

Estatística Espacial (MI418) / Geoestatística (ME907)

Guilherme Ludwig

2019-02-07

Processos de Gibbs

Simulando processos de Gibbs

Inferência por pseudo-verossimilhança

Processos finitos com densidades

Considere um processo pontual X em $W \subset \mathbb{R}^d$ com densidade f com respeito ao processo de Poisson de taxa 1 em W . Isto é, para uma configuração finita F , o evento $X \in F$ tem probabilidade

$$\mathbb{P}(X \in F) = \sum_{n=0}^{\infty} \frac{e^{-|W|}}{n!} \int_W \cdots \int_W \mathbf{1}(\{x_1, \dots, x_n\}) f(\{x_1, \dots, x_n\}) dx_1 \cdots dx_n.$$

Por exemplo, o processo de Poisson de intensidade ρ tem densidade (com respeito ao processo de Poisson com intensidade 1) dada por

$$f(x) = \exp\{|W| - \mu(W)\} \prod_{\xi \in x} \rho(\xi).$$

Processos finitos com densidades condicionais

Note que f é uma densidade, logo $f = c^{-1}h \propto h$ onde c é uma constante de normalização

$$c = \sum_{n=0}^{\infty} \frac{e^{-|W|}}{n!} \int_W \cdots \int_W h(\{x_1, \dots, x_n\}) dx_1 \cdots dx_n.$$

Note também a densidade condicional $f_n \propto f$ que é usada em inferência corresponde a

$$\mathbb{P}(X \in F | N(X) = n) = \int_W \cdots \int_W \mathbf{1}(\{x_1, \dots, x_n\}) f_n(\{x_1, \dots, x_n\}) dx_1 \cdots dx_n.$$

A constante c_n e a função h_n são definidas igualmente.

Pairwise Interaction Gibbs Processes

Seja X um processo pontual localmente finito em $\mathcal{D} \subset \mathbb{R}^2$, com densidade

$$f_{\theta}(x) = c(\theta_1, \theta_2) \prod_{\xi \in x} \rho_{\theta_1}(\xi) \prod_{\{\xi, \eta\} \subset x} \phi_{\theta_2}(\xi, \eta),$$

com respeito ao processo de Poisson padrão, e indexado por um vetor de parâmetros $\theta = (\theta_1^t, \theta_2^t)^t$, $\theta_1 \in \mathbb{R}^p$, $\theta_2 \in \mathbb{R}^q$. Então X é um processo de Gibbs com interações de segunda ordem.

Há processos de Gibbs com interações de ordem maior, mas são bastante complicados.

O processo de Strauss

O processo de Strauss é o exemplo mais comum de processo com repulsão. Ele é definido por ter densidade com respeito ao processo de Poisson dada por

$$f(x) \propto \beta^{n(x)} \gamma^{S_R(x)},$$

onde

$$S_R(x) = \sum_{\{\xi, \eta\} \subset x} \mathbf{1}\{\|\xi - \eta\| < R\},$$

e $\beta > 0$ e $0 \leq \gamma \leq 1$. Quando $S_R = 0$ e $\gamma = 0$, tome a indeterminação $\gamma^{S_R} := 1$. A constante de normalização c não tem forma fechada.

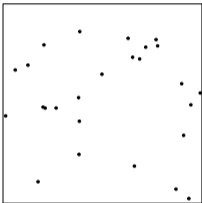
O processo de Strauss

- ▶ Note que $\gamma > 1$ resulta em um processo explosivo, com constante de normalização indefinida (i.e. integral diverge).
- ▶ $\gamma = 1$ resulta no processo de Poisson homogêneo com intensidade ρ .
- ▶ $\gamma = 0$ é o chamado processo *hardcore*. No processo hardcore, nenhuma configuração com pares de pontos a uma distância inferior a R é possível.

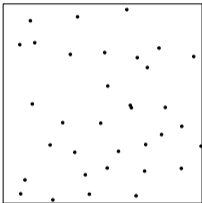
A densidade condicional do processo de Strauss é $f_n \propto \gamma^{S_R(x)}$, e existe mesmo para $\gamma > 1$. Portanto é possível usar o processo de Strauss para modelar atração ($\gamma > 1$), mas só condicionalmente (n fixo).

O processo de Strauss

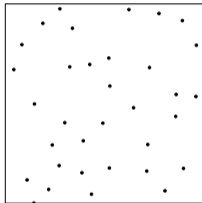
Processo de Poisson, $\beta = 30$



Processo de Strauss, $\beta = 70, 0.2, R = 0.1$

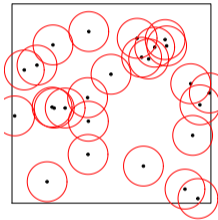


Processo Hardcore, $\beta = 100, R = 0.1$

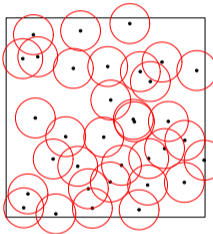


O processo de Strauss

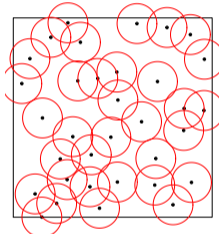
Processo de Poisson, $\beta = 30$



Processo de Strauss, $\beta = 70, 0.2, R = 0.1$



Processo Hardcore, $\beta = 100, R = 0.1$



Processos de nascimento e morte

Um processo de nascimento e morte espacial é um método dinâmico de geração de processos pontuais espaciais. Considere uma taxa de nascimento $b(x, \xi) > 0$ e uma taxa de morte $d(x, \xi) > 0$, com $\beta(x) = \int_W b(x, \xi) d\xi$, $\delta(\emptyset) = 0$, $\delta(x) = \sum_{\xi \in x} d(x, \xi)$ se $x \neq \emptyset$. Seja $\alpha(x) = \beta(x) + \delta(x)$. Dada uma configuração inicial¹ x_0 em $t_0 = 0$, para $m = 1, 2, \dots$

- ▶ $t_m = t_{m-1} + T$, onde $T \sim \text{Exp}(\alpha(x_m))$.
- ▶ Jogar uma moeda com probabilidade $\beta(x_m)/\alpha(x_m)$.
 - ▶ Se tiver sucesso, nasce 1 ponto ξ_m com densidade $b(x_m, \xi_m)/\beta(x_m)$. $X_m = x_{m-1} \cup \xi_m$.
 - ▶ Se tiver fracasso, e a configuração não estiver vazia, morre 1 ponto $\xi \in x$ com probabilidade $d(x_m, \xi)/\delta(x_m)$. Se a configuração estiver vazia, continua vazia.

As taxas b e d são chamadas de taxa de nascimento e taxa de morte, respectivamente.

¹Esqueci de dizer na aula anterior: com probabilidade não-nula.

Processos de nascimento e morte

A condição de balanço detalhado implica que o processo tem uma medida invariante. Suponha que h é a densidade não-normalizada de X com respeito ao processo de Poisson. Então se $\mathbb{E}(\beta(X)) < \infty$, então a equação de balanço detalhado é

$$b(x, \xi)h(x) = d(x, \xi)h(x \cup \xi).$$

Há mais de uma escolha de b , d com a mesma medida invariante. Mas em particular, se $d = 1$, então b é a intensidade de Papangelou.

Intensidade de Papangelou

Considere um processo X com densidade f_θ com respeito ao processo de Poisson. Então X tem intensidade condicional de Papangelou

$$\lambda_\theta^*(u, x) = \begin{cases} f_\theta(x \cup u)/f_\theta(x), & \text{se } u \notin x, \text{ e} \\ f_\theta(x)/f_\theta(x \setminus x_i), & \text{para } x_i \in x. \end{cases}$$

Você pode enxergar a intensidade de Papangelou como a probabilidade de um ponto ξ ocorrer, condicionalmente ao resto da configuração de pontos. Processos onde $\lambda^*(x, \xi) \leq \lambda^*(y, \xi)$ são atrativos quando $x \subset y$; e repulsivos se $\lambda^*(x, \xi) \geq \lambda^*(y, \xi)$ quando $x \subset y$.

Note: a intensidade condicional de Papangelou de um processo de Poisson é $\lambda_\theta(u, x) = f_\theta(u)$.

Intensidade de Papangelou

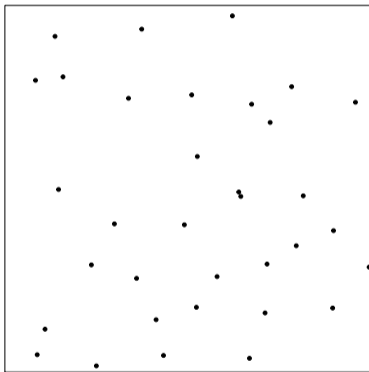
A intensidade de Papangelou do processo de Strauss é

$$\lambda^*(u, x) = \beta\gamma \sum_{i=1}^{n(x)} \mathbf{1}_{\{\|x_i - u\| < R\}}$$

e não depende de constante de normalização. Logo, é fácil simular processos de Strauss a partir de processos de nascimento e morte espaciais.

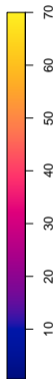
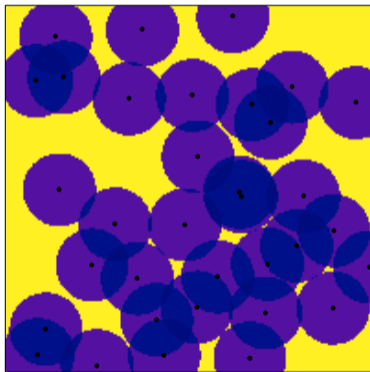
Intensidade de Papangelou

Processo de Strauss, $\beta = 70, 0.2, R = 0.1$



Intensidade de Papangelou

Processo de Strauss, $\beta = 70, 0.2, R = 0.1$



Intensidade de Papangelou

Segundo Møller and Waagepetersen (2003), um processo X com intensidade de Papangelou $\lambda^*(u, x)$ tem intensidade

$$\rho(\xi) = \mathbb{E}[\lambda^*(\xi, X)],$$

e densidade do produto de segunda ordem

$$\rho^{(2)}(\xi, \eta) = \mathbb{E}[\lambda^*(\xi, X)\lambda^*(\eta, X \cup \xi)].$$

Alguns modelos comuns de processos de Gibbs

- ▶ **Processos hardcore** tem densidade $c\beta^{n(x)}\mathbf{1}\{\text{nenhum par de pontos está a menos de } R \text{ de distância}\}$. Disponível em `spatstat::rHardcore`.
- ▶ **Processo de Strauss** tem densidade $c\beta^{n(x)}\gamma^{S_R(x)}$, onde $S_R(x) = \sum_{\{\xi, \eta\} \subset X} \mathbf{1}\{\|\xi - \eta\| < R\}$. Disponível em `spatstat::rStrauss`, neste caso em particular com simulação perfeita.
- ▶ **Processo de Strauss não-homogêneo** tem densidade $c\beta(x_1) \cdots \beta(x_n)\gamma^{S_R(x)}$. É usado em regressão, por exemplo. Disponível em `spatstat::rStrauss` com uma fórmula pra média.
- ▶ **Processo softcore** tem densidade $c\beta^{n(x)} \exp\{-\sum_{\{\xi, \eta\} \subset X} (\sigma/\|\xi - \eta\|)^{2/\kappa}\}$. Disponível em `spatstat::rSoftcore`.

Há outros processos com aplicações específicas, mas estes acima são suficientes para uma visão geral. O `spatstat::Pairwise` permite que você customize sua função de interação.

Inferência

O processo de Strauss têm uma vantagem sobre outros processos de Gibbs: ele está intrínsecamente conectado à função K empírica. Em particular,

$$s_R(x) = \frac{n(n-1)}{2|W|} \hat{K}(R).$$

De fato, há uma aproximação (Penttinen, 1984, apud Baddeley et al., 2015) que aproxima o estimador de máxima verossimilhança

$$\hat{\gamma} \approx \frac{\hat{K}(R)}{\pi R^2}.$$

Note que a inferência é feita baseada em R fixo.

Na prática, máxima verossimilhança só funciona se você conseguir aproximar c_θ . Métodos Bayesianos são comumente usados, pois só precisam de algo proporcional à verossimilhança, e devido à conexão natural com o mecanismo de simulação do processo de Gibbs. Uma alternativa popular é a chamada pseudo-verossimilhança.

Pseudo-verossimilhança

Dentro da caixa $W \subset \mathcal{D}$, a pseudo-verossimilhança de X é

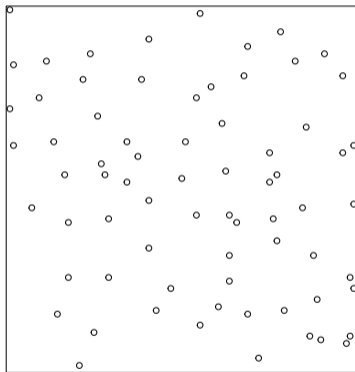
$$PL(\boldsymbol{\theta}; \mathbf{x}) = \left(\prod_{x_i \in W} \lambda_{\boldsymbol{\theta}}(x_i, \mathbf{x}) \right) \exp \left\{ - \int_W \lambda_{\boldsymbol{\theta}}(\xi, \mathbf{x}) d\xi \right\}.$$

No caso de regressão, se $\log(\lambda_{\boldsymbol{\theta}}(\xi; \mathbf{x})) \approx \boldsymbol{\beta}' \mathbf{z}(\xi; \mathbf{x})$, então $\hat{\boldsymbol{\beta}}$ é solução de

$$\sum_{x_i \in W} \mathbf{z}(x_i, \mathbf{x}) = \int_W \mathbf{z}(\xi, \mathbf{x}) \exp\{\boldsymbol{\beta}' \mathbf{z}(\xi; \mathbf{x})\} du.$$

Exemplo: Swedish Pines

Swedish Pines



Os dados estão em `spatstat::swedishpines`.

Exemplo: Swedish Pines

```
library(spatstat)
# Default for ppm with interaction is pseudolikelihood
ppm(swedishpines ~ 1, Strauss(r = 5))
```

```
## Stationary Strauss process
##
## First order term:  beta = 0.0123665
##
## Interaction distance:      5
## Fitted interaction parameter gamma:  0.2725979
##
## Relevant coefficients:
## Interaction
##   -1.299758
##
## For standard errors, type coef(summary(x))
```

Exemplo: Swedish Pines

```
coef(summary(ppm(swedishpines ~ 1, Strauss(r = 5))))
```

##		Estimate	S.E.	CI95.lo	CI95.hi	Ztest	Zval
##	(Intercept)	-4.392764	0.2132260	-4.810679	-3.9748485	***	-20.601444
##	Interaction	-1.299758	0.5075298	-2.294498	-0.3050174	*	-2.560948

Exemplo: Regressão

```
library(spatstat)
ppm(swedishpines ~ x * y, Strauss(r = 5))
```

```
## Nonstationary Strauss process
```

```
##
```

```
## Log trend: ~x * y
```

```
##
```

```
## Fitted trend coefficients:
```

```
##      (Intercept)           x           y           x:y
## -5.2646287402  0.0186195126  0.0143272135 -0.0003053618
```

```
##
```

```
## Interaction distance:      5
```

```
## Fitted interaction parameter gamma:  0.2605748
```

```
##
```

```
## Relevant coefficients:
```

```
## Interaction
```

```
##      1.244265
```

Parâmetros irregulares

As estimativas são sensíveis ao parâmetro R

```
coef(summary(ppm(swedishpines ~ 1, Strauss(r = 4))))
```

##		Estimate	S.E.	CI95.lo	CI95.hi	Ztest	Zval
##	(Intercept)	-4.7483122	0.1674391	-5.076487	-4.4201376	***	-28.358450
##	Interaction	-0.9207604	0.5229416	-1.945707	0.1041864		-1.760733

```
coef(summary(ppm(swedishpines ~ 1, Strauss(r = 5))))
```

##		Estimate	S.E.	CI95.lo	CI95.hi	Ztest	Zval
##	(Intercept)	-4.392764	0.2132260	-4.810679	-3.9748485	***	-20.601444
##	Interaction	-1.299758	0.5075298	-2.294498	-0.3050174	*	-2.560948

Parâmetros irregulares

De fato, R é chamado de parâmetro irregular. Para interações *hardcore*, é fácil achar uma estimativa de R (basta escolher $R = \min\{x_i - x_j\}$), ainda que mostrar que essa estimativa é não-viesada ou ótima seja bastante difícil.

Para repulsão probabilística ou *softcore*, é bem difícil. Uma possibilidade é considerar a verossimilhança perfilada, em que escolhemos R que maximiza $PL_R(\hat{\theta}; x)$.

Parametros irregulares

```
sqrt(area(swedishpines)) # "Maximum"
```

```
## [1] 97.97959
```

```
min(dist(cbind(swedishpines$x, swedishpines$y))) # "min"
```

```
## [1] 2.236068
```

```
R <- data.frame(r = seq(2.5, 20, by = 0.1))
```

```
ps <- profilepl(R, Strauss, swedishpines)
```

```
## (computing rbord)
```

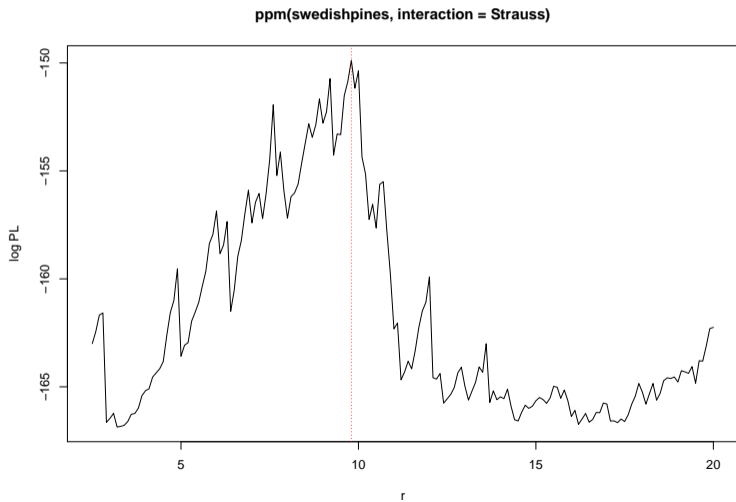
```
## comparing 176 models...
```

```
## 1, 2, 3, 4.6.8.10.12.14.16.18.20.22.24.26.28.30.32.34.36.38.
```

```
## 40.42.44.46.48.50.52.54.56.58.60.62.64.66.68.70.72.74.76.78
```

Parametros irregulares

```
plot(ps, col.opt = "Red")
```



Fim das aulas, obrigado!

Referências I

- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. CRC Press.
- Møller, J. and Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall, Baton Rouge.