); Reich and Hodges (2008)

Treat the weights $w_{ij}$ as random variables

- Ma, Carlin and Banerjee (2010); Lee and Mitchell (2013)

Use a clustering or grouping prior for the random effects

- Richardson and Green (2002); Knorr-Held and Raßer (2000)
Starting with Poisson model and Intrinsic CAR for random effects:

\[ Y_i|E_i, R_i \sim \text{Poisson}(E_i R_i) \]

\[ \log(R_i) = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j + \phi_i, \quad i = 1, \ldots, N \]

\[ p(\phi|\sigma, \{w_{ij}\}) = C(\sigma, \{w_{ij}\}) \exp \left( -\frac{1}{2\sigma^2} \sum_{i \sim j} w_{ij}(\phi_i - \phi_j)^2 \right) \]

Normalising term \( C(\sigma, \{w_{ij}\}) \) included to emphasise that \( w_{ij} \) is now being treated as unknown.
Idea: Call the non-zero $w_{ij}$, $w_{ij}^{+}$, and treat as Bernoulli random variables with unknown probabilities, $p_{ij}$.

Transform and smooth $p_{ij}$ using a further CAR prior.

\[
  w_{ij}^{+} \mid p_{ij} \sim \text{Bernoulli}(p_{ij}) \quad \text{and} \quad \logit(p_{ij}) = z_{ij}' \gamma + \theta_{ij}
\]

\[
p(\theta \mid \zeta) \propto \exp \left( \frac{-1}{2\zeta^2} \sum_{ij \sim kl} (\theta_{ij} - \theta_{kl})^2 \right)
\]

Pros: Maintains binary nature of $w_{ij}$

Cons: Assumes relatively strong smoothness over the $w_{ij}$; a lot of parameters to estimate $\phi$, $w_{ij}$, $\theta$ that can be hard to identify.
A localised model in space and time

Poisson model for counts in space and time (indexed by $i$ and $j$, respectively)

\[
Y_{ij} | E_{ij}, R_{ij} \sim \text{Poisson}(E_{ij}R_{ij})
\]

\[
\log(R_{ij}) = \beta_0 + \phi_{ij}, \quad i = 1, \ldots, N \text{ and } j = 1, \ldots, T
\]

Letting $\phi_t = (\phi_{1t}, \ldots \phi_{Nt})$, prior for spatial random effects is

\[
p(\phi_t | \sigma, \{w_{ij}\}) = C(\sigma, \{w_{ij}\}) \exp \left( -\frac{1}{2\sigma^2} \sum_{i \sim j} w_{ij}(\phi_{it} - \phi_{jt})^2 \right)
\]
A localised model in space and time

**Idea:** Treat the non-zero elements of $w_{ij}$, $w_{ij}^+$, as random variables on $[0, 1]$.

Then transform using $\text{logit}(w_{ij}^+) = z_{ij}$ and smooth these using a Leroux prior.

\[
p(z_{ij}) \propto \exp \left[ -\frac{1}{2\tau^2} \left( \rho \sum_{ij \sim kl} (z_{ij} - z_{kl})^2 + (1 - \rho) \sum_{ij} (z_{ij})^2 \right) \right]
\]

**Pros:** Treating $w_{ij}$ as non-binary makes things slightly easier; Leroux prior is a bit more flexible than the Intrinsic CAR model.

**Cons:** Still a lot of parameters to estimate
A localised model in space and time - computation

Need to update $\phi_{N \times T}$

- Can be done very quickly using one-at-a-time updating in C++ and exploiting matrix sparsity

Need to update $z_{ij}$

- Block updating using each blocks conditional prior variance
- Main bottleneck: recalculating determinant of $\phi$ prior precision for each block proposal
- 85% of MCMC time even with sparse Cholesky routine
Some illustrative simulations

- Using Glasgow intermediate geographies as a spatial template
- To investigate effect of increasing time points, generate 10 or 50 repeated realisations of this surface
- Simulate spatial counts step-function using Poisson with mean of either 100 or 200.
- Interested in effect of $\rho$: fix at 0, or treat as unknown.
$T = 10; \rho = 0.993$
$T = 50; \rho = 0.983$
$T = 10; \rho = 0$
$T = 50; \rho = 0$
DIC and parameter estimates for each model

**DIC table for** $T = 10$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.993</td>
<td>25093.07</td>
</tr>
<tr>
<td>0</td>
<td>25331.94</td>
</tr>
</tbody>
</table>

**DIC table for** $T = 50$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.983</td>
<td>127717.5</td>
</tr>
<tr>
<td>0</td>
<td>128425.7</td>
</tr>
</tbody>
</table>
Some more (simulation) results

![Box plot showing risk and RMSE for different model types with rho fixed and rho varies.](image)
Some more (simulation) results

Model type
- rho fixed
- rho varies

Risk
DIC

1 2 3

47500
50000
52500
55000
Application to England LHAs - Data

Date = 12–2004

Newcastle
Manchester
Birmingham
London

0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Application to England LHAs - Results

Date = 12−2004

Newcastle
Birmingham
Manchester
London

0.19
0.37
0.56
0.74
0.92
1.1
1.28
1.46
1.64
1.82
2
Application to England LHAs - Results

Distribution of posterior mean $w_{ij}$

$w_{ij} \cdot (1 - P[\text{discontinuity}_{ij}])$

Frequency

0.0 0.2 0.4 0.6 0.8 1.0
0 100 200 300 400 500
Current work

Model is obviously overparameterised

- Set of spatial random effects for each time and region
- Doesn’t take account of temporal structure

Letting $\tilde{\phi}_t = (\phi_{t1}, \ldots, \phi_{tN})$, where $t = 1, \ldots, T$. Assume a model of the form

$$ f(\tilde{\phi}_1, \ldots, \tilde{\phi}_T) = f(\tilde{\phi}_1) \prod_{t=2}^{T} f(\tilde{\phi}_t | \tilde{\phi}_{t-1}), $$

where

$$ \tilde{\phi}_1 \sim ICAR(0, \sigma^2 Q^{-1}) $$
$$ \tilde{\phi}_t | \tilde{\phi}_{t-1} \sim ICAR(\alpha \tilde{\phi}_{t-1}, \sigma^2 Q^{-1}). $$
More work required to understand properties of these models

- Discontinuity detection rates under realistic surfaces
- How many replications make this analysis worthwhile?
- Simulations in progress...

Difficult to incorporate information about likely discontinuity structure into prior - compromise sometimes is required due to computational constraints:

- This is the reason CAR models are so nice: (matrix) sparsity; well established tricks for model-fitting; easy to write very efficient code.

- Knorr-Held and Raßer (2000) use a nice clustering prior that can induce smoothness; but requires RJMCMC.
References


Thanks for listening!