Self-Exciting Point Processes for Crime

Craig Gilmour Prof. D. Higham

University of Strathclyde

craig.gilmour@strath.ac.uk

Monday 17th March

Why Self-exciting Point Processes?

- Studies have shown certain types of crime, e.g. burglaries and gang violence are more likely to occur soon after a previous crime has been committed.
- 'Repeat victimisation' and 'near repeat victimisation' offences target past victims or occur in nearby locations.
- Self-exciting point processes used in earthquake modelling (e.g. Vere-Jones 1970) aftershocks occur in the vicinity of the primary earthquake.
- These types of models have been more recently used in the context of crimes (e.g. Mohler 2011).

Self-exciting Point Processes

Self-exciting point processes can be defined by the *conditional intensity function*

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathbb{E}(N(t, t + \Delta t)|\mathcal{H}_t)}{\Delta t}.$$

ntensity

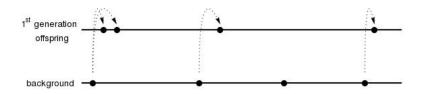
Poisson process

$$\lambda(t)=\mu.$$

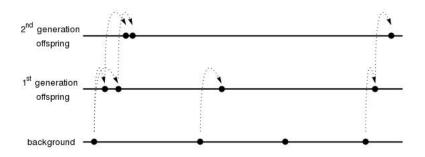


Hawkes process (Hawkes 1971) $\lambda(t) = \mu + \sum_{t_i < t} g(t - t_i).$

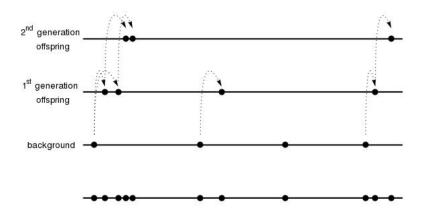




Self-Exciting Point Process



Self-Exciting Point Process



Parametric Formulation

Often a parametric form is assumed for the triggering function, e.g.

$$g(t - t_i) = \alpha \omega e^{-\omega(t - t_i)}$$
$$\lambda(t) = \mu + \sum_{t_i < t} \alpha \omega e^{-\omega(t - t_i)}$$

Given a point process $(t_1, t_2, ..., t_n)$ on an interval [0, T), the likelihood function is given by

$$L = \left(\prod_{i=1}^n \lambda(t_i)\right) \exp\left(-\int_0^T \lambda(t) \, dt\right).$$

For the above ETAS model the log-likelihood can be given as

$$\log L = \sum_{i=1}^{n} \log \left(\mu + \alpha \omega \sum_{t_j < t_i} e^{-\omega(t_i - t_j)} \right) - \mu T + \sum_{i=1}^{n} \alpha (e^{-\omega(T - t_i)} - 1).$$

Often the log-likelihood function is very flat, meaning standard iterative methods may not be suitable.

The model can be viewed instead as an incomplete data problem where we don't know whether an event is a background event or has been triggered by a previous event.

For the above ETAS model the log-likelihood can be given as

$$\begin{split} \log L &= \sum_{i=1}^{n} p_{ii} \log(\mu(t_i)) - \int_{0}^{T} \mu(t) \mathrm{d}t \\ &+ \sum_{i=1}^{n} \sum_{t_i < t_i} p_{ij} \log(g(t_i - t_j)) - \sum_{i=1}^{n} \int_{t_j}^{T} g(t - t_i) \mathrm{d}t \end{split}$$

EM-algorithm

For a spatio-temporal point process

$$\lambda(x, y, t) = \mu(x, y) + \sum_{t>t_i} g(t - t_i) f(x - x_i, y - y_i)$$
$$g(t)f(x, y) = \alpha \omega e^{-\omega t} \cdot \frac{1}{2\pi\sigma^2} e^{-(x^2 + y^2)/2\sigma^2}$$

We can create the following EM-algorithm

$$\mu = \frac{\sum_{i=1}^{n} p_{ii}}{T}, \quad \alpha = \frac{\sum_{i=1}^{n} \sum_{j=1}^{i-1} p_{ij}}{n},$$
$$\omega = \frac{\sum_{i=1}^{n} \sum_{j=1}^{i-1} p_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{i-1} p_{ij}(t_i - t_j)},$$
$$\sigma^2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{i-1} p_{ij}((x_i - x_j)^2 + (y_i - y_j)^2)}{\sum_{i=1}^{n} \sum_{j=1}^{i-1} 2p_{ij}(t_i - t_j)}.$$

Image: Image:

э

We iterate between this maximisation step, and the following expectation step until convergence

$$p_{ii} = rac{\mu}{\mu + \sum_{j=1}^{i-1} g(t_i - t_j)}, \ p_{ij} = rac{g(t_i - t_j)}{\mu + \sum_{j=1}^{i-1} g(t_i - t_j)}.$$

Crime Data

Dine:-2015 passet			Clines - 20	inpaset _ × +					
D · https://doi.org/whitege.org/Tuble-Selety/Crime-2001-to-present-Map/Mep-edm	D 0 Card	수 승 후 후 후 🗗 시 🚍 🖸 -	€ 0.4 hts	which also display and the	Safety/Orimes-2001-to-pr	0475/32-982		ci e escon	0 0 0 4 A 🗗 A 🗏
City of Chicago Deserves	🛞 City of Chicago tourous					· D . How And he links by Develops "Averation DystDates lights I			
Jesendo View See 42. Root Law Sector State	(and) Contract (B) (See The	Tradic International Contract	Climes -	2001 to present I reflects reported incidents of clim	a failth the exception of murd	ers where data exists for each	uction that P		2 II B B B Q Fini a Dis Odesel
	anatore		0	O II Case Number O I	1 Date 0 1	Dock O	NOR	O II Primary Type	O II Description
A Opr Staandung Des Players		Reading Dentile Territ		NUMBER AND	EXCOUNT II IN SUPE	62810 N 2082-31	0400	INTERV.	DOMESTIC INCTITY SMPLE
Sull Days		1	2.0	10000000 34200771		1000 3 (DRIVEN AND	0400	PUTER	DOMESTIC INCIDENT DAPLE
THE CONDUCTORY			2.00	10000074 14200014	\$3090011711A338PM	64710 10753 LERION 415	1416	NE4PON9 HOL4DON	UNLASTICE POSS OF HANDOURS
0 0 0 0		1		10095854 34296779	600002117 11 A1 31 PM	CONTRACTOR DATA	2022	NARCOTICS	P009-000446
Gilbarias Casif Bears		1	5 12	900064TA 34297955	655902117 H AD 31 PM	640KR W 318T 5T	6004	ORWISEGIAL ASSAULT	NON-AGGRANITED
S floor S & Bolow	NO 10.	1	4.00	40686585 JA206700	65592911115638PM	OSBID S AGA ST	1453	MEMPONS HOL/DOM	POST-FREEPAINABO NO FOE-CARD
		1	7.00	TORMERA LAQUENT	ESTOCUTT 11 38 30 PM	CT213 IL LOOME BLID	2000	NERCOTICS	POUND BURPROTHWISCOTICS
Earth		j.	1.12	TORRETTE JACONET	63/30/2011 11:30:30 PM	tonis s vicce sr	T11.3	DECEPTVE PRACTICE	PERFECTA DENTITY THEFT CARE \$ 200
B Nym Lang D Daw		1	+ II	10096100 34207054	63/50/2017 11:50:30 PM	DOTED STREAM ALC:	0613	34677	0168358
		(and bit	10 10	10095900 (4295754	600002117 11 29 36 PM	00000 N LOTUS AC	8559	ADSAIR.T	OWPLE
igina suprime Arota Naparile				40685940 34296755	6559(2)(T H 2)(3) PK	MORE IN CHICKGO AVE	6629	11217	\$518 AND UNCER
		implants	42 10	40686832 34206763	6559201111538PB	65500 5 CARPENTER ST	(343	CRIMINAL GAMAGE	10 PROPERTY
		Michael Cra-	13 10	10080810 14208244	65100011111038PB	OHEID N KEHLAND AR	6410	BATTERS .	INPLE
Comp Comp	and the second second	Analytican B	14.50	10000000 14200702	ESSOCIETT IT OR 35 PM	COD SKELLERAS	2024	NINCOTICS	PDD:HERDRONHTED
Personal Presents	utility and the second		10.00	*0000000 JA205705	63/59/2017 11:07:30 PM	ETSILS VENTYORTHAN	1414	1/54P019 HOL4TON	UNLAWFUL POSS OF HANDRUN
Pandara Orand Park		0	10.00	10095954 (4296750	655962117 Ht 0238 PM	05010 INFEROMAND ST	1310	CRIMINAL DAMAGE	TO PROPERTY
	s haven Church Church	disease . 83	7000	6003mi	9				

Crime data from Chicago is publicly available at https://data.cityofchicago.org/

Data includes locations and details of every crime recorded in Chicago since 2001.

We can use Voronoi residual analysis to assess the fit of a spatio-temporal point process (Bray 2014).

For an event (x_i, y_i) , the associated Voronoi cell C_i can be defined as

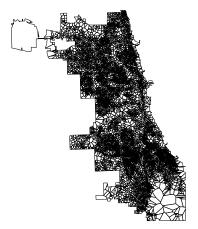
$$C_i = \{(x, y) \mid d((x, y), (x_i, y_i)) \leq d((x, y), (x_j, y_j)) \quad \forall i \neq j\},\$$

We can evaluate Voronoi residuals R_{vor}^{i} for each Voronoi cell *i*, where

$$R_{\text{vor}}^{i} = 1 - \int_{0}^{T} \int \int_{(x,y) \in C_{i}} \lambda(x,y,t) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t.$$

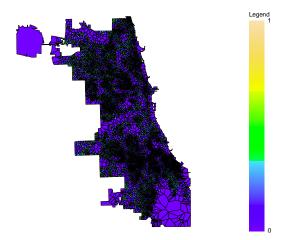
13 / 24

Voronoi Residual Analysis

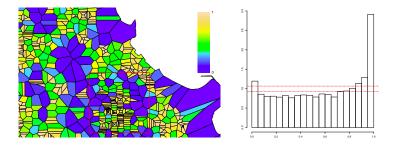


- 一司

Voronoi Residual Analysis



Voronoi Residual Analysis

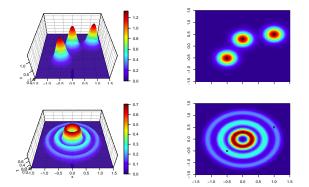


Modelling crime with a homogeneous background rate is not suitable.

Marsan and Lengliné (2008) introduced Model Independent Stochastic Declustering (MISD) in order to find a non-parametric form for the triggering function without any prior assumptions.

This has been used to model crime events (e.g. Mohler et al. 2011). We use different adaptations of this to model crime in Chicago.

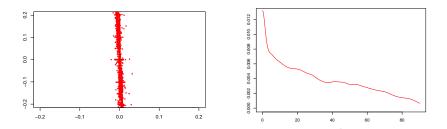
We model the crime assuming isotropy in the triggering component.



Results on Chicago Data

Results using model from Mohler et al. 2011.

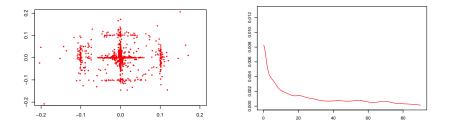
$$\lambda(x, y, t|\mathcal{H}_t) = \nu(t)\mu(x, y) + \sum_{t>t_i} g(x - x_i, y - y_i, t - t_i).$$



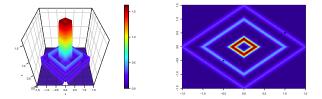
Results on Chicago Data

Results when assuming isotropy.

$$\lambda(x, y, t|\mathcal{H}_t) = \nu(t)\mu(x, y) + \sum_{t>t_i} g(r - r_i, t - t_i).$$



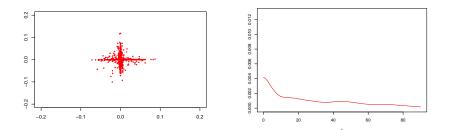
We also model the triggering function using Manhattan distance to measure the distance between events.



Results on Chicago Data

Results when using Manhattan distance.

$$\lambda(x,y,t|\mathcal{H}_t) =
u(t)\mu(x,y) + \sum_{t>t_i} g(d-d_i,t-t_i).$$



- Tested self-exciting point process models on Chicago crime data
- Tested variants of the standard approaches
- Could maintain predictive power with reduced dimensionality

23 / 24

The End

< □ > < ---->

æ