Self-Exciting Spatio-Temporal Models for Count Data

Nicholas Clark

Iowa State University
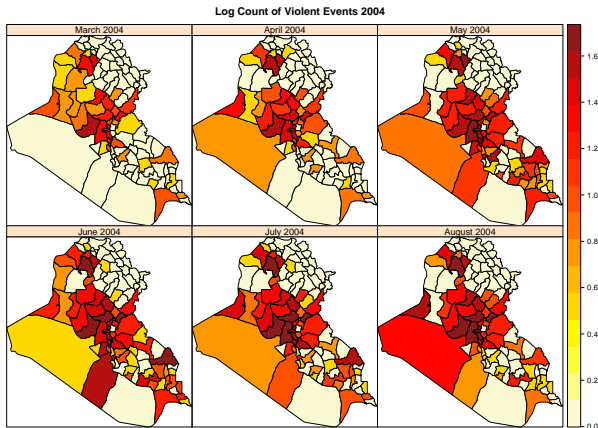
nclark1@iastate.edu

January 29, 2018

Motivation: The Evolution of Violence in Space and Time

At present, the most under-researched area of spatial criminology is that of spatio-temporal crime patterns... the temporal component of the underlying crime distributions has languished as a largely ignored area of study - **Crime mapping: Spatial and Temporal Challenges**, Ratcliffe (2010)

The Spread of Violence in Iraq 2004

Log Count of Violent Events 2004

Burglaries South Side of Chicago

Burglaries on South Side of Chicago Mar–May 2011

Goals
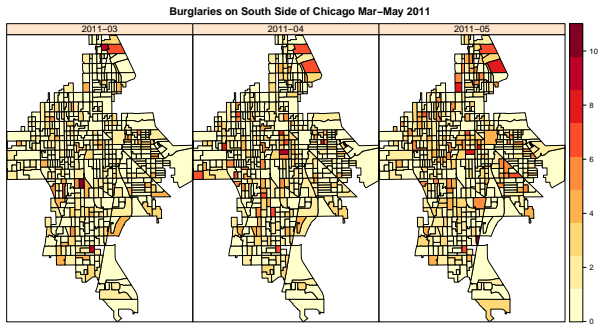
- General statistical model for diffusion of violence in space-time

  - Accurately reflects beliefs on how violence/crime evolves

  - Extends traditional statistical models for count data

  - Stationary with extremely flexible second order properties

  - Inference via traditional MCMC techniques

Overview

A Model of Criminal Behavior (Short et al. 2008)

- $Z(s_i, t)$ - number of observed burglaries from $(t - \Delta t, t)$
- $s_i \in \{s_1, \cdots, s_n\}$ - fixed regions in $\mathbb{R}^2$
- $t \in \{1, \cdots, T\}$ - discrete time
- Define $A(s_i, t) \equiv A(s_i, 0) + B(s_i, t)$ as attractiveness

$$B(s_i, t) = (1 - \chi \Delta t) B(s_i, t - \Delta t) + \eta Z(s_i, t - \Delta t) \qquad (1)$$

- Probability of occurrence at each time interval, $(t, t + \Delta t)$ is Poisson with rate, $A(s_i, t)$
- Three factors impact change in crime rate, base attractiveness , decay χ, and repeat victimization, η

Relationship to INGARCH Model

Integer Auto-Regressive Conditionally Heteroskedastic, INGARCH (1,1), or Poisson Auto-Regression Model

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right) \tag{2}$$

$$\lambda(s_i, t) = d + a\lambda(s_i, t-1) + bZ(s_i, t-1) \tag{3}$$

- Unlike GARCH, not solely a variance property
- Short model is similar to INGARCH(1,1) with $A(s_i, 0) = \sum_{k=0}^{t} a^k d$, $a = (1 - \chi\Delta t)$, and $b = \eta$

Relationship to Self-Exciting Models

- Point process introduced by Alan Hawkes with intensity

$$\lambda(t) = \nu(t) + \int_0^t g(t-u)N(ds) \tag{4}$$

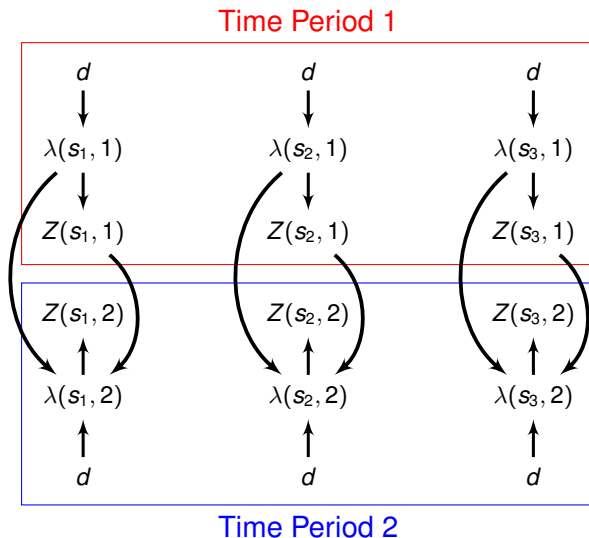- Commonly discretized as

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right) \tag{5}$$

$$\lambda(s_i, t) = \nu + \sum_{j<t} \eta^{t-j} Z(s_i, t-j) \tag{6}$$

- η essentially determines p in INGARCH$(0, p)$, $\eta < .5$ for stationarity

- Motivated by Reaction-Diffusion PDE
- Recall $A(s_i, t) \equiv A(s_i, 0) + B(s_i, t)$ as attractiveness, let N_{s_i} be spatial adjacency matrix

$$
B(s_i, t) = \kappa B(s_i, t - \Delta t) + \frac{\psi}{|N_{s_i}|} \sum_{s_j \in N_{s_i}} \left[B(s_j, t - \Delta t) - B(s_i, t - \Delta t) \right]
$$
$$
+ \eta Z(s_i, t - \Delta t) \tag{7}
$$

- Four factors impact change in crime rate, base attractiveness, decay χ, and repeat victimization, η, and spatial spread ψ
- Resulting model is MINGARCH (1,1)
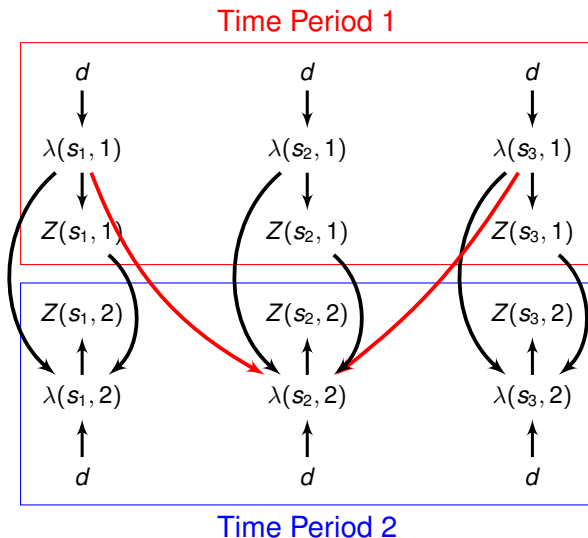
Applied to Residential Burglaries in Chicago

Burglaries on South Side of Chicago Mar–May 2011
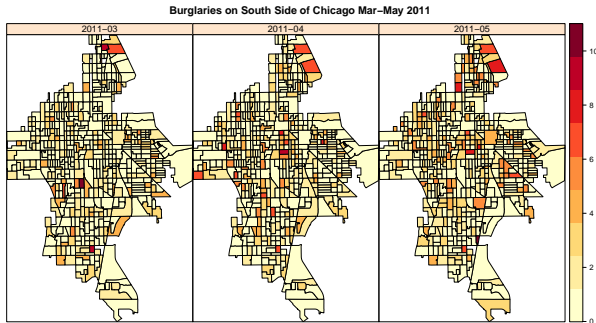
552 Spatial Locations, 72 Months, residential burglaries

$$B(s_i, t) = \kappa B(s_i, t - \Delta t)$$
$$+ \frac{\psi}{|N_{s_i}|} \sum_{s_j \in N_{s_i}} \left[B(s_j, t - \Delta t) - B(s_i, t - \Delta t) \right] + \eta Z(s_i, t - \Delta t) \quad (8)$$

- Further structure $A(s_i, 0)$ to account for socio-economic factors
- MLEs are $\hat{\psi} = .02$, $\hat{\eta} = .155$ $\hat{\kappa} = .773$
- Simulate data from asymptotic distribution - unable to replicate lag-one autocorrelation, spatial correlation, or variance to mean ratio of original data

Properties of INGARCH Model

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right)$$
$$\lambda(s_i, t) = d + a\lambda(s_i, t-1) + bZ(s_i, t-1)$$

Stationarity yields:

$$E[Z(s_i, t)] = \frac{d}{1 - (a+b)} \tag{9}$$

$$\text{Var}[Z(s_i, t)] = \frac{1 - (a+b)^2 + b^2}{1 - (a+b)^2} E[Z(s_i, t)] \tag{10}$$

$$\text{Cov}[Z(s_i, t), Z(s_i, t-h)] = \frac{b(1 - a(a+b))(a+b)^h}{1 - (a+b)^2} E[Z(s_i, t)] \tag{11}$$

$$\text{Var-Mean Ratio}[Z(s_i, t)] = 1 + \frac{b^2}{1 - (a+b)^2} \tag{12}$$

Issues

Allows for Overdispersion... But at a cost!

$$\text{Cor}[Z(s_i, t), Z(s_i, t-1)] = \frac{b(a+b)(a^2+ab-1)}{a^2+2ab-1} \qquad (13)$$

$$\text{Var-Mean Ratio}[Z(s_i, t)] = 1 + \frac{b^2}{1-(a+b)^2} \qquad (14)$$

$$(15)$$

For fixed Var-Mean Ratio at $2 \implies b = 1/2(-a + \sqrt{2-a^2})$.
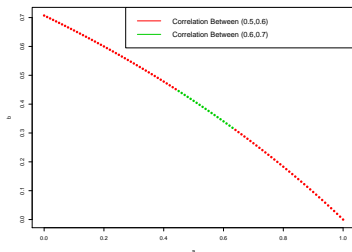
$Y(s_i, t)$ - Spatially Correlated Latent Gaussian
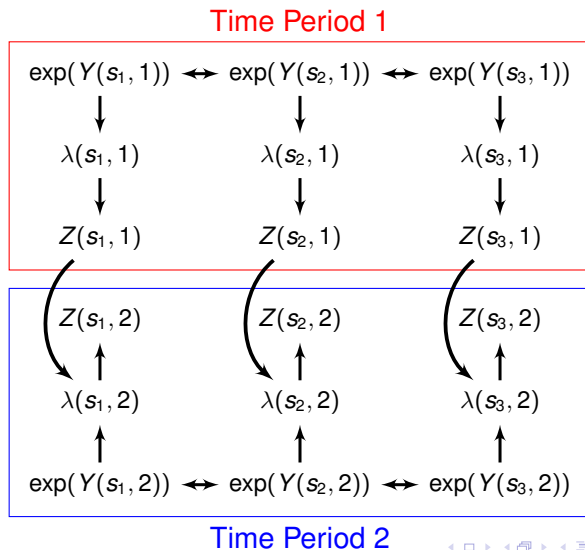
Time Period 1

$\exp(Y(s_1, 1)) \leftrightarrow \exp(Y(s_2, 1)) \leftrightarrow \exp(Y(s_3, 1))$

$\lambda(s_1, 1)$ $\lambda(s_2, 1)$ $\lambda(s_3, 1)$

$Z(s_1, 1)$ $Z(s_2, 1)$ $Z(s_3, 1)$

$Z(s_1, 2)$ $Z(s_2, 2)$ $Z(s_3, 2)$

$\lambda(s_1, 2)$ $\lambda(s_2, 2)$ $\lambda(s_3, 2)$

$\exp(Y(s_1, 2)) \leftrightarrow \exp(Y(s_2, 2)) \leftrightarrow \exp(Y(s_3, 2))$

Time Period 2

- **Theory:** Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation

- **Theory:** Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation
- Mixture of two processes that influence expectation : LGCP and Hawkes process
- Hawkes process letting $g(t - j) = \eta$ if $(t - j) = 1$, 0 otherwise
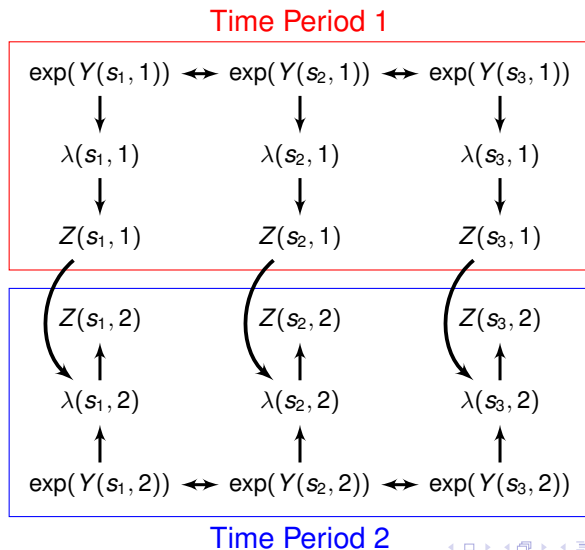
$$Z(\mathbf{s_i}, t) | \lambda(\mathbf{s_i}, t) \sim \text{Pois} \left(\lambda(\mathbf{s_i}, t) \right) \quad (16)$$
$$\lambda(\mathbf{s_i}, t) = \exp(Y(\mathbf{s_i}, t)) + \eta Z(\mathbf{s_i}, t - 1)$$
$$Y(\mathbf{s_i}, t) = \theta_1 \sum_{\mathbf{s_j} \in N(\mathbf{s_i})} Y(\mathbf{s_j}, t) + \epsilon(\mathbf{s_i}, t)$$
$$\epsilon(\mathbf{s_i}, t) \sim Gau(0, \sigma^2)$$

$Y(s_i, t)$ - Spatially Correlated Latent Gaussian

Time Period 1

$\exp(Y(s_1, 1)) \leftrightarrow \exp(Y(s_2, 1)) \leftrightarrow \exp(Y(s_3, 1))$

$\lambda(s_1, 1) \qquad \lambda(s_2, 1) \qquad \lambda(s_3, 1)$

$Z(s_1, 1) \qquad Z(s_2, 1) \qquad Z(s_3, 1)$

$Z(s_1, 2) \qquad Z(s_2, 2) \qquad Z(s_3, 2)$

$\lambda(s_1, 2) \qquad \lambda(s_2, 2) \qquad \lambda(s_3, 2)$

$\exp(Y(s_1, 2)) \leftrightarrow \exp(Y(s_2, 2)) \leftrightarrow \exp(Y(s_3, 2))$

Time Period 2

SPINGARCH(1,1) Model

Spatially Correlated INGARCH(1,1) Model

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim \text{Pois}(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i, t) = \exp[Y(s_i, t)] + \eta Z(s_i, t-1) + \kappa \lambda(s_i, t-1) \tag{17}$$

- Define: $N_i = \{s_j : s_j \text{ is a spatial neighbor of } s_i\}$

$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i, t), \sigma_{sp}^2) \tag{18}$$
$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j, t) - \alpha(s_j)\}.$$

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim$ Pois $(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i, t) = \exp[Y(s_i, t)] + \eta Z(s_i, t-1) + \kappa \lambda(s_i, t-1) \tag{17}$$

- Define: $N_i = \{s_j : s_j \text{ is a spatial neighbor of } s_i\}$

$$Y(s_i, t) | \mathbf{Y}(N_i) \sim N(\mu(s_i, t), \sigma_{sp}^2) \tag{18}$$
$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j, t) - \alpha(s_j)\}.$$

- $\eta = 0, \kappa = 0$ Poisson - CAR, $\sigma_{sp}^2 \to 0$, INGARCH(1,1)/Short model

$$\frac{\lambda(s_i, t) - \lambda(s_i, t-1)}{\Delta t} = d - \chi\lambda(s_i, t-1) + \eta Z(s_i, t-1) \qquad (19)$$

- Change in violence due to exogeneous d, natural decay, χ, and excitement, η
- Assume each time period, exogenous impact is stochastic and spatially correlated yields SPINGARCH

$$\frac{\lambda(s_i, t) - \lambda(s_i, t-1)}{\Delta t} = \exp(Y(s_i, t)) - \chi\lambda(s_i, t-1) + \eta Z(s_i, t-1) \qquad (20)$$

- Change in intensity due to three factors, CAR, natural decay, and excitement

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim$ Pois $(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

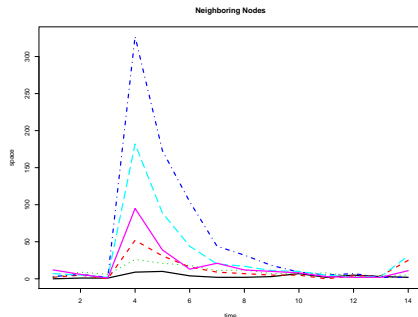$$\lambda(s_i, t) = \exp[Y(s_i, t)] + \eta Z(s_i, t - 1) + \kappa \lambda(s_i, t - 1)$$
$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i, t), \sigma_{sp}^2)$$
$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j, t) - \alpha(s_j)\}.$$

- $\zeta \in (\psi_{(1)}^{-1}, \psi_{(n)}^{-1})$ where $\psi_{(i)}$ is the ith largest eigenvalue of adjacency matrix
- For stationarity, $\eta > 0$, $\kappa > 0$, $\eta + \kappa < 1$

Data Realizations

50 Spatial Observations on $\mathbb{R}^1$, 100 Temporal Observations
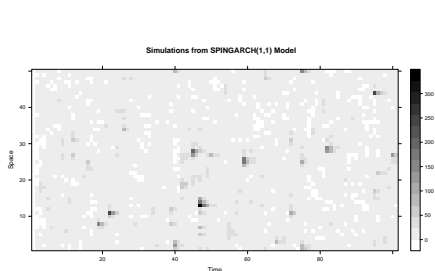
$$\lambda(s_i, t) = \exp\left[Y(s_i, t)\right] + 0.1\, Z(s_i, t-1) + 0.4\, \lambda(s_i, t-1)$$
$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim \text{Gau}(\mu(s_i, t), 0.5)$$
$$\mu(s_i, t) = 0 + 0.49 \sum_{s_j \in N_i} \{Y(s_j, t)\}. \tag{21}$$

SPINGARCH Model as Markov Chain

- Let $\boldsymbol{\lambda_t} = (\lambda(s_1, t), \lambda(s_2, t), \cdots, \lambda(s_{n_d}, t))^T$

$$Z(s_i, t)|\lambda(s_i, t) \sim \text{Pois}(\lambda(s_i, t))$$
$$E[Z(s_i, t)] = \lambda(s_i, t)$$
$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} + \kappa \boldsymbol{\lambda_{t-1}}$$
$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d, n_d} - \boldsymbol{C})^{-1} \boldsymbol{M}) \qquad (22)$$

- Markov chain for $\boldsymbol{\lambda_t}$ on State space, $(\mathbb{R}^+)^{n_d}$

Impact of Initial Condtions and Recursion

By recursion

$$[\lambda(s_i, t)|\lambda(s_i, 0) = B] = \exp(Y(s_i, t)) + \kappa\lambda(s_i, t-1) + \eta Z(s_i, t-1)$$
$$= \exp(Y(s_i, t)) + \kappa\left[\exp(Y(s_i, t-1)) + \kappa\lambda(s_i, t-2)\right.$$
$$\left. + \eta Z(s_i, t-2)\right] + \eta Z(s_i, t-1)$$
$$\cdots$$
$$= \sum_{k=0}^{t-1} \kappa^k \exp(Y(s_i, t-k)) + \sum_{k=0}^{t-1} \kappa^k \eta Z(s_i, t-k-1) + \kappa^t B.$$

Geometric Ergodicity with Finite Moments

Under the parameter space restriction, $\eta, \kappa \geq 0$ and $\eta + \kappa < 1$, the SPINGARCH (1,1) is geometrically ergodic and admits a unique stationary distributions that has finite first two moments.

Sketch of Proof.

Meyn and Tweedie (15.0.1) need to show aperiodic, ϕ-irreducible and $\exists$ test function $V(.)$ such that
$E[V(\lambda_{t+1})|\lambda_t = \lambda_*] \leq \psi V(\lambda_*) + L\, \mathsf{I}(\lambda_* \in C)$ holds where $\psi \in (0, 1)$, $L \in (0, \infty)$ and $I(.)$ is the indicator function and C is a petite set.

Basic Idea: With positive probability, $\exists$ a realization
$Z(s_i, 1) = Z(s_i, 2) = \cdots = Z(s_i, t-1) = 0$. Along that chain,
$P(\lambda(s_i, t)) \in A = P(\exp(Y(s_i, t)) \in A - \kappa^t B)$. If $\kappa^T B > \sup A$ run chain longer. $\qquad\square$

Geometric Ergodicity with Finite Moments

Under the parameter space restriction, $\eta, \kappa \geq 0$ and $\eta + \kappa < 1$, the SPINGARCH (1,1) is geometrically ergodic and admits a unique stationary distributions that has finite first two moments.

Sketch of Proof Cont.

Test function $V(\lambda) = 1 + \lambda^2$ works for
$E[V(\lambda_{t+1})|\lambda_t = \lambda_*] \leq \psi V(\lambda_*) + L \, \mathrm{I}(\lambda_* \in C).$

$\implies$ Unique stationary distribution, goes to geometrically fast.
Specific choice of $V(.)$ gives (at least) finite first two moments (can be extended likely as in Fokianos, 2009.)

$\square$

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right)$$
$$\lambda(s_i, t) = \exp\left[Y(s_i, t)\right] + \eta Z(s_i, t-1) + \kappa E\left[Z(s_i, t-1)\right]$$
$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i, t), \sigma_{sp}^2)$$
$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \left\{Y(s_j, t) - \alpha(s_j)\right\}$$

Define $\Sigma_{i,j}$ as i, j entry of $(I_{n_d, n_d} - \boldsymbol{C})^{-1}\boldsymbol{M}$

$$E\left[Z(s_i, t)\right] = \frac{1}{1 - \eta - \kappa} \exp(\alpha + \frac{\Sigma_{1,1}}{2}) \tag{23}$$

$$\text{Var}\left(Z(s_i, t)\right) = \frac{1}{1 - (\kappa + \eta)^2}\text{Var}\left(\exp(Y(s_i, t))\right) + \frac{1 - \kappa^2 - 2\kappa\eta}{1 - (\kappa + \eta)^2}E(Z(s_i, t)) \tag{24}$$

Temporal Covariance:

$$\text{Cov}\left(Z(s_i, t), Z(s_i, t-1)\right) = (\eta + \kappa)\text{Var}(Z(s_i, t)) - \kappa E\left[Z(s_i, t)\right] \quad (25)$$

Temporal Covariance:

$$\text{Cov}\left(Z(s_i, t), Z(s_i, t-1)\right) = (\eta + \kappa)\text{Var}(Z(s_i, t)) - \kappa E\left[Z(s_i, t)\right] \quad (25)$$

Recall: Var-Mean Ratio at 2 $\implies$ Lag-1 correlation is between 0.5 and $\sqrt{1/2}$ for INGARCH(1,1)

Increased Modeling Flexibility with SPINGARCH(1,1)

Temporal Covariance:

$$\text{Cov}\left(Z(s_i, t), Z(s_i, t-1\right) = (\eta + \kappa)\text{Var}(Z(s_i, t)) - \kappa E\left[Z(s_i, t)\right] \quad (25)$$

Recall: Var-Mean Ratio at 2 $\implies$ Lag-1 correlation is between 0.5 and $\sqrt{1/2}$ for INGARCH(1,1)

Let $\kappa = 0$ (SPINGARCH(0,1)), Var-Mean Ratio at 2

$$\implies 2 = \frac{\text{Var}\left(\exp(Y(s_i, t))\right)}{(1 - \eta)^2 E\left[Z(s_i, t)\right]} + \frac{1}{1 - \eta^2} \quad (26)$$

$$\text{Cor}\left(Z(s_i, t), Z(s_i, t-1)\right) = \eta \quad (27)$$

$\forall \eta \in (0, \sqrt{1/2}) \quad \exists \alpha, \sigma_{sp}^2$ such that equality holds

Spatial Correlation

$$\text{Corr}(Z(s_i, t), Z(s_j, t)) = \frac{\left(\exp(\Sigma_{i,i} + \Sigma_{i,j}) - \exp(\Sigma_{i,i})\right)}{\exp(2\Sigma_{i,i}) - \exp(\Sigma_{i,i}) + \exp(-\alpha + \frac{\Sigma_{i,i}}{2})\frac{1}{1-(\kappa+\eta)}}$$

Spatial Correlation For Differing Spatial Parameter Values
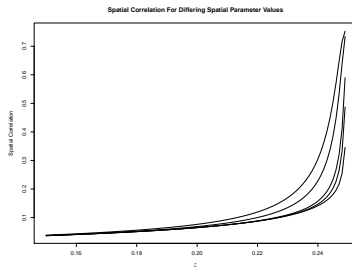
$\eta = .3$, $\sigma_{sp}^2 = .5$, 4×4 to 15×15 size lattice

Inference

- Likelihood roots for INGARCH(1,1) easily found, asymptotically Gaussian
- Inclusion of latent process in SPINGARCH(1,1) complicates
- $\boldsymbol{\theta} \equiv (\eta, \alpha, \zeta, \sigma_{sp}^2)$

$$\pi(\boldsymbol{\theta}|\boldsymbol{Z}, \boldsymbol{Y}) \propto \prod_t \pi(\boldsymbol{Z}_t|\boldsymbol{\lambda}_t)\pi(\boldsymbol{\lambda}_t|\boldsymbol{\lambda}_{t-1}, \boldsymbol{Z}_{t-1}, \boldsymbol{\theta}, \boldsymbol{Y}_t)\pi(\boldsymbol{Y}_t|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \quad (28)$$

$$\pi(\boldsymbol{Y}|\boldsymbol{Z}, \boldsymbol{\theta}) \propto \prod_t \pi(\boldsymbol{Z}_t|\boldsymbol{\lambda}_t)\pi(\boldsymbol{\lambda}_t|\boldsymbol{\lambda}_{t-1}, \boldsymbol{Z}_{t-1}, \boldsymbol{\theta}, \boldsymbol{Y}_t)\pi(\boldsymbol{Y}_t|\boldsymbol{\theta}). \quad (29)$$

$$\log(\boldsymbol{Y}|\boldsymbol{\alpha}, \sigma_{sp}, \zeta) \propto \frac{1}{2}\log|\Sigma_f^{-1}(\theta)|$$
$$-\frac{1}{2}(Y-\alpha)^T\Sigma_f^{-1}(\theta)(Y-\alpha), \tag{30}$$

- $\Sigma_f^{-1} \equiv \left(I_{n_d \times T, n_d \times T} - I_{t,t} \otimes \boldsymbol{C}\right)^{-1} I_{t,t} \otimes \boldsymbol{M}$
- $\log|\Sigma^{-1}(\theta)| = \frac{n_d}{2\log\sigma_{sp}^2} + \log|I_{n_d,n_d} - \zeta N|$
- Letting $V\Lambda V^T$ be the spectral decomposition of N we have $|I_{n_d,n_d} - \zeta N| = |V||I_{n_d,n_d} - \zeta\Lambda||V^T| = \prod_{j=1}^{n_d}\left(1 - \zeta\chi_j\right)$ where χ_j are the eigenvalues of the neighborhood matrix

$$\log|\Sigma_f^{-1}(\theta)| = T\log|\Sigma^{-1}(\theta)| \tag{31}$$

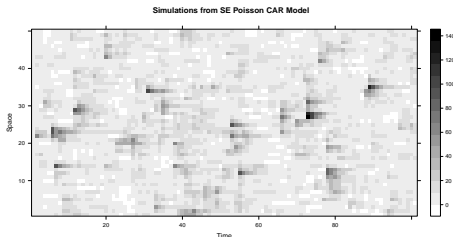$$\propto \frac{n_d \times T}{\log\sigma_{sp}^2} + T\sum_{j=1}^{n_d}(1 - \zeta\chi_j) \tag{32}$$

Simulation and Estimation

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t))$$
$$\lambda(s_i, t) = \exp[Y(s_i, t)] + 0.66Z(s_i, t - 1)$$
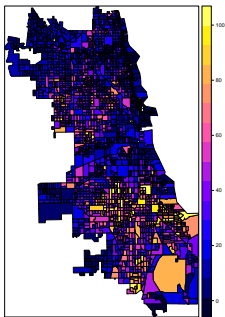$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim \text{Gau}(\mu(s_i, t), 0.5)$$
$$\mu(s_i, t) = 0 + 0.49 \sum_{s_j \in N_i} \{Y(s_j, t)\}. \tag{33}$$

Simulations from SE Poisson CAR Model

Simulation and Estimation

- SPINGARCH(0,1) 95% credible intervals: $\alpha \in (-0.24, 0.1)$, $\sigma^2 \in (0.46, 0.59)$, $\zeta \in (0.486, 0.492)$, and $\eta \in (0.64, 0.66)$
- SPINGARCH(1,0) 95% credible intervals: $\alpha \in (-0.54, -0.2)$, $\sigma^2 \in (0.96, 1.2)$, $\zeta \in (0.47, 0.48)$, and $\kappa \in (0.65, 0.67)$

	SPINGARCH(1,0)	SPINGARCH(0,1)
p_1 - Moran's I	.05	.46
p_2 - Var to Mean	.99	.65
p_3 - Lag 1 Corr	.45	.7

Burglaries in South Side of Chicago

Aggregated Burglaries
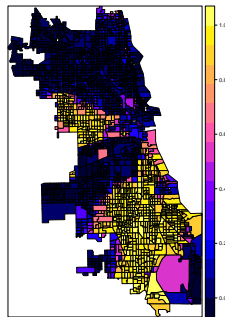
Racial Segregation

Burglaries South Side of Chicago

Burglaries on South Side of Chicago Mar–May 2011
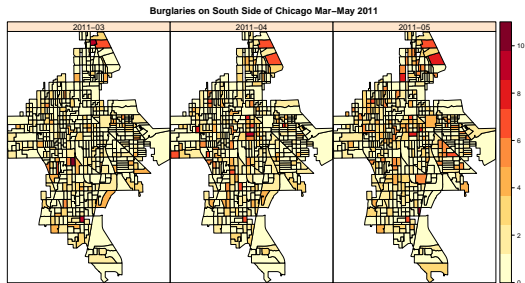
- Crime data from city of Chicago
- 72 months (2010-2015), 552 locations (Census block groups)
- Demographic data from Census bureau

SPINGARCH(1,1) Model

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t)) \tag{34}$$
$$E[Z(s_i, t)] = \lambda(s_i, t)$$
$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t} + \boldsymbol{U}) + \eta \boldsymbol{Z_{t-1}} + \kappa \boldsymbol{\lambda_{t-1}}$$
$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha}, \sigma_{ind}^2 \boldsymbol{I}_{n_d, n_d})$$
$$\boldsymbol{U} \sim \text{Gau}(\boldsymbol{0}, [\sigma_{sp}^2(\boldsymbol{N} - \boldsymbol{C})]^{-1})$$

- Removed temporal trend and seasonality
- Single spatial effect, ζ fixed near edge of parameter space
- Additional small scale effect captured in σ_{ind}^2

$$\alpha_{s_i} = \exp\big(\beta_0 + \beta_{pop} \log(\text{Pop}_{s_i})$$
$$+ \beta_{ym} \text{Young Men}_{s_i} + \beta_{wealth} \text{Wealth}_{s_i} + \beta_{unemp} \text{Unemp}_{s_i}\big) \tag{35}$$

Impacts of Including Spatial Correlation

Parameter	SPINGARCH(1,1)	INGARCH(1,1)
β_0	(-3.8,-2.2)	(-4.2,-3.4)
β_{pop}	(0.12,0.34)	(0.33,0.46)
β_{ym}	$(-0.77, 0.16)$	$(0.06, 0.09)$
β_{wealth}	$(0.04, 0.16)$	$(-0.04, 0.01)$
β_{unemp}	(0.005,0.07)	(0.002,0.03)
η	$(0.02, 0.05)$	$(0.22, 0.24)$
κ	(0.42,0.49)	(0.44,0.48)
σ_{sp}^2	(0.36,0.50)	-
σ_{ind}^2	(0.49,0.57)	-

	SPINGARCH(1,1)	INGARCH(1,1)
p_1 - Moran's I Statistic	0.20	0
p_2 - Variance to Mean Ratio	0.42	0
p_3 - Lag 1 Auto Correlation	0.52	0.74

- INGARCH(1,1)/Short Model unable to replicate second order properties
- Correct Lag 1 Auto Cor dictates incorrect Var/Mean ratio
- Conclusions on self-excitement

Summary

- Addition of spatially correlated effects naturally extends INGARCH model
- Both Poisson-CAR and INGARCH arise from SPINGARCH in limit
- Failure to specify random structure may result in differing conclusions
- Precomputing eigenvalues allows for relatively efficient Bayesian inference (8 hrs for 552 spatial locations, 104 time locations)

Future Work

- Impacts of aggregation
- Laplace approximations greatly speed up SPINGARCH(0,1) - Can extend to SPINGARCH(1,1)?
- Reaction Diffusion Self-Exciting Model from (Clark & Dixon, 2018) does not fit in framework (temporally correlated errors)
    - RDSEM captures reaction diffusion process of Short in Latent Process
- Dropping self-excitement leads to SPDE with exact solution - Sparse approximation?
- Questions?

Post Talk Slides

- Research Chapter 1 - SCSEM and RDSEM Models applied to Iraq Data, Laplace Approximation based exploration of posterior density of parameters
- Research Chapter 2 - Extend SCSEM and put in context of other statistical models, apply to Burglaries in Chicago
- Research Chapter 3 - Discovered bias in Laplace approximation for subset of parameter space. Explains why bias occurs and how to fix (in some instances)

Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise **The Future of Data Analysis.** Annals of Mathematical Statistics 33: 1-67 John W. Tukey 1962

SPINGARCH (0,1) - Spatially Correlated Self-Exciting Model

Previously used in inference on violence in Iraq

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t)) \tag{36}$$

$$E[Z(s_i, t)] = \lambda(s_i, t) \tag{37}$$

$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} \tag{38}$$

$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d, n_d} - \boldsymbol{C})^{-1} \sigma_{sp}^2). \tag{39}$$

Integration Free Technique

Laplace Approximation to marginals

$$\tilde{\pi}(\eta, \zeta, \sigma_{sp}^2, \alpha | \mathbf{Z}) \propto \frac{\pi(\mathbf{Z}|\eta, \mathbf{Y})\pi(\mathbf{Y}|\alpha, \zeta, \sigma_{sp}^2)\pi(\zeta)\pi(\alpha)\pi(\sigma_{sp}^2)\pi(\zeta)}{\pi_G(\mathbf{Y}|\alpha, \eta, \zeta, \sigma_{sp}^2, \mathbf{Z})}, \quad (40)$$

$\pi_G(\mathbf{Y}|\alpha, \eta, \zeta, \sigma_{sp}^2, \mathbf{Z})$ is Gaussian approximation from Taylor series expansion of full conditional

$$\pi_G(\mathbf{Y}|\eta, \zeta, \sigma_{sp}^2, \mathbf{Z}) \propto (2\pi)^{n/2} \det(\Sigma(\theta))^{1/2} \exp\left(-\frac{1}{2}(\mathbf{Y})^t \Sigma^{-1}(\theta)\mathbf{Y} \right.$$
$$\left. + \sum_{s_i, t} f(\mu(s_i, t))(Y(s_i, t)) + 1/2 k(\mu(s_i, t))(Y(s_i, t))^2\right)$$
$$(41)$$

Issues

Evaluate $\pi_G(\boldsymbol{Y}|\eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})$ at mode $\boldsymbol{Y} = \boldsymbol{\mu}^*$:

$$\log(\pi(\boldsymbol{\theta}|\boldsymbol{Z})) \propto \log \pi(\boldsymbol{Z}|\eta, \boldsymbol{\mu}^*) + \log \pi(\boldsymbol{\mu}^*|\zeta, \sigma^2_{sp}, \alpha) - 1/2 \log |\Sigma(\boldsymbol{\theta})|$$
$$+ \log \pi(\boldsymbol{\theta}) + 1/2 \log |\Sigma(\boldsymbol{\theta}) + \text{diag } k(\mu(\boldsymbol{s}_i, t)^*)| \tag{42}$$

- Even for small T, $\pi(\theta|\boldsymbol{Z})$ appears to be Gaussian
- $\pi_G(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})$, not a good approximation
- σ^2_{sp} in particular appears biased (low)
- Bias increases as η or σ^2_{sp} increases

Likelihood

Full likelihood has latent dimensionality that increases as *n* increases:

$$L(\eta, \alpha, \zeta, \sigma_{sp}^2 | \boldsymbol{Z}) \propto \int_{\boldsymbol{\Omega}_y} \prod_{i=1}^{n} \prod_{t=1}^{T} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t)))$$
$$\times (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i,t)} \, d\mu_{\boldsymbol{Y}}. \quad (43)$$

Given data, $\boldsymbol{Z}$:

$$L(\eta, \alpha, \zeta, \sigma_{sp}^2 | \boldsymbol{Z}) \propto \prod_{t=1}^{T} \int_{\boldsymbol{\Omega}_{y_t}} \prod_{i=1}^{n} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t)))$$
$$\times (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i,t)} \, d\mu_{\boldsymbol{Y_t}} \quad (44)$$

Still high dimensional, intractable

Standard Laplace Approximation

$$\int_{\mathbf{\Omega}_{y_t}} \prod_{i=1}^{n} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t)))$$

$$\times \; (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i, t)} \, d\mu_{\mathbf{Y_t}}$$

Let

$$g(\mathbf{Y_t}) = \log \pi(\mathbf{Y_t}|\boldsymbol{\theta}) + \log(\mathbf{Z_t}|\boldsymbol{\theta}, \mathbf{Z_{t-1}}) \tag{45}$$

$$\frac{\partial g}{\partial Y(s_i, t)} = g_i \tag{46}$$

$$\frac{\partial g}{\partial Y(s_i, t) \partial Y(s_j, t)} = g_{ij} \tag{47}$$

$$g_{\mathbf{YY}} = \text{Hessian matrix of } g \tag{48}$$

$$g^{ij} = (i, j)\text{element of inverse Hessian} \tag{49}$$

$$\tag{50}$$

$$M_t = \int \exp(-g(\boldsymbol{Y}_t))dY_t$$

- Multivariate Taylor Series expansion of g about unique minimum followed by Taylor Series expansion of exp about zero yields

$$\boldsymbol{M} = \exp(-\hat{g})\left|\frac{\hat{g}_{\boldsymbol{Y}\boldsymbol{Y}}}{2\pi}\right|E\left[1 - \frac{1}{3!}\hat{g}_{i,j,k}U(s_1,t)U(s_2,t)U(s_3,t)\right.$$
$$\left. -\frac{1}{4!}\hat{g}_{i,j,k,l}U(s_1,t)U(s_2,t)U(s_3,t)U(s_4,t) - \cdots\right] \quad (51)$$

- $\boldsymbol{U} \sim \text{Gau}(\boldsymbol{0}, \hat{g}_{\boldsymbol{Y}\boldsymbol{Y}})$
- $\hat{g}_{i,j,k,l} = 0$ unless $i = j = k = l$, $E[U(s_1,t)^4] = 3(g^{ii})^2$

$$\frac{1}{4!}\hat{g}_{i,j,k,l}E[U(s_1,t)U(s_2,t)U(s_3,t)U(s_4,t)] = \frac{1}{8}\sum_i \hat{g}_{iiii}(\hat{g}^{ii})^2 \tag{52}$$

$$= -\frac{1}{72}\sum_{i,j\leq i}\hat{g}_{iii}\hat{g}_{jjj}\left(6\left(\hat{g}^{ij}\right)^3 + 9\hat{g}^{ii}\hat{g}^{jj}\hat{g}^{ij}\right) \tag{53}$$

$$= \frac{1}{48}\sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 \tag{54}$$

Issues when $\hat{g}^{ii} > 1$

Terms increase as η and σ_{sp}^2 increase

Extended Laplace Approximation

(Shun & McCullagh, 1995) and (Evangelou et al. 2011)

$$M_t = \int \exp(-g(\mathbf{Y}_t)) dY_t$$

- Taylor series expansion of $g(.)$, $\mathbf{M_t} = \exp(-\hat{g}) E \exp(-\hat{g}_i Y(s_i) - \hat{g}_{ij} Y(s_i) Y(s_j)/2! - \hat{g}_{ijk} Y(s_i) Y(s_j) Y(s_k)/3! - \cdots)$
- $\log M_t$ is joint cumulant-generating function of $Y(s_i)$, $Y(s_i) Y(s_j)$, $Y(s_i) Y(s_j) Y(s_k)$, etc.
- Using Shun & McCullagh (1995) for $\log M$

$$\log M \propto -\hat{g} - \frac{1}{2}|\hat{g}_{\mathbf{YY}}| - \sum_t \sum_i \frac{1}{8}\hat{g}_{iiii}(\hat{g}^{ii})^2 -$$
$$\sum_t \sum_i \frac{1}{48}\hat{g}_{iiiiii}(\hat{g}^{ii})^4 + \frac{1}{72}\sum_t \sum_{i,j \leq i} \hat{g}_{iii}\hat{g}_{jjj}\left(6\left(\hat{g}^{ij}\right)^3 + 9\hat{g}^{ii}\hat{g}^{jj}\hat{g}^{ij}\right)$$

$$\tilde{\pi}(\eta, \zeta, \sigma_{sp}^2, \alpha | \boldsymbol{Z}) \propto \frac{\pi(\boldsymbol{Z}|\eta, \boldsymbol{Y})\pi(\boldsymbol{Y}|\alpha, \zeta, \sigma_{sp}^2)\pi(\zeta)\pi(\alpha)\pi(\sigma_{sp}^2)\pi(\zeta)}{\pi_E(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma_{sp}^2, \boldsymbol{Z})}, \quad (55)$$

Algorithm 1 pseudocode for the calculation of marginal density of $\boldsymbol{\theta}$

1: Fix $\boldsymbol{\theta} = \boldsymbol{\theta_m}$ (Near Moment Based Estimates of $\boldsymbol{\theta}$)
2: Repeatedly solve $\left(\Sigma^{-1}(\boldsymbol{\theta}) + \text{diag } g^{ii}(\mu_n)\right)\mu_{n+1} = f(\mu_n)$, yielding $\boldsymbol{\mu}^*$ mode of $\pi_G(\boldsymbol{Y}|\boldsymbol{\theta})$
3: Evaluate $g_{iii}, g_{iiii}, g_{iiiiii}, g^{ii}, g^{ij}$ at $\boldsymbol{\mu}^*$ and $\log \pi(\boldsymbol{\theta})$ to approximate $\log M$
4: Numerically estimate Hessian of $\boldsymbol{\theta}$
5: Conduct Newton-Raphson update of $\boldsymbol{\theta}$
6: Repeat until convergence

When it Works

10×10 Lattice, 100 Time Points, $\zeta = .245, \alpha = 0$

	$\eta = .1, \sigma_{sp}^2 = .4$	$\eta = .4, \sigma_{sp}^2 = .6$
Relative Bias in LA(1)	.12	.2
Time to Fit LA(1) (min.)	10-15	16-20
Extended LA Without 6th Order	.03	.1
Extended LA With 6th Order	.03	.05
Time to Fit Extended LA	20-30	20-30
MCMC	.02	.02
Time to Fit MCMC	150-250	400-650

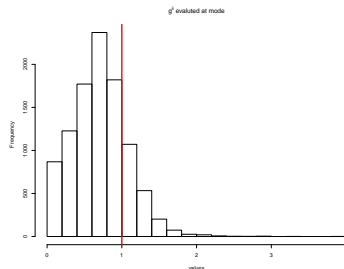|	$\eta = .7, \sigma_{sp}^2 = 1$
Relative Bias in LA(1)	.46
Time to Fit LA(1) (min.)	10-15 16-20
Extended LA Without 6th Order	.2
Extended LA With 6th Order	.2
Time to Fit Extended LA	25-35
MCMC	.06
Time to Fit MCMC	500-650

- Examine $\frac{1}{48} \sum_i \hat{g}_{iiiiii} (\hat{g}^{ii})^4$, $\frac{1}{8} \sum_i \hat{g}_{iiii} (\hat{g}^{ii})^2$
- Higher order terms $(\hat{g}^{ii})^k$ for $k > 2$
- $\sigma_{sp}^2 = 1, \eta = .7$ yield $\hat{g}^{ii} > 1$

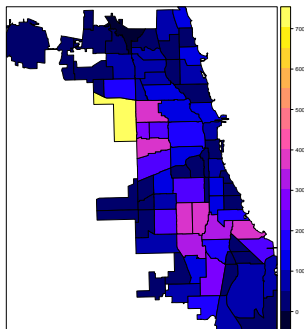$\hat{g}^{ii}$ terms for $\eta = .4$, $\sigma_{sp}^2 = .6$ and $\eta = .7$, $\sigma_{sp}^2 = 1$
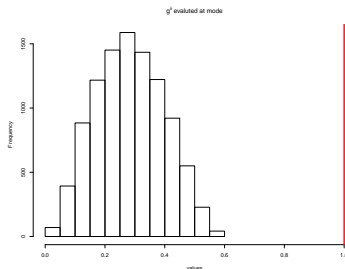
$$\frac{1}{48} \sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 = -41$$

$$\frac{1}{48} \sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 = -440$$

Weekly from (December 28 2014 - January 2, 2016)

Model

SPINGARCH(0,1) or Spatially Correlated Self-Exciting model

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t)) \tag{56}$$
$$E[Z(s_i, t)] = \lambda(s_i, t) \tag{57}$$
$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} \tag{58}$$
$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d, n_d} - C)^{-1} M) \tag{59}$$

- Assume both spatial and temporal covariates
  $\alpha(s_i, t) = \beta_0 + \beta_{temp} \, \text{Temp}(t) + \beta_{pop} \, \text{Pop}(s_i)$
- Vague proper priors for all parameters

Posterior Median Estimates

Point Estimates	σ_{sp}^2	ζ	η	β_0	β_1	β_2
LA	.38	.180	.50	-5.6	.17	.50
Extended LA	.52	.179	.50	-5.6	.18	.49
MCMC	.50	.179	.50	-5.6	.18	.49

- Extended LA fit without 6th order term
- Stan - 3 chains, 15000 samples, no evidence of non-convergence, 3 hours run in parallel
- Extended LA/LA - 15 minutes

Profile of σ_{sp}^2 for differing η values
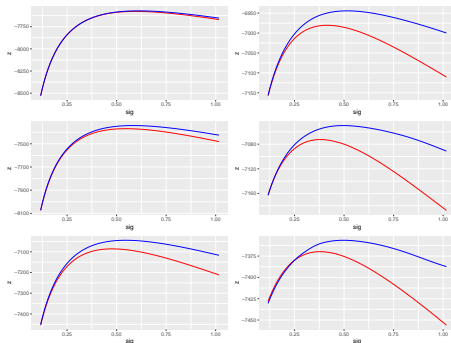
Red is LA(1), Blue is Extended LA, $\eta \in \{0, .1, .3, .5, .7, .9\}$

Comparison of 95% Credible Intervals

	σ^2_{sp}	ζ	η
Extended LA	(.43,.61)	(.176,.182)	(.47,.53
MCMC	(.42,.59)	(.176,.182)	(.47,.53)

	β_0	β_1	β_2
Extended LA	(-6.3,-4.9)	(.09,.27)	(.42,.55)
MCMC	(-6.3,-5.0)	(.09,.27)	(.42,.56)

Summary

- Essential to examine magnitude of $\hat{g}^{ii}$ terms, if large, LA will have non-negligible issues
- Extending Laplace approximation removes bias over wide range of parameter space and credible intervals comparable to MCMC
- Pay price in front end coding and derivations

Dissertation Contributions

- Models for violence that are consistent with sociological theories on how violence spreads
- New class of models that extends existing INGARCH models to spatial-temporal problems and accurately capture beliefs on how violence and crime spreads
- Methodology for inference that is quick and relatively accurate

- Extended LA for SPINGARCH(1,1)
- Theory for RDSEM models (INGARCH with spatio-temporally correlated latent structure)
- Impacts of aggregation (SPDE Approach)