A fast sampler for data simulation from spatial, and other, Markov random fields

Andee Kaplan

Iowa State University
ajkaplan@iastate.edu

June 22, 2017

Slides available at http://bit.ly/kaplan-phd

Joint work with M. Kaiser, S. Lahiri, and D. Nordman

Overview

Thesis: On advancing MCMC-based methods for Markovian data structures with applications to deep learning, simulation, and resampling

Goal: Develop statistical inference via Markov chain Monte Carlo (MCMC) techniques in complex data problems related to statistical learning, the analysis of network/graph data, and spatial resampling.

Challenge: Develop implementations which are both *statistically rigorous* and *computationally scalable* by exploiting conditional independence.

1. Statistical quantification of graph models used in deep machine learning and image classification
(Ch. 2 & 3)
2. Fast methods for simulating spatial, network, and other data
(Ch. 4 & 5)

Goal

- Markov random field models are popular for spatial or network data
- Rather than specifying a joint distribution directly, a model is specified through a set of full conditional distributions for each spatial location
- Assume the spatial data are on a regular lattice (wrapped on a torus)

Goal: A new, provably fast approach for simulating spatial/network data.

Spatial Markov random field (MRF) models

Notation

- Variables $\{Y(\boldsymbol{s}_i) : i = 1, \ldots, n\}$ at locations $\{\boldsymbol{s}_i : i = 1, \ldots, n\}$
- Neighborhoods: $\mathcal{N}_i$ specified according to some configuration
- Neighboring Values: $\boldsymbol{y}(\mathcal{N}_i) = \{y(\boldsymbol{s}_j) : \boldsymbol{s}_j \in \mathcal{N}_i\}$
- Full Conditionals: $\{f_i(y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i), \boldsymbol{\theta}) : i = 1, \ldots, n\}$
  - $f_i(y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i), \boldsymbol{\theta})$ is conditional pmf/pdf of $Y(\boldsymbol{s}_i)$ given values for its neighbors $\boldsymbol{y}(\mathcal{N}_i)$
  - Often assume a common conditional cdf $F_i = F$ form $(f_i = f)$ for all i

Exponential family examples

1. Conditional Gaussian (3 parameters):

$$f_i(y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i), \alpha, \eta, \tau) = \frac{1}{\sqrt{2\pi}\tau} \exp\left(-\frac{[y(\boldsymbol{s}_i) - \mu(\boldsymbol{s}_i)]^2}{2\tau^2}\right)$$

   $Y(\boldsymbol{s}_i)$ given neighbors $\boldsymbol{y}(\mathcal{N}_i)$ is normal with variance τ^2 and mean

$$\mu(\boldsymbol{s}_i) = \alpha + \eta \sum_{\boldsymbol{s}_j \in \mathcal{N}_i} [y(\boldsymbol{s}_j) - \alpha]$$

2. Conditional Binary (2 parameters):
   $Y(\boldsymbol{s}_i)$ given neighbors $\boldsymbol{y}(\mathcal{N}_i)$ is Bernoulli $p(\boldsymbol{s}_i, \kappa, \eta)$ where

$$\mathrm{logit}[p(\boldsymbol{s}_i, \kappa, \eta)] = \mathrm{logit}(\kappa) + \eta \sum_{\boldsymbol{s}_j \in \mathcal{N}_i} [y(\boldsymbol{s}_j) - \kappa]$$

In both examples, η represents a dependence parameter.

Concliques

Cliques – Hammersley and Clifford (1971)

Singletons and sets of locations such that each location in the set is a neighbor of all other locations in the set
Example: Four nearest neighbors gives cliques of sizes 1 and 2

The Converse of Cliques – Concliques

Sets of locations such that no location in the set is a neighbor of any other location in the set

4 Nearest Neighbors

·	*	·
*	*s*	*
·	*	·

Concliques 4 Nearest Neighbors

1	2	1	2
2	1	2	1
1	2	1	2
2	1	2	1

8 Nearest Neighbors

*	*	*
*	*s*	*
*	*	*

Concliques 8 Nearest Neighbors

1	2	1	2
3	4	3	4
1	2	1	2
3	4	3	4

Illustrative Example

- Spatial dataset from Besag (1977)
- Binary observations located on a 14×179 indicating the presence or absence of footrot in endive plants

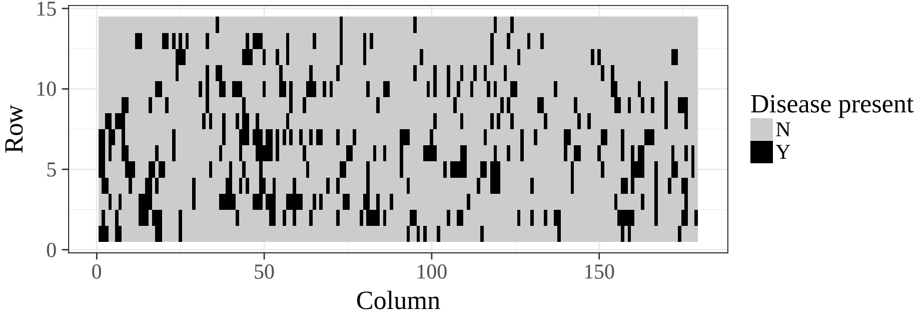


Figure 1: The endive dataset, a 14×179 rectangular lattice with binary data encoding the presence or absence of footrot in endive plants from Besag (1977).

Three models

1. Isotropic centered autologistic model (Caragea and Kaiser 2009; Besag 1972; Besag 1977)
2. Centered autologistic model with two dependence parameters
3. Centered autologistic model as in (2) but having large scale structure determined by regression on the horizontal coordinate u_i of each spatial location $\boldsymbol{s}_i = (u_i, v_i)$.

Three models (Cont'd)

Conditional mass function of the form

$$f_i(y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i), \boldsymbol{\theta}) = \frac{\exp[y(\boldsymbol{s}_i)A_i\{\boldsymbol{y}(\mathcal{N}_i)\}]}{1 + \exp[y(\boldsymbol{s}_i)A_i\{\boldsymbol{y}(\mathcal{N}_i)\}}, \quad y(\boldsymbol{s}_i) = 0, 1,$$

with

Model	Natural parameter function
(1)	$A_i\{\boldsymbol{y}(\mathcal{N}_i)\} = \log\left(\frac{\kappa}{1-\kappa}\right) + \eta \sum_{\boldsymbol{s}_j \in \mathcal{N}_i} \{y(\boldsymbol{s}_j) - \kappa\}$
(2)	$A_i\{\boldsymbol{y}(\mathcal{N}_i)\} = \log\left(\frac{\kappa}{1-\kappa}\right) + \eta_u \sum_{\boldsymbol{s}_j \in N_{u,i}} \{y(\boldsymbol{s}_j) - \kappa\} + \eta_v \sum_{\boldsymbol{s}_j \in N_{v,i}} \{y(\boldsymbol{s}_j) - \kappa\}$
(3)	$A_i\{\boldsymbol{y}(\mathcal{N}_i)\} = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta_u \sum_{\boldsymbol{s}_j \in N_{u,i}} \{y(\boldsymbol{s}_j) - \kappa_i\} + \eta_v \sum_{\boldsymbol{s}_j \in N_{v,i}} \{y(\boldsymbol{s}_j) - \kappa_i\},$ $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \beta_0 + \beta_1 \boldsymbol{u}_i$

Table 1: Full conditional distributions of three binary MRF models for the endive data.

Bootstrap percentile confidence intervals

- Fit three models of increasing complexity to these data via pseudo-likelihood (Besag 1975)
- Apply simulation (parametric bootstrap) to obtain reference distributions for statistics based on the resulting estimators

	Model (1)		Model (2)			Model (3)			
	η	κ	η_u	η_v	κ	η_u	η_v	β_0	β_1
2.5%	0.628	0.107	0.691	0.378	0.106	-0.225	-0.221	-1.822	-0.003
50%	0.816	0.126	0.958	0.660	0.125	0.000	0.004	-1.600	-0.001
97.5%	1.001	0.145	1.220	0.921	0.145	0.209	0.214	-1.391	0.001

Table 2: Bootstrap percentile confidence intervals in all three autologistic models.

Sampling distributions of dependence parameters

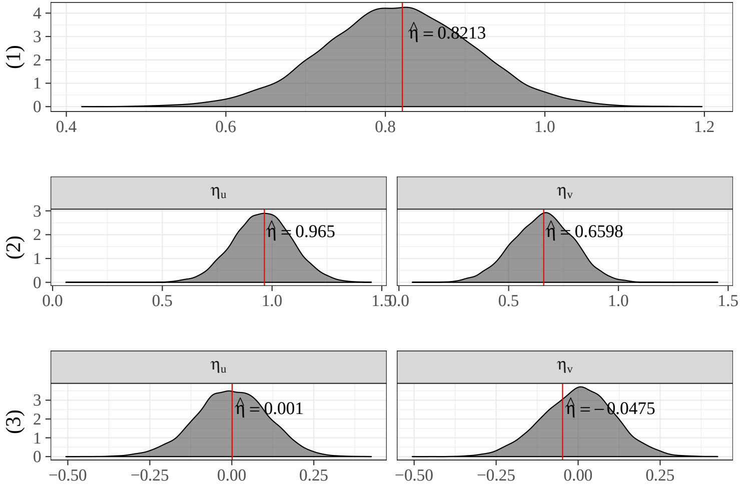


Figure 2: Sampling distribution of the dependence parameters (η, η_u, and η_v) for the three centered autologistic models.

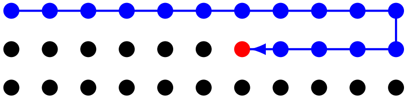
Common Spatial Simulation Approach

With common conditionally specified models for spatial lattice, standard MCMC simulation approach via Gibbs sampling is:

Starting from some initial $\boldsymbol{Y}_*^{(j)} \equiv \{Y_*^{(j)}(\boldsymbol{s}_1), \dots, Y_*^{(j)}(\boldsymbol{s}_n)\}$,

1. Moving row-wise, for $i = 1, \dots, n$, individually simulate/update $Y_*^{(j+1)}(\boldsymbol{s}_i)$ for each location $\boldsymbol{s}_i$ from conditional cdf F given

$$Y_*^{(j+1)}(\boldsymbol{s}_1), \dots, Y_*^{(j+1)}(\boldsymbol{s}_{i-1}), \quad Y_*^{(j)}(\boldsymbol{s}_{i+1}), \dots, Y_*^{(j)}(\boldsymbol{s}_n)$$



2. n individual updates provide 1 full Gibbs iteration.
3. Repeat 1-2 to obtain M resampled spatial data sets $\boldsymbol{Y}_*^{(j)}$, $j = 1, \dots, M$ (e.g., can burn-in, thin, etc.)

Conclique-based Gibbs sampler

Using the conditional independence of random variables at locations within a conclique we propose a conclique-based Gibbs sampling algorithm for sampling from a MRF.

1. Split locations into Q disjoint concliques, $\mathcal{D} = \cup_{i=1}^{Q} \mathcal{C}_i$.
2. Initialize the values of $\{Y^{(0)}(\boldsymbol{s}) : \boldsymbol{s} \in \{\mathcal{C}_2, \dots, \mathcal{C}_Q\}\}$.
3. Starting from $\mathcal{C}_1$ for the i^{th} iteration, draw $\{Y^{(i)}(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{C}_1\}$ as random sample where $Y^{(i)}(\boldsymbol{s}) \overset{iid}{\sim} F(y(\boldsymbol{s}) | Y^{(i-1)}(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{N}(\boldsymbol{s}))$
4. Update observations conclique-wise (using previous conclique updates).
   - For $j = 2, \dots, Q$, draw $\{Y^{(i)}(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{C}_j\}$ as random sample where $Y^{(i)}(\boldsymbol{s}) \overset{iid}{\sim} F(y(\boldsymbol{s}) | \{Y^{(i)}(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{N}(\boldsymbol{s}) \cap \mathcal{C}_k \text{ where } k < j\}, \{Y^{(i-1)}(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{N}(\boldsymbol{s}) \cap \mathcal{C}_k \text{ where } k > j\})$

This works by conditional independence & because neighbors for updating one conclique always belong to other concliques.

It's (computationally) fast!

- Because we are using batch updating vs. sequential updating of each location, this approach is **computationally fast**.
- A flexible `R` package using `Rcpp` (called `conclique`, to appear on CRAN) that implements a conclique-based Gibbs sampler while allowing the user to specify an arbitrary model.

It's (provably) fast!

- While computationally fast, the MCMC sampler is also provably geometrically ergodic (i.e., the MCMC mixes at a fast rate) in a general sense, which is unusual for spatial data.
- State-of-the-art general theory for proving geometric ergodicity of Gibbs samplers exists only for two-state samplers (i.e., drift & minorization conditions) (Johnson and Burbank 2015).
  - For common 4-nearest neighbor spatial models, there are exactly 2 concliques (two stages in the conclique-based Gibbs sampler).
  - One can formally prove that the spatial sampler proposed is geometrically ergodic for many conditional spatial models (Gaussian, Gamma, Inverse-gamma, Beta, Binomial, etc.)

Simulation comparisons

Quantitative framework from Turek et al. (2017) to compare conclique-based and sequential Gibbs sampler efficiency

1. Mixing effectiveness (algorithmic efficiency)
2. Computational demands of the algorithm (computational efficiency)

Algorithmic efficiency:

$$A = \min_{1 \leq i \leq n} \left\{ \left(1 + 2 \sum_{j=1}^{\infty} \rho_i(j) \right)^{-1} \right\},$$

Computational efficiency:

$$C = \begin{cases} \sum_{k=1}^{Q} \text{samp}(\{Y(\boldsymbol{s}_i) : \boldsymbol{s}_i \in \mathcal{C}_k\} | \mathcal{C}_j, j \neq k) & \text{Conclique-based} \\ \sum_{k=1}^{n} \text{samp}(Y(\boldsymbol{s}_k) | Y(\boldsymbol{s}_j), j \neq k) & \text{Sequential} \end{cases}$$

Simulation comparisons (Cont'd)

Gibbs	Model (a)		Model (b)		Model (c)	
	A	C	A	C	A	C
Conclique	0.807	2.9×10^{-4}	0.745	2.7×10^{-4}	0.72	3×10^{-4}
Sequential	0.809	0.029	0.749	0.029	0.704	0.024

Table 3: Measures of algorithmic and computational efficiency, A and C, for three autologistic models on a 40×40 grid. We compare the metrics for a conclique-based Gibbs sampler and a sequential sampler.

Endive data timing

- Endive example dataset simulations performed with the proposed (conclique-based) Gibbs sampler
- Reported results would have been virtually identical with the same number of iterations to the standard sequential Gibbs sampler
- Generation of the reference distribution using the standard sampler would have taken approximately
    1. 25.31 minutes longer
    2. 31 minutes longer
    3. 40.7 minutes longer
- Conclique MRF sampler had running times
    1. 8.15 seconds
    2. 14.74 seconds
    3. 95.71 seconds

Timing simulations

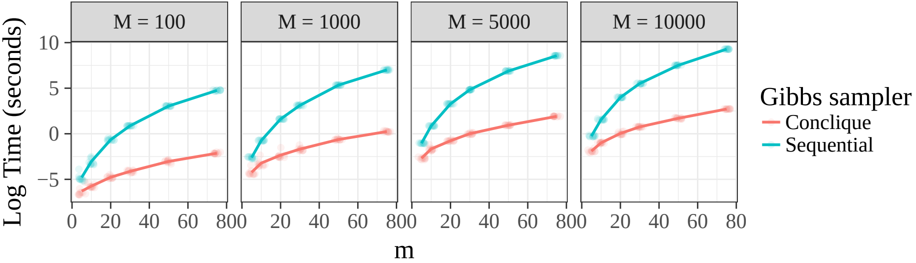


Figure 3: Comparisons of log time for simulation of $M = 100, 1000, 5000, 10000$ four-nearest neighbor Gaussian MRF datasets on a lattice of size $m \times m$ for various size grids, $m = 5, 10, 20, 30, 50, 75$, using sequential and conclique-based Gibbs samplers.

For $10,000$ iterations/samples on 75×75 grid, conclique-based took 15.05 seconds and sequential took 1.076197×10^4 seconds ≈ 2.99 hours.

Application (Goodness of Fit)

- An important question for Markov random field models with spatial data is

  How to assess/diagnose fit?

- Kaiser, Lahiri, and Nordman (2012) provide a methodology for performing GOF tests using concliques
- Conclique-based Gibbs sampling allows for fast approximation of the reference distribution for the GOF test statistics in this methodology

Generalized spatial residuals

Definition

- $F(y|\boldsymbol{y}(\mathcal{N}_i), \boldsymbol{\theta})$ is the conditional cdf of $Y(\boldsymbol{s}_i)$ under the model
- Substitute random variables, $Y(\boldsymbol{s}_i)$ and neighbors $\{Y(\boldsymbol{s}_j) : \boldsymbol{s}_j \in \mathcal{N}_i\}$, into (continuous) conditional cdf to define residuals:

$$R(\boldsymbol{s}_i) = F(Y(\boldsymbol{s}_i)|\{Y(\boldsymbol{s}_j) : \boldsymbol{s}_j \in \mathcal{N}_i\}, \boldsymbol{\theta}).$$

Key Property

Let $\{\mathcal{C}_j : j = 1, \ldots, q\}$ be a collection of concliques that partition the integer grid. Under the conditional model, **spatial residuals** ***within*** **a conclique are iid Uniform**$(0, 1)$**-distributed**:

$$\{R(\boldsymbol{s}_i) : \boldsymbol{s}_i \in \mathcal{C}_j\} \stackrel{iid}{\sim} \text{Uniform}(0, 1) \qquad \text{for } j = 1, \ldots, q$$

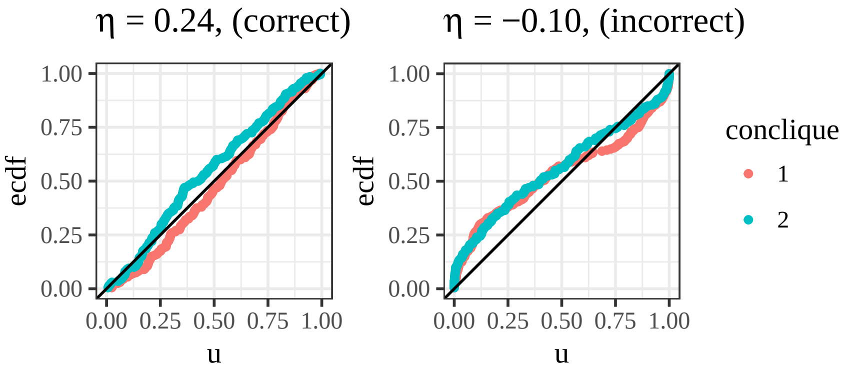
(Kaiser, Lahiri, and Nordman 2012)

Simple example

Gaussian Conditional Model - 20 $\times$ 20 Lattice, 4-nearest Neighbors

Let $Y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i) \sim N(\mu(\boldsymbol{s}_i), \tau^2)$, where $\mu(\boldsymbol{s}_i) = \alpha + \eta \sum_{\boldsymbol{s}_j \in \mathcal{N}_i} (y(\boldsymbol{s}_j) - \alpha)$.

Truth: $\alpha = 10, \tau^2 = 2, \eta = 0.24$.



From residuals to test statistics

Residual Empirical Distribution

Divide locations $\{\boldsymbol{s}_i\}_{i=1}^n$ into concliques: $\mathcal{C}_j$, $j = 1, \dots, q$
For j^{th} conclique, empirical cdf and and its difference to Uniform$(0, 1)$ cdf

$$G_{jn}(u) = \frac{1}{|\mathcal{C}_j|} \sum_{\boldsymbol{s}_i \in \mathcal{C}_j} I[R(\boldsymbol{s}_i) \leq u]$$
$$W_{jn}(u) \equiv n^{1/2} \left[G_{jn}(u) - u\right]; \quad u \in [0, 1]$$

Test Statistics

$$\begin{aligned} T_{1n} &= \max_{j=1,\dots,q} \sup_{u \in [0,1]} |W_{jn}(u)| \\ T_{2n} &= \frac{1}{q} \sum_{j=1}^{q} \left(\int_0^1 |W_{jn}(u)|^2 du \right)^{1/2} \end{aligned}$$

Hypothesis testing

Composite Hypothesis

$$H_0(C): \text{The conditional distributions of } \{Y(\boldsymbol{s}_i): i = 1, \ldots, n\} \text{ are } F(y(\boldsymbol{s}_i)|\boldsymbol{y}(\mathcal{N}_i), \boldsymbol{\theta})$$

where $\boldsymbol{\theta} \in \Theta$ is some *unknown* parameter value

Theoretical Challenge

Centered residual edfs $W_{jn}(u)$ are *not* independent over concliques & residuals/test statistics computed from estimated parameter, $\hat{\boldsymbol{\theta}}$.

- Asymptotic behavior of test statistics T_{kn} is non-trivial
- Resampling is helpful for approximating test statistic T_{kn} distributions

In practice

In application, a conditional distribution F model is formulated/specified.

1. Fit model $\hat{\theta}$ to original data $Y_1, \dots, Y_n$
2. Compute generalized residuals and test statistics: T_{kn}
3. Simulate spatial data $Y_1^*, \dots, Y_n^*$ from fitted cond. cdf: $F_{\hat{\theta}}$
4. Fit model to simulated data: $\hat{\theta}^*$
5. Compute generalized residuals and test statistics: T_{kn}^* from $Y_1^*, \dots, Y_n^*$ and $F_{\hat{\theta}^*}$
6. Do 3-5 many times
7. Result is reference distribution for test statistic T_{kn}

In simulating/resampling step 3 for spatial data, can use conclique-based Gibbs sampler due to the conditional specification F for each location.

Theory for the spatial simulation method

Let $P_{n^*}^{(M)}$ denote the joint distribution of spatial data $\boldsymbol{Y}_{n*}^{(M)}$ at the Mth iteration of the conclique-based Gibbs sampler from cond. cdf $F \equiv F_{\hat{\boldsymbol{\theta}}_n}$.

The bootstrap approximation for the GOF statistic is theoretically valid

- As $M \to \infty$, $P_{n^*}^{(M)}(T_{kn}^* \leq x) \to P_{n^*}(T_{kn}^* \leq x)$

  Gibbs sampler approximates test distribution from fitted cond. cdf $F_{\hat{\boldsymbol{\theta}}_n}$ because the conclique-based Gibbs sampler is *Harris ergodic*.

- As $n \to \infty$, $F_{\hat{\boldsymbol{\theta}}_n} \xrightarrow{p} F_{\boldsymbol{\theta}_0}$ & $P_{n^*}(T_{kn}^* \leq x) - P(T_{kn} \leq x) \xrightarrow{p} 0$

  T_{kn}^*-distribution (from joint data distribution induced by fitted cond. cdf $F_{\hat{\boldsymbol{\theta}}_n}$) converges to T_{kn}-distribution (from joint distribution induced by true cond. cdf $F_{\boldsymbol{\theta}_0}$)

This is work in progress with regards to the conclique Gibbs sampler.

Simulated example

The GOF procedure is good for distribution discrimination

- Simulated one realization of lognormal conditionals on 20×20:

  $\log Y(\boldsymbol{s}_i)$ given neighbors $\{\boldsymbol{s}_i + (0, \pm 1), \boldsymbol{s}_i + (\pm 1, 0)\}$ is normal with variance τ^2 and mean $\mu(\boldsymbol{s}_i) = \alpha + \eta \sum_{\boldsymbol{s}_j \in \mathcal{N}_i} [\log y(\boldsymbol{s}_j) - \alpha]$
- Fit Gaussian MRF & fit log Gaussian MRF to data $Y(\boldsymbol{s}_i)$ using pseudo-likelihood

Model	Expected Value α	Conditional Variance τ^2	Dependence η	Model $p-$value
True	10	2	0.24	
Log-Gaussian	9.83	2.3	0.21	0.4121176
Gaussian	8.70362×10^4	3.5162355×10^{10}	0.17	0.00019996

Reference distributions

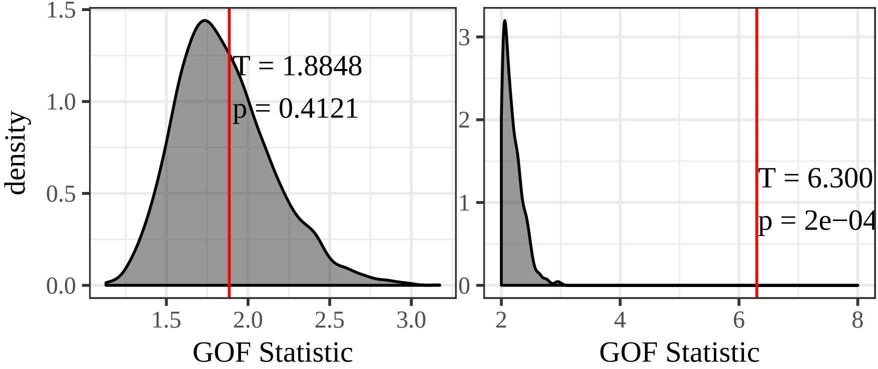


Figure 3: Bootstrapped reference distributions for the maximum across concliques of the Kolmogorov-Smirnov statistic from data generated from a four-nearest neighbor lognormal MRF with $\tau^2 = 2, \alpha = 10, \eta = 0.24$ and fit with a lognormal (left) and Gaussian (right) model.

Agricultural field trials example

The Problem

- Besag and Higdon (1999) *JRSS B* **36**, 691-746 (with discussion)
- Six agricultural field trials with corn
- They discuss appropriate Gaussian MRF model of spatial structure

GOF Procedure

- Can a simple one parameter isotropic Gaussian model be discounted?

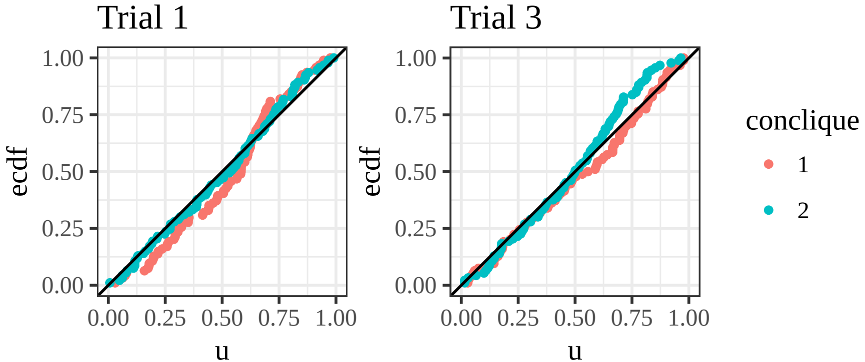
$$\mu(\boldsymbol{s}_i) = \alpha + \eta \sum_{\boldsymbol{s}_i \in \mathcal{N}_i} \{y(\boldsymbol{s}_i) - \alpha\}$$

- Four nearest neighbors, 2 concliques of sizes 93 and 94
- Maximum pseudo-likelihood estimation (e.g., Besag, 1974)
- Parametric bootstrap for 5000 data sets
- Gibbs, burn-in of 500

Agricultural field trials results

	Trial					
Statistic	1	2	3	4	5	6
T_{1n}	0.2511	0.2414	0.195	0.5935	0.8034	0.6611
T_{2n}	0.03	0.2919	0.5133	0.5801	0.8242	0.6551

Table 4: GOF test statistic p-values for the one-parameter Gaussian model.



conclique

R package (to appear on CRAN) can be installed via GitHub using the following R code.

```
devtools::install_github("andeek/conclique")
```

- Convenience functions `lattice_4nn_torus` and `min_conclique_cover`
- Gibbs samplers `run_conclique_gibbs` and `run_sequential_gibbs`
- GOF functions `spatial_residuals` and `gof_statistics`
- Bootstrap function `bootstrap_gof`

Extending `conclique`

One of the **key advantages** to using conclique-based approaches for simulation (and GOF tests) is the ability to consider non-Gaussian conditional models that go beyond a four-nearest neighbor structure.

`conclique` is generalizable in

- Dependence structure - beyond four-nearest neighbor
- Conditional distribution for each spatial location - beyond Gaussian and binary
- Generalized spatial residuals - for a user-supplied conditional distribution
- GOF statistics - aggregation beyond mean and max

Perks

Geometric Ergodicity

- Guaranteed convergence rate to the target joint data distribution for many (common) spatial MRF models
- With other established results, can obtain CLTs and Monte Carlo sample size assessments (Chan and Geyer 1994; Jones and others 2004; Hobert et al. 2002; Roberts, Rosenthal, and others 1997)

Speed & Flexibility

- Computationally more efficient alternative to the standard (sequential) Gibbs sampler
- Same general applicability in allowing accessible simulation for a wide variety of MRFs
  - Not limited to any one model or family or models
  - Can be applied to irregular lattices and non-standard neighborhoods

Future work and ideas

- Goodness-of-fit test for network data
  - The model-based method of resampling re-frames network into a collection of (Markovian) neighborhoods by using covariate information
  - Creates concliques on a graph structure
  - Use a conditionally specified network distribution (Casleton, Nordman, and Kaiser (2017)) to sample network data in a blockwise conclique-based Gibbs sampler.
- Bootstrap theory for approximating GOF statistics is ongoing work
- More user friendly API for `conclique` to appear on CRAN

Thank you

Questions?

- Slides – http://bit.ly/kaplan-phd
- Contact
  - Email – ajkaplan@iastate.edu
  - Twitter – http://twitter.com/andeekaplan
  - GitHub – http://github.com/andeek

References I

Besag, Julian. 1972. "Nearest-Neighbour Systems and the Auto-Logistic Model for Binary Data." *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 75–83.

———. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 192–236.

———. 1975. "Statistical Analysis of Non-Lattice Data." *The Statistician*. JSTOR, 179–95.

———. 1977. "Some Methods of Statistical Analysis for Spatial Data." *Bulletin of the International Statistical Institute* 47 (2): 77–92.

Besag, Julian, and David Higdon. 1999. "Bayesian Analysis of Agricultural Field Experiments." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61 (4). Wiley Online Library: 691–746.

Caragea, Petruța C, and Mark S Kaiser. 2009. "Autologistic Models with Interpretable Parameters." *Journal of Agricultural, Biological, and Environmental Statistics* 14 (3). Springer: 281.

Casleton, Emily, Daniel J Nordman, and Mark S Kaiser. 2017. "A Local Structure Model for Network Analysis." *Statistics and Its Interface* 10 (2). International Press of Boston, Inc.: 355–67.

Chan, Kung Sik, and Charles J Geyer. 1994. "Discussion: Markov Chains for Exploring Posterior Distributions." *The Annals of Statistics* 22 (4). JSTOR: 1747–58.

Hammersley, John M, and Peter Clifford. 1971. "Markov Fields on Finite Graphs and Lattices." *Unpublished*.

Hobert, James P, Galin L Jones, Brett Presnell, and Jeffrey S Rosenthal. 2002. "On the Applicability of Regenerative Simulation in Markov Chain Monte Carlo." *Biometrika*. JSTOR, 731–43.

Johnson, Alicia A, and Owen Burbank. 2015. "Geometric Ergodicity and Scanning Strategies for Two-Component Gibbs

References II

Samplers." *Communications in Statistics - Theory and Methods* 44 (15): 3125–45.

Jones, Galin L, and others. 2004. "On the Markov Chain Central Limit Theorem." *Probability Surveys* 1 (299-320): 5–1.

Kaiser, Mark S. 2007. "Statistical Dependence in Markov Random Field Models." *Preprint* 1. Citeseer.

Kaiser, Mark S, Soumendra N Lahiri, and Daniel J Nordman. 2012. "Goodness of Fit Tests for a Class of Markov Random Field Models." *The Annals of Statistics* 40 (1). Institute of Mathematical Statistics: 104–30.

Roberts, Gareth O, Jeffrey S Rosenthal, and others. 1997. "Geometric Ergodicity and Hybrid Markov Chains." *Electron. Comm. Probab* 2 (2): 13–25.

Turek, Daniel, Perry de Valpine, Christopher J Paciorek, Clifford Anderson-Bergman, and others. 2017. "Automated Parameter Blocking for Efficient Markov Chain Monte Carlo Sampling." *Bayesian Analysis* 12 (2). International Society for Bayesian Analysis: 465–90.