Estimation semi-paramétrique de l'intensité de Poisson d'un processus ponctuel à interaction par paires homogène

Nadia Morsli et Jean-François Coeurjolly.
Laboratoire Jean Kuntzmann, Université de Grenoble, France.
nadia.morsli@imag.fr; jean-francois.coeurjolly@upmf-grenoble.fr

Intensité conditionnelle de papangelou

Une densité f est dite héréditaire si pour toute configuration finie de points $\mathbf{x}$

$$f(\mathbf{x}) > 0 \text{ et pour tout y } \subseteq x \Longrightarrow f(y) > 0.$$

Cette condition, signifie que si une configuration peut se produire, alors toutes les sous-configurations qu'elle contient peuvent se produire aussi.
Si f est héréditaire, l'intensité de Papangelou de $u \notin \mathbf{x}$ conditionnelle à $\mathbf{x}$ est définie par

$$\lambda_\beta(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})}.$$

λ présente l'avantage sur f d'éliminer la constante de normalisation, c'est pourquoi son usage sera central. $\lambda_\beta(u, \mathbf{x})$ peut être interprétée comme étant la densité de probabilité qu'il y ait un point de $\mathbf{X}$ en u, sachant que la réalisation de $\mathbf{X}$ est $\mathbf{x}$ ailleurs.

Processus ponctuel à interaction par paires homogène

- Un processus ponctuel est "à interaction par paires" lorsque son intensité conditionnelle de Papangelou en $\mathbf{x} \in \Omega$, Ω étant l'espace des configurations des points localement finies, est de la forme:

$$\lambda(u, \mathbf{x}) = \phi(u) \prod_{v \in \mathbf{x} \setminus u} \phi(\{u, v\}), \quad u \in \mathbb{R}^d$$

  où ϕ est une fonction d'interaction.

- Un processus ponctuel à interaction par paires est dit homogène si
  - $\phi(u)$ est constant
    et
  - $\phi(\{u, v\}) = \phi_2(||u - v||)$
    où
    - $\phi_2 : (0, \infty) \longrightarrow [0, \infty)$
    - $||.||$ symbolise la distance euclidienne.

Exemples des processus à interaction par paires

- **Processus ponctuel de Strauss:**

$$\lambda(u, \mathbf{x}) = \beta \gamma^{n_{[0,R]}(u,\mathbf{x})}$$

où $\beta > 0, \gamma \in [0, 1]$ et $n_{[0,R]}(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \mathbb{1}(\|v - u\| \leq R)$ représente le nombre de R-voisins fermés de u dans $\mathbf{x}$.

- **Processus ponctuel de Strauss avec hard-core:**

$$\lambda(u, \mathbf{x}) = \begin{cases} 0 & \text{si } ||u - v|| \leq \delta \\ \beta \gamma^{n_{]\delta,R]}(u,\mathbf{x}\setminus u)} & \text{ailleurs} \end{cases}$$

où $\beta > 0, \gamma \in [0, 1]$ et $n_{]\delta,R]}(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \mathbb{1}(\|v - u\| \in]\delta, R]$.

- **Processus ponctuel de Piecewise Strauss:**

$$\lambda(u, \mathbf{x}) = \beta \prod_{j=1}^{p} \gamma_j^{n_{[R_{j-1},R_j]}(u,\mathbf{x}\setminus u)}$$

où
$\beta > 0, \gamma_j \in [0, 1], n_{[R_{j-1},R_j]}(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \mathbb{1}(\|v - u\| \in [R_{j-1}, R_j])$
et $0 = R_0 < R_1 < \ldots < R_p = R < \infty$.

Position du problème

On considère un modèle à interaction par paires homogène caractérisé par sa densité conditionnelle de Papangelou

$$\lambda(u, \mathbf{x}) = \beta \cdot \exp(-\sum_{v \in \mathbf{x} \setminus u} g(||v-u||)), \ \forall u \in \mathbb{R}^d, \mathbf{x} \in \Omega \quad (1)$$

où

- $\beta > 0$ est le paramètre d'intensité de Poisson.
- g est le potentiel d'interaction de paires.
- **[BG]:** g admet une portée d'interaction finie R, i.e. $g(||v-u||) = 0 \quad \text{pour } ||v-u|| > R, \ \forall u \in \mathbb{R}^d, v \in \mathbf{x} \in \Omega.$

Objectif:

- Estimation du paramètre β indépendamment de la fonction g basèe sur une seule observation du processus à interaction par paires homogène dans $\mathbb{R}^d$, noté $\mathbf{X}$, dans un domaine $\Lambda_{n_{\ominus \widetilde{R}}}$.
  - $\Lambda_{n_{\ominus \widetilde{R}}} = \{u \in \Lambda_n : B(u, \widetilde{R}) \subset \Lambda_n\}$,
  - $\lim_{n \to \infty} |\Lambda_{n_{\ominus \widetilde{R}}}| = \infty$.
  - $\widetilde{R}$ est la borne supérieure de R supposée connue.

Définition de l'estimateur du paramètre d'intensité de Poisson

- Formule de Georgii-Nguyen-Zessin: $\mathbf{X}$ est supposé un processus ponctuel de Gibbs d'intensité conditionnelle de Papangelou λ , tel que,

$$\mathbf{E}\left[\sum_{u \in \mathbf{X}} h(u, \mathbf{X} \backslash u)\right] = \mathbf{E}\left[\int_{\mathbb{R}^d} h(u, \mathbf{X}) \lambda(u, \mathbf{X}) du\right] \qquad (2)$$

pour toute fonction mesurable non négative h de $\mathbb{R}^d \times \Omega$.

- Pour le choix de la fonction mesurable non négative h définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ par

$$\textbf{[CH]: } h(u, \mathbf{x}) = \mathbb{1}\left(\inf_{v \in \mathbf{x}} ||v - u|| > \widetilde{R}\right) = \mathbb{1}\left(d(u, \mathbf{x}) > \widetilde{R}\right)$$

il découle que:

$$
\begin{aligned}
\mathbf{E} \sum_{u \in \mathbf{X}_{\Lambda_{n_{\ominus \widetilde{R}}}}} h(u, \mathbf{X} \backslash u) &= \mathbf{E} \overbrace{\sum_{u \in \mathbf{X}_{\Lambda_{n_{\ominus \widetilde{R}}}}} \mathbb{1}\left(\inf_{v \in \mathbf{X}} ||v - u|| > \widetilde{R} \right)}^{N_{\Lambda_{n_{\ominus \widetilde{R}}}}} \\
&= \mathbf{E} \int_{\Lambda_{n_{\ominus \widetilde{R}}}} \mathbb{1}\left(\inf_{v \in \mathbf{X}} ||v - u|| > \widetilde{R} \right) \lambda(u, \mathbf{X}) du \\
&= \beta \mathbf{E} \int_{\{u; \inf_{v \in \mathbf{X}} ||v-u|| > \widetilde{R}\}} \exp(- \sum_{v \in \mathbf{X} \backslash u} g(||v - u||)) du \\
&= \beta \mathbf{E} \int_{\Lambda_{n_{\ominus \widetilde{R}}}} \mathbb{1}(\inf_{v \in \mathbf{X}} ||v - u|| > \widetilde{R}) \exp(0) du \\
&= \beta \mathbf{E} \overbrace{\int_{\Lambda_{n_{\ominus \widetilde{R}}}} \mathbb{1}\left(\inf_{v \in \mathbf{X}} ||v - u|| > \widetilde{R} \right) du}^{V_{\Lambda_{n_{\ominus \widetilde{R}}}}} . \qquad (3)
\end{aligned}
$$

Consistance forte de l'estimateur de l'interaction du premier ordre

Il résulte de l'equation (3) qu'un estimateur naturel du paramètre β, noté $\hat{\beta}_n$ est défini par

$$\widehat{\beta}_n(\mathbf{X};\widetilde{R}) = \frac{|\Lambda_{n_{\ominus \widetilde{R}}}|^{-1} N_{\Lambda_{n_{\ominus \widetilde{R}}}}}{|\Lambda_{n_{\ominus \widetilde{R}}}|^{-1} V_{\Lambda_{n_{\ominus \widetilde{R}}}}}.$$

Proposition

Soit $\mathbf{X}$ *un processus ponctuel à interaction par paires homogène, de potentiel d'interaction de paires* g *défini par* (1) *satisfait la condition* **[BG]** *et sous la condition* **[CH]**. *Alors pour tout* $\widetilde{R}$ *fixé* ($\widetilde{R} \geq R$), *l'estimateur* $\widehat{\beta}_n$ *du paramètre* β *est fortement consistant.*

La normalité asymptotique de l'estimateur du paramètre d'intensité de Poisson

Proposition

Soit $\mathbf{X}$ *un processus ponctuel à interaction par paires homogène ergodique, de potentiel d'interaction de paires g défini par* (1) *satisfait la condition* **[BG]** *et sous la condition* **[CH]***. Alors pour tout* $\widetilde{R}$ *fixé* ($\widetilde{R} \geq R$),

$$|\Lambda_{n_{\ominus \widetilde{R}}}|^{1/2}(\widehat{\beta}_n - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \tag{4}$$

où

- $\sigma^2 = \sigma^2(\beta, \widetilde{R}) = \frac{\beta}{1-F(\widetilde{R})} + \frac{\beta^2 \Omega_1(\widetilde{R})}{(1-F(\widetilde{R}))^2}$.
- $\Omega_1(\widetilde{R}) = \int_{B(0,\widetilde{R})} P_\beta(d(o, \mathbf{X}) > \widetilde{R}, d(v, \mathbf{X}) > \widetilde{R}) dv$.
- $F(\widetilde{R}) = P_\beta\left(d(0, \mathbf{X}) \leq \widetilde{R}\right)$.

$$|\Lambda_{n_{\ominus \widetilde{R}}}|^{1/2} \frac{(\widehat{\beta}_n - \beta)}{\widehat{\sigma}_n} \xrightarrow{d} \mathcal{N}(0,1) \tag{5}$$

$\widehat{\sigma}_n^2$ est un estimateur fortement consistant de σ^2 défini par

- $\widehat{\sigma}_n^2 = |\Lambda_{n_{\ominus \widetilde{R}}}| \left(\frac{\widehat{\beta}_n}{V_{\Lambda_{n_{\ominus \widetilde{R}}}}} + \frac{\widehat{\beta}_n^2 W_{\Lambda_{n_{\ominus \widetilde{R}}}}}{V^2_{\Lambda_{n_{\ominus \widetilde{R}}}}} \right)$
  avec
- $$W_{\Lambda_{n_{\ominus \widetilde{R}}}} = \int_{\Lambda_{n_{\ominus \widetilde{R}}}} \int_{B(u,\widetilde{R}) \cap \Lambda_{n_{\ominus \widetilde{R}}}} \mathbb{1}\left(d(u,\mathbf{x}) > \widetilde{R}\right) \mathbb{1}\left(d(v,\mathbf{x}) > \widetilde{R}\right) du dv$$

Etude par Simulations

Processus ponctuel de Strauss: $\lambda(u, \mathbf{x}) = \beta\gamma^{n_{[0,R]}(u,\mathbf{x})}$

- `s1`: $\gamma = 0.2$.
- `s2`: $\gamma = 0.8$.

Processus ponctuel de Strauss avec hard-core:

$$\lambda(u, \mathbf{x}) = \beta\gamma^{n_{]\delta,R]}(u,\mathbf{x}\backslash u)}$$

- `shc1`: $\gamma = 0.2$, $\delta = R/2$.
- `shc2`: $\gamma = 0,8$, $\delta = R/2$.

Processus ponctuel de Piecewise Strauss:

$$\lambda(u, \mathbf{x}) = \beta \prod_{j=1}^{p} \gamma_j^{n_{[R_{j-1},R_j]}(u,\mathbf{x}\backslash u)}$$

- $(R_1, R_2, R_3) = (R/3, 2R/3, R)$.
- `ps1`: $(\gamma_1, \gamma_2, \gamma_3) = (0.8, 0.5, 0.2)$.
- `ps2`: $(\gamma_1, \gamma_2, \gamma_3) = (0.2, 0.8, 0.2)$.

Modèle	$\overline{n}$	Moyenne (Ecart type) pour $\widetilde{R} = R \times p$, $R = 0,05$ $p = 0.9$	$p = 1$	$p = 1.1$	$p = 1.2$
s1 $L = 1$	99	174.0 (26.2)	203.6 (34.2)	204.3 (38.3)	205.5 (44.4)
$L = 2$	393	172.6 (12.7)	200.5 (16.4)	200.3 (18.5)	200.6 (20.8)
s2 $L = 1$	156	192.0 (31.3)	203.3 (41.8)	203.9 (48.1)	204.7 (55.8)
$L = 2$	622	190.7 (16.0)	200.7 (19.5)	200.3 (22.8)	200.4 (26.5)
shc1 $L = 1$	94	174.3 (23.2)	201.9 (30.3)	203.3 (36.8)	205.0 (41.9)
$L = 2$	379	174.1 (12.1)	201.6 (15.7)	201.9 (17.8)	202.1 (21.1)
shc2 $L = 1$	130	194.3 (30.2)	205.4 (36.7)	207.0 (42.8)	209.1 (51.6)
$L = 2$	514	190.2 (15.2)	198.5 (17.8)	199.5 (21.1)	200.6 (24.4)
ps1 $L = 1$	111	172.2 (25.1)	202.6 (33.2)	203.5 (38.7)	206.5 (47.2)
$L = 2$	445	171.5 (12.4)	201.6 (16.3)	201.7 (18.6)	202.4 (21.6)
ps2 $L = 1$	134	191.1 (30.7)	201.7 (37.7)	206.2 (45.2)	208.8 (54.2)
$L = 2$	535	192.8 (14.5)	201.7 (17.5)	202.1 (20.6)	202.0 (24.4)

Table: Moyennes et écarts-types des estimations du paramètre d'intensité de Poisson $\beta = 200$ pour différents paramètres $\widetilde{R}$ basés sur 500 réplications des modèles $s1, s2, sch1, sch2, ps1, ps1$ générés sur la fenêtre $[0, L]^2$ et estimés sur la fenêtre $[\widetilde{R}, L - \widetilde{R}]^2$ pour $L = 1, 2$.

		$\widetilde{R} = p \times R$			
		$p = 0.9$	$p = 1$	$p = 1.1$	$p = 1.2$
s1	$L = 1$	77.4%	94.8%	95.0%	94.2%
	$L = 2$	40.0%	94.8%	93.8%	95.4%
s2	$L = 1$	82.6%	93.2%	94.6%	93.8%
	$L = 2$	69.8%	95.0%	93.6%	92.8%

Table: Les taux de couverture empiriques (i.e. la fraction de 95% des intervalles de confiance de couverture du paramètre $\beta = 200$) basés sur 500 réplications des modèles de Strauss sur la fenêtre $[0, L]^2$.

Les objectifs

Nous illustrons l'intérêt et l'efficacité de ce travail pour un autre développement:

- L'estimation non paramétrique de la fonction g pour notre modèle d'interaction par paires homogène caractérisé par sa densité conditionnelle de Papangelou:

$$\lambda(u,\mathbf{x}) = \beta\cdot \exp(-\sum_{v\in\mathbf{x}\setminus u} g(||v-u||)),\ \forall u \in \mathbb{R}^d, \mathbf{x}\in\Omega.$$

J.-F. Coeurjolly and F. Lavancier.
Residuals and goodness-of-fit test for stationary marked Gibbs point processes.
Journal of the Royal Statistical Society, Series B, 75(2):247–276, 2013.

J.-F. Coeurjolly, N. Morsli.
Poisson intensity parameter estimation for stationary Gibbs point processes.
To appear in Spatial Statistics, 2013.

J.-F. Coeurjolly and E. Rubak.
Fast covariance estimation for innovations computed from a spatial Gibbs point process.
To appear in Scandinavian Journal of Statistics, 2013.

X.X. Nguyen and H. Zessin.
Ergodic theorems for Spatial Process.
Z. Wahrscheinlichkeitstheorie verw. Gebiete, 48:133–158, 1979.