

Approximation du maximum de vraisemblance des processus déterminantaux

Arnaud Poinas

Travail en collaboration avec Frédéric Lavancier (LMJL, Nantes)

22 Juin 2022



- 1 Processus ponctuels déterminantaux
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

Définition (Processus ponctuel sur $\mathbb{R}^d$)

Un processus ponctuel (ou PP) sur $\mathbb{R}^d$ est un sous-ensemble aléatoire X de $\mathbb{R}^d$ localement fini. C'est à dire que pour tout $A \subset \mathbb{R}^d$ borné,

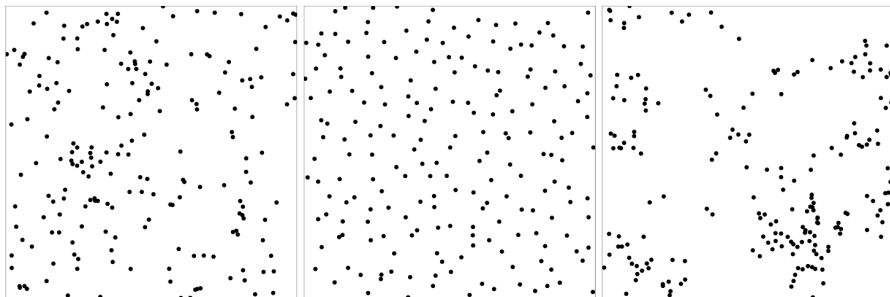
$$\text{card}(X \cap A) < +\infty \text{ p.s.}$$

Définition (Processus ponctuel sur $\mathbb{R}^d$)

Un processus ponctuel (ou PP) sur $\mathbb{R}^d$ est un sous-ensemble aléatoire X de $\mathbb{R}^d$ localement fini. C'est à dire que pour tout $A \subset \mathbb{R}^d$ borné,

$$\text{card}(X \cap A) < +\infty \text{ p.s.}$$

On distingue trois principaux types de comportement pour les PPs.



Exemples de jeux de données ponctuelles répulsives

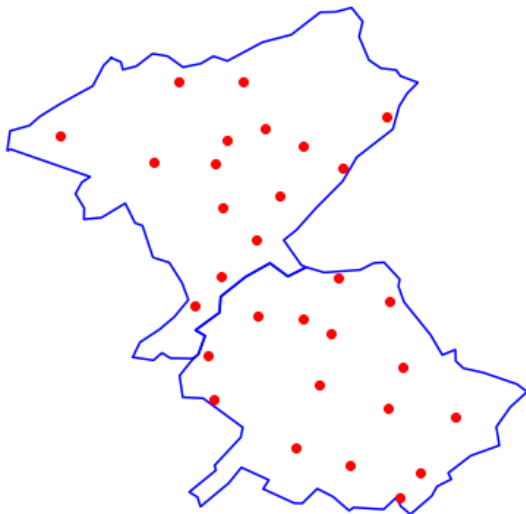


Figure: *Emplacement des antennes SFR à Roubaix et Tourcoing*

Exemples de jeux de données ponctuelles répulsives

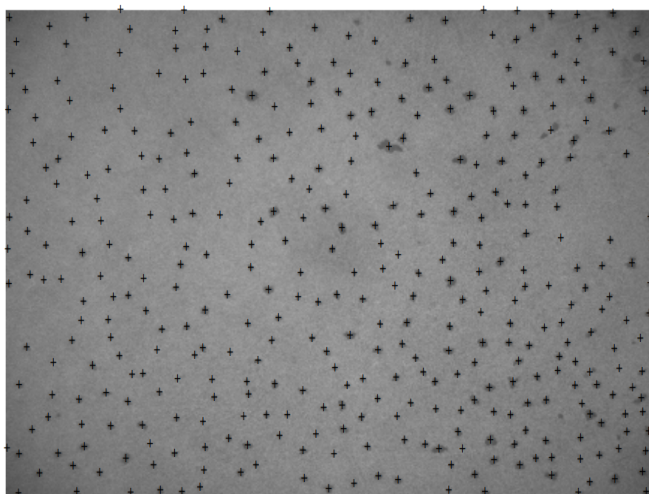


Figure: *Emplacement de glandes sudoripares.*

Exemples de jeux de données ponctuelles répulsives

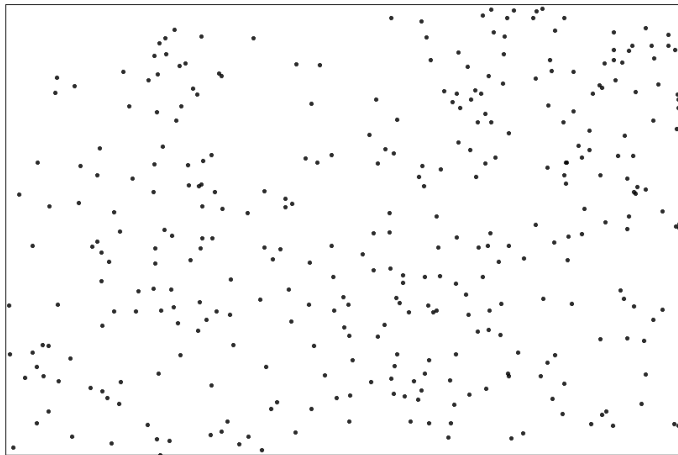


Figure: *Emplacement de 340 arbres dans une région de 60m par 90m du Parc Naturel d'Urkiola.*

Exemples de jeux de données ponctuelles répulsives

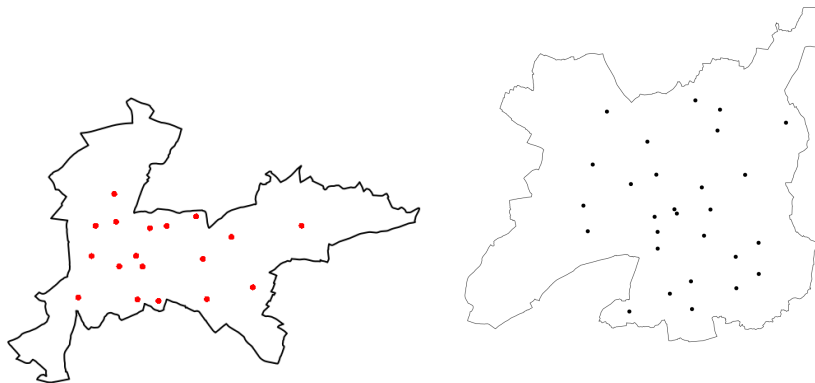


Figure: *Emplacement des bureaux de vote à Poitiers et Rennes*

Exemples de modèles de processus ponctuels répulsifs

- Processus ponctuels hardcore de Matérn
- Grilles perturbées
- Processus de Gibbs (ex: Strauss)
- Processus déterminantaux

Outline

- 1 **Processus ponctuels determinantaux**
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 **Approximation asymptotique de la vraisemblance des DPPs**
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 **Conclusion et questions ouvertes**

Définition (Fonctions d'intensité jointe d'un PP)

Soit $n \in \mathbb{N}$ non nul, la fonction d'intensité jointe d'ordre n d'un PP X est la fonction $\rho^{(n)}$ qui, si elle existe, vérifie

$$\mathbb{E} \left[\sum_{(x_1, \dots, x_n) \in X}^{\neq} h(x_1, \dots, x_n) \right] = \int_{(\mathbb{R}^d)^n} h(x_1, \dots, x_n) \rho^{(n)}(x_1, \dots, x_n) dx$$

pour n'importe quelle fonction $h : \mathbb{R}^{dn} \rightarrow \mathbb{R}$ intégrable.

La quantité

$$\rho^{(n)}(x_1, \dots, x_n)$$

peut être interprétée comme la **densité de probabilité** que X possède un point en chaque x_i .

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx$$

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx = \rho |A| \quad \text{si } \rho^{(1)}(x) = \rho.$$

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx = \rho |A| \quad \text{si } \rho^{(1)}(x) = \rho.$$

- $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y)$ peut-être interprété comme une mesure de dépendance entre deux points.

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx = \rho|A| \quad \text{si } \rho^{(1)}(x) = \rho.$$

- $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y)$ peut-être interprété comme une mesure de dépendance entre deux points.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) = 0$ pour les processus de Poisson.

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx = \rho|A| \quad \text{si } \rho^{(1)}(x) = \rho.$$

- $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y)$ peut-être interprété comme une mesure de dépendance entre deux points.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) = 0$ pour les processus de Poisson.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) \leq 0$ pour les processus répulsifs.

Deux quantités intéressantes:

- $\rho^{(1)}(x)$ est l'intensité d'un processus ponctuel.

$$\mathbb{E}[\text{card}(X \cap A)] = \int_A \rho^{(1)}(x) dx = \rho|A| \quad \text{si } \rho^{(1)}(x) = \rho.$$

- $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y)$ peut-être interprété comme une mesure de dépendance entre deux points.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) = 0$ pour les processus de Poisson.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) \leq 0$ pour les processus répulsifs.
 - ▶ $\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) \geq 0$ pour les processus attractifs.

Définition (Processus ponctuel déterminantal (O. Macchi, 1975))

Un processus ponctuel X est dit déterminantal (ou DPP) sur $\mathbb{R}^d$ de noyau K s'il existe une fonction $K : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$ telle que pour tout $n \in \mathbb{N}$:

$$\rho^{(n)}(x_1, \dots, x_n) = \det \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

Définition (Processus ponctuel déterminantal (O. Macchi, 1975))

Un processus ponctuel X est dit déterminantal (ou DPP) sur $\mathbb{R}^d$ de noyau K s'il existe une fonction $K : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$ telle que pour tout $n \in \mathbb{N}$:

$$\rho^{(n)}(x_1, \dots, x_n) = \det \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

En particulier:

- $\rho^{(1)}(x) = K(x, x)$

Définition (Processus ponctuel déterminantal (O. Macchi, 1975))

Un processus ponctuel X est dit déterminantal (ou DPP) sur $\mathbb{R}^d$ de noyau K s'il existe une fonction $K : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$ telle que pour tout $n \in \mathbb{N}$:

$$\rho^{(n)}(x_1, \dots, x_n) = \det \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

En particulier:

- $\rho^{(1)}(x) = K(x, x)$
- Si $K(x, y) = K(y, x)$ alors

$$\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) = \det \begin{pmatrix} K(x, x) & K(x, y) \\ K(y, x) & K(y, y) \end{pmatrix} - K(x, x)K(y, y)$$

Définition (Processus ponctuel déterminantal (O. Macchi, 1975))

Un processus ponctuel X est dit déterminantal (ou DPP) sur $\mathbb{R}^d$ de noyau K s'il existe une fonction $K : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$ telle que pour tout $n \in \mathbb{N}$:

$$\rho^{(n)}(x_1, \dots, x_n) = \det \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

En particulier:

- $\rho^{(1)}(x) = K(x, x)$
- Si $K(x, y) = K(y, x)$ alors

$$\rho^{(2)}(x, y) - \rho^{(1)}(x)\rho^{(1)}(y) = -K(x, y)^2 \leq 0$$

Hypothèses sur le noyau $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$

Hypothèses sur le noyau $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$

- K est symétrique: $\forall x, y \in \mathbb{R}^d, K(x, y) = K(y, x)$.
- K est continu et localement intégrable.

Hypothèses sur le noyau $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$

- K est symétrique: $\forall x, y \in \mathbb{R}^d, K(x, y) = K(y, x)$.
- K est continu et localement intégrable.

Pour tout compact $W \subset \mathbb{R}^d$ on définit l'opérateur intégral

$$\mathcal{K}|_W : f \mapsto \int_W K(., y)f(y)dy$$

- $\mathcal{K}|_W$ est à valeurs propres dans $[0, 1]$
- $\text{Tr}(\mathcal{K}|_W) < \infty$.

Construction de familles de DPPs stationnaires

Théorème (Lavancier, Møller, Rubak, 2015)

Soit $K_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de covariance continue de transformée de Fourier $0 \leq \hat{K}_0 \leq 1$, alors $K(x, y) := K_0(y - x)$ est le noyau d'un DPP stationnaire.

Construction de familles de DPPs stationnaires

Théorème (Lavancier, Møller, Rubak, 2015)

Soit $K_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de covariance continue de transformée de Fourier $0 \leq \hat{K}_0 \leq 1$, alors $K(x, y) := K_0(y - x)$ est le noyau d'un DPP stationnaire.

Noyau Gaussien: $K^{\rho, \alpha}(x, y) := \rho \exp\left(-\left\|\frac{y-x}{\alpha}\right\|^2\right)$ où $0 < \rho \leq (\sqrt{\pi}\alpha)^{-d}$

Construction de familles de DPPs stationnaires

Théorème (Lavancier, Møller, Rubak, 2015)

Soit $K_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de covariance continue de transformée de Fourier $0 \leq \hat{K}_0 \leq 1$, alors $K(x, y) := K_0(y - x)$ est le noyau d'un DPP stationnaire.

Noyau Gaussien: $K^{\rho, \alpha}(x, y) := \rho \exp\left(-\left\|\frac{y-x}{\alpha}\right\|^2\right)$ où $0 < \rho \leq (\sqrt{\pi}\alpha)^{-d}$

Noyau de Cauchy: $K^{\rho, \alpha, \nu}(x, y) := \frac{\rho}{\left(1 + \left\|\frac{y-x}{\alpha}\right\|^2\right)^{\nu+d/2}}$ où $0 < \rho \leq \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)(\sqrt{\pi}\alpha)^d}$

Construction de familles de DPPs stationnaires

Théorème (Lavancier, Møller, Rubak, 2015)

Soit $K_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de covariance continue de transformée de Fourier $0 \leq \hat{K}_0 \leq 1$, alors $K(x, y) := K_0(y - x)$ est le noyau d'un DPP stationnaire.

Noyau Gaussien: $K^{\rho, \alpha}(x, y) := \rho \exp\left(-\left\|\frac{y-x}{\alpha}\right\|^2\right)$ où $0 < \rho \leq (\sqrt{\pi}\alpha)^{-d}$

Noyau de Cauchy: $K^{\rho, \alpha, \nu}(x, y) := \frac{\rho}{\left(1 + \left\|\frac{y-x}{\alpha}\right\|^2\right)^{\nu+d/2}}$ où $0 < \rho \leq \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)(\sqrt{\pi}\alpha)^d}$

Noyau de Bessel: $K^{\rho, \alpha, \sigma}(x, y) := \rho 2^{(\sigma+d)/2} \Gamma\left(\frac{\sigma+d+2}{2}\right) \frac{J_{(\sigma+d)/2}(\sqrt{2(\sigma+d)}\left\|\frac{y-x}{\alpha}\right\|)}{(\sqrt{2(\sigma+d)}\left\|\frac{y-x}{\alpha}\right\|)^{(\sigma+d)/2}}$

où $0 < \rho \leq \frac{(\sigma+d)^{d/2} \Gamma((\sigma+2)/2)}{\Gamma((\sigma+d+2)/2) (\sqrt{2\pi}\alpha)^d}$

Construction de familles de DPPs stationnaires

Théorème (Lavancier, Møller, Rubak, 2015)

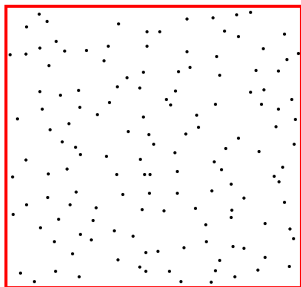
Soit $K_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de covariance continue de transformée de Fourier $0 \leq \hat{K}_0 \leq 1$, alors $K(x, y) := K_0(y - x)$ est le noyau d'un DPP stationnaire.

Noyau Gaussien: $\hat{K}_0^{\rho, \alpha}(x) = \rho(\sqrt{\pi}\alpha)^d \exp\left(-\|\pi\alpha x\|^2\right)$

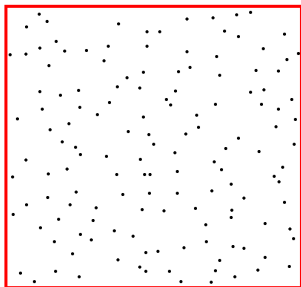
Noyau de Cauchy: $\hat{K}_0^{\rho, \alpha, \nu}(x) = \frac{\rho(\sqrt{\pi}\alpha)^d 2^{1-\nu}}{\Gamma(\nu+d/2)} \|2\pi\alpha x\|^\nu K_\nu(\|2\pi\alpha x\|)$

Noyau de Bessel: $\hat{K}_0^{\rho, \alpha, \sigma}(x) = \rho \frac{(2\pi)^{d/2} \alpha^d \Gamma((\sigma+d+2)/2)}{(\sigma+d)^{d/2} \Gamma((\sigma+2)/2)} \left(1 - \frac{2\pi^2 \alpha^2 \|x\|^2}{\sigma+d}\right)_+^{\sigma/2}$

Problème: On considère l'observation d'une unique réalisation d'un DPP de noyau K^{θ^*} provenant d'une famille paramétrique $\{K^{\theta}, \theta \in \Theta\}, \Theta \subset \mathbb{R}^p$, observée sur une fenêtre W et on cherche à estimer θ^* par **maximum de vraisemblance** des DPPs.



Problème: On considère l'observation d'une unique réalisation d'un DPP de noyau K^{θ^*} provenant d'une famille paramétrique $\{K^{\theta}, \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^p$, observée sur une fenêtre W et on cherche à estimer θ^* par **maximum de vraisemblance** des DPPs.



Ici, l'asymptote se fait sur la **taille de la fenêtre** et donc indirectement sur le **nombre de points observés**.

Outline

- 1 Processus ponctuels determinantaux
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

Théorème (O. Macchi, '75)

Soit X un DPP de noyau K à valeurs propres dans $[0, 1[$ et W un compact de $\mathbb{R}^d$, alors $X \cap W$ est absolument continu par rapport au processus de Poisson d'intensité 1, de densité

$$\forall x \in \cup_n W^n, \quad f(x) = \exp(|W|) \det(\text{Id} - \mathcal{K}|_W) \det((L_W(x_i, x_j))_{i,j})$$

où L_W est solution de l'équation de Fredholm du 2nd type:

$$\forall x, y \in W, \quad L_W(x, y) = K(x, y) + \int_W L_W(x, z) K(z, y) dz.$$

Théorème (O. Macchi, '75)

Soit X un DPP de noyau K à valeurs propres dans $[0, 1[$ et W un compact de $\mathbb{R}^d$, alors $X \cap W$ est absolument continu par rapport au processus de Poisson d'intensité 1, de densité

$$\forall x \in \cup_n W^n, \quad f(x) = \exp(|W|) \det(\text{Id} - \mathcal{K}|_W) \det((L_W(x_i, x_j))_{i,j})$$

où L_W est solution de l'équation de Fredholm du 2nd type:

$$\forall x, y \in W, \quad L_W(x, y) = K(x, y) + \int_W L_W(x, z) K(z, y) dz.$$

Alors, la log-vraisemblance (normalisée) de $X \cap W$ s'écrit:

$$l(\theta|X) = 1 + \frac{1}{|W|} \log \det(\text{Id} - \mathcal{K}^\theta|_W) + \frac{1}{|W|} \log \det((L_W^\theta(x, y))_{x,y \in X \cap W})$$

Théorème (O. Macchi, '75)

Soit X un DPP de noyau K à valeurs propres dans $[0, 1[$ et W un compact de $\mathbb{R}^d$, alors $X \cap W$ est absolument continu par rapport au processus de Poisson d'intensité 1, de densité

$$\forall x \in \cup_n W^n, \quad f(x) = \exp(|W|) \det(\text{Id} - \mathcal{K}|_W) \det((L_W(x_i, x_j))_{i,j})$$

où L_W est solution de l'équation de Fredholm du 2nd type:

$$\forall x, y \in W, \quad L_W(x, y) = K(x, y) + \int_W L_W(x, z) K(z, y) dz.$$

Alors, la log-vraisemblance (normalisée) de $X \cap W$ s'écrit:

$$l(\theta|X) = 1 + \frac{1}{|W|} \log \det(\text{Id} - \mathcal{K}^\theta|_W) + \frac{1}{|W|} \log \det((L_W^\theta(x, y))_{x,y \in X \cap W})$$

Outline

- 1 Processus ponctuels déterminants
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

Idée de la méthode

Regarder le comportement de L_W et $\log\det(\text{Id} - \mathcal{K}|_W)$ lorsque $W \approx \mathbb{R}^d$

Hypothèses supplémentaires

On considère des DPPs stationnaires. Leurs noyaux s'écrivent sous la forme $K(x, y) = K_0(y - x)$.

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\log\det(I_d - \mathcal{K}|_W) = - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k)$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}\log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x\end{aligned}$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}\log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\ &\approx - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W \times (\mathbb{R}^d)^{k-1}} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x\end{aligned}$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}\log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\ &\approx - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W \times (\mathbb{R}^d)^{k-1}} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_W K_0^{*k}(x_1 - x_1) dx_1\end{aligned}$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}\log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\ &\approx - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W \times (\mathbb{R}^d)^{k-1}} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_W K_0^{*k}(x_1 - x_1) dx_1 \\ &= - \sum_{k=1}^{\infty} \frac{|W|}{k} K_0^{*k}(0)\end{aligned}$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}
 \log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\
 &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\
 &\approx - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W \times (\mathbb{R}^d)^{k-1}} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\
 &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_W K_0^{*k}(x_1 - x_1) dx_1 \\
 &= - \sum_{k=1}^{\infty} \frac{|W|}{k} K_0^{*k}(0) = -|W| \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^d} \hat{K}_0^k(x) dx
 \end{aligned}$$

Approximation de $\log\det(I_d - \mathcal{K}|_W)$

$$\begin{aligned}
 \log\det(I_d - \mathcal{K}|_W) &= - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}((\mathcal{K}|_W)^k) \\
 &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W^k} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\
 &\approx - \sum_{k=1}^{\infty} \frac{1}{k} \int_{W \times (\mathbb{R}^d)^{k-1}} K_0(x_2 - x_1) \cdots K_0(x_k - x_{k-1}) K_0(x_1 - x_k) d^k x \\
 &= - \sum_{k=1}^{\infty} \frac{1}{k} \int_W K_0^{*k}(x_1 - x_1) dx_1 \\
 &= - \sum_{k=1}^{\infty} \frac{|W|}{k} K_0^{*k}(0) = -|W| \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^d} \hat{K}_0^k(x) dx
 \end{aligned}$$

$$\log\det(I_d - \mathcal{K}|_W) \approx |W| \int_{\mathbb{R}^d} \log(1 - \hat{K}_0(t)) dt$$

Approximation de L_W

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$

Approximation de L_W

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$
$$\Rightarrow L_W(x, y) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz$$

Approximation de L_W

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_W(x, y) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_0(y - x) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_0(z - x) K_0(y - z) dz$$

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_W(x, y) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_0(y - x) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_0(z - x) K_0(y - z) dz$$

$$\Rightarrow L_0(x) = K_0(x) + K_0 * L_0(x)$$

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_W(x, y) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_0(y - x) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_0(z - x) K_0(y - z) dz$$

$$\Rightarrow L_0(x) = K_0(x) + K_0 * L_0(x)$$

$$\Rightarrow \hat{L}_0(x) = \hat{K}_0(x) + \hat{K}_0(x) \hat{L}_0(x)$$

$$L_W(x, y) = K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_W(x, y) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz$$

$$\Rightarrow L_0(y - x) \approx K_0(y - x) + \int_{\mathbb{R}^d} L_0(z - x) K_0(y - z) dz$$

$$\Rightarrow L_0(x) = K_0(x) + K_0 * L_0(x)$$

$$\Rightarrow \hat{L}_0(x) = \hat{K}_0(x) + \hat{K}_0(x) \hat{L}_0(x)$$

$$\Rightarrow \hat{L}_0(x) = \frac{\hat{K}_0(x)}{1 - \hat{K}_0(x)}$$

$$\begin{aligned}
 L_W(x, y) &= K_0(y - x) + \int_W L_W(x, z) K_0(y - z) dz \\
 \Rightarrow L_W(x, y) &\approx K_0(y - x) + \int_{\mathbb{R}^d} L_W(x, z) K_0(y - z) dz \\
 \Rightarrow L_0(y - x) &\approx K_0(y - x) + \int_{\mathbb{R}^d} L_0(z - x) K_0(y - z) dz \\
 \Rightarrow L_0(x) &= K_0(x) + K_0 * L_0(x) \\
 \Rightarrow \hat{L}_0(x) &= \hat{K}_0(x) + \hat{K}_0(x) \hat{L}_0(x) \\
 \Rightarrow \hat{L}_0(x) &= \frac{\hat{K}_0(x)}{1 - \hat{K}_0(x)}
 \end{aligned}$$

$$\Rightarrow L_W(x, y) \approx L_0(y - x) = \int_{\mathbb{R}^d} \frac{\hat{K}_0(t)}{1 - \hat{K}_0(t)} \exp(2\pi i \langle t, y - x \rangle) dt$$

Au final, on obtient l'approximation suivante de la vraisemblance:

$$\tilde{l}(\theta|X) := 1 + \int_{\mathbb{R}^d} \log(1 - \hat{\kappa}_0^\theta(x)) dx + \frac{1}{|W|} \log \det((L_0^\theta(y-x))_{x,y \in X \cap W})$$

où

$$L_0^\theta(x) = \int_{\mathbb{R}^d} \frac{\hat{\kappa}_0^\theta(t)}{1 - \hat{\kappa}_0^\theta(t)} \exp(2\pi i \langle t, x \rangle) dt$$

Simplification du calcul de L_0 dans certains cas particuliers

- Lorsque K_0 est une fonction radiale alors L_0 peut s'exprimer comme une intégrale 1-dimensionnelle.

Simplification du calcul de L_0 dans certains cas particuliers

- Lorsque K_0 est une fonction radiale alors L_0 peut s'exprimer comme une intégrale 1-dimensionnelle.
- Les noyaux Gaussien

$$K_0^{\rho,\alpha}(x) = \rho \exp\left(-\frac{\|x\|^2}{\alpha^2}\right)$$

vérifient

$$L_0^{\rho,\alpha}(x) = \sum_{n \geq 1} \rho^n \frac{(\sqrt{\pi}\alpha)^{d(n-1)}}{n^{d/2}} \exp\left(-\frac{\|x\|^2}{n\alpha^2}\right).$$

Simplification du calcul de L_0 dans certains cas particuliers

- Lorsque K_0 est une fonction radiale alors L_0 peut s'exprimer comme une intégrale 1-dimensionnelle.
- Les noyaux Gaussien

$$K_0^{\rho, \alpha}(x) = \rho \exp\left(-\frac{\|x\|^2}{\alpha^2}\right)$$

vérifient

$$L_0^{\rho, \alpha}(x) = \sum_{n \geq 1} \rho^n \frac{(\sqrt{\pi} \alpha)^{d(n-1)}}{n^{d/2}} \exp\left(-\frac{\|x\|^2}{n \alpha^2}\right).$$

- Les noyaux de Bessel ($\sigma = 0$) de la forme

$$K_0^{\rho, \alpha}(x) = \rho 2^{d/2} \Gamma(d/2 + 1) \frac{J_{d/2}(\sqrt{2d} \|x\| / \alpha)}{(\sqrt{2d} \|x\| / \alpha)^{d/2}}$$

vérifient

$$L_0^{\rho, \alpha}(x) = \frac{\rho 2^{d/2} \Gamma(d/2 + 1)}{1 - \rho \frac{(2\pi)^{d/2} \alpha^d \Gamma(d/2 + 1)}{d^{d/2}}} \frac{J_{d/2}(\sqrt{2d} \|x\| / \alpha)}{(\sqrt{2d} \|x\| / \alpha)^{d/2}}.$$

Simplification du calcul de L_0 dans certains cas particuliers

- Les noyaux de Cauchy ($\nu = 1/2$) de la forme

$$K_0^{\rho, \alpha}(x) = \frac{\rho}{\left(1 + \left\| \frac{x}{\alpha} \right\| \right)^{\frac{d+1}{2}}},$$

vérifient

$$L_0^{\rho, \alpha}(x) = \sum_{n \geq 1} \frac{\rho^n}{n^d} \left(\frac{\pi^{(d+1)/2} \alpha^d}{\Gamma((d+1)/2)} \right)^{n-1} \frac{1}{\left(1 + \left\| \frac{x}{\alpha} \right\|^2\right)^{(d+1)/2}}.$$

- Les noyaux de Whittle-Matérn de la forme

$$K_0^{\rho, \alpha, \nu}(x) = \rho \frac{2^{1-\nu}}{\Gamma(\nu)} \left\| \frac{x}{\alpha} \right\|^\nu K_\nu \left(\left\| \frac{x}{\alpha} \right\| \right),$$

vérifient

$$L_0^{\rho, \alpha, \nu}(x) = \sum_{n \geq 0} \frac{\rho^n (\sqrt{\pi} \alpha)^{d(n-1)} \Gamma(\nu + d/2)^n}{2^{n\nu-1-(n-1)d/2} \Gamma(\nu)^n \Gamma(n\nu + nd/2)} \left\| \frac{x}{\alpha} \right\|^{n\nu+(n-1)d/2} K_{n\nu+(n-1)d/2} \left(\left\| \frac{x}{\alpha} \right\| \right)$$

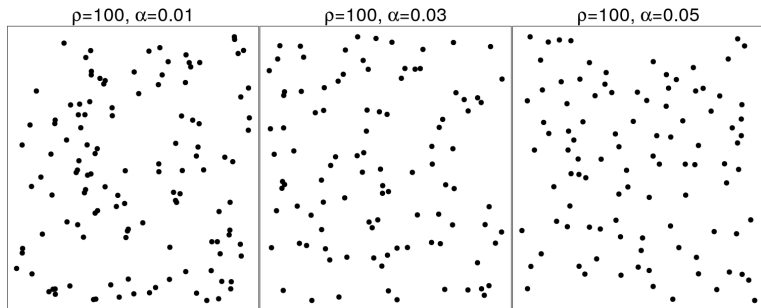
Outline

- 1 Processus ponctuels determinantaux
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - **Résultats numériques**
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

On considère le problème de l'estimation de α pour des DPPs avec noyaux de type Gaussien sur une fenêtre W de $\mathbb{R}^2$:

$$K^{\rho,\alpha}(x,y) = \rho \exp\left(-\frac{\|y-x\|^2}{\alpha^2}\right).$$

On considère les cas où $W \in \{[0,1]^2, [0,2]^2, [0,3]^2\}$ de vrais paramètres $\rho^* = 100$ et $\alpha^* \in \{0.01, 0.03, 0.05\}$.



On compare alors l'estimation de α par notre **approximation du MLE** avec les méthodes de minimum de contraste basées sur la **fonction de corrélation par paire (pcf)** où la **fonction K de Ripley**.

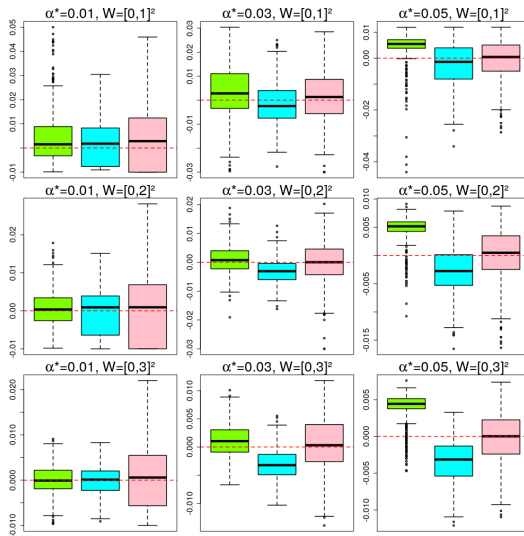


Figure: Boxplots de $\hat{\alpha} - \alpha^*$ pour 500 simulations de DPPs de type **Gaussien**. De gauche à droite: **Approximation du MLE**, **MCE avec la pcf** et **MCE avec K**.

Si on compare l'approximation de la vraisemblance avec sa vraie valeur, on se rend compte que cette approximation ne fonctionne pas bien pour les valeurs de α les plus élevées. Le problème est certainement dû à des effets de bord.

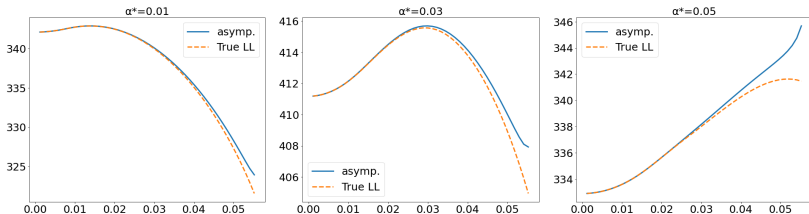


Figure: Graphe de la log-vraisemblance approchée comparée à la vraie vraisemblance pour trois simulations sur $[0, 1]^2$ d'un DPP de paramètre $\rho^* = 100$ et $\alpha \in \{0.01, 0.03, 0.05\}$.

L'utilisation d'une **correction périodique** des effets de bords permet d'avoir une approximation de la log-vraisemblance avec un comportement qui se rapproche beaucoup plus de la vraie vraisemblance.

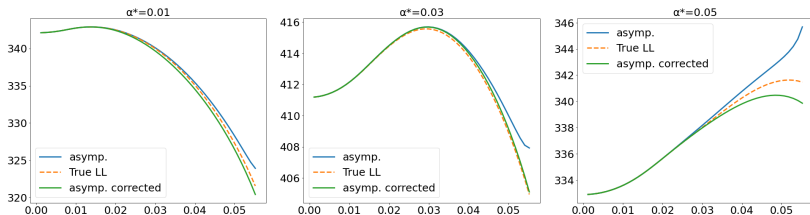


Figure: Graphe de la log-vraisemblance approchée, avec et sans la correction d'effets de bords, comparée à la vraie vraisemblance pour trois simulations sur $[0, 1]^2$ d'un DPP de paramètre $\rho^* = 100$ et $\alpha \in \{0.01, 0.03, 0.05\}$.

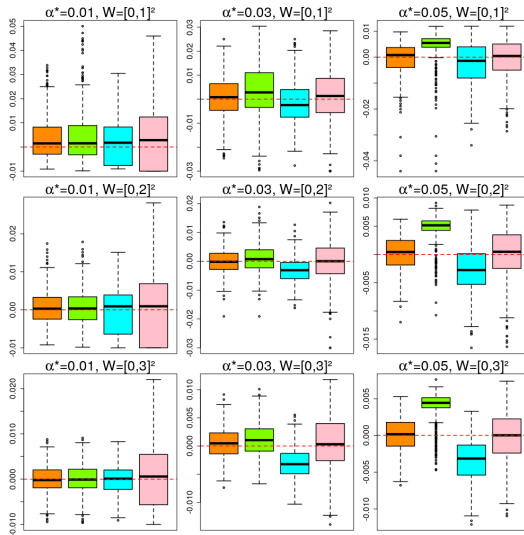
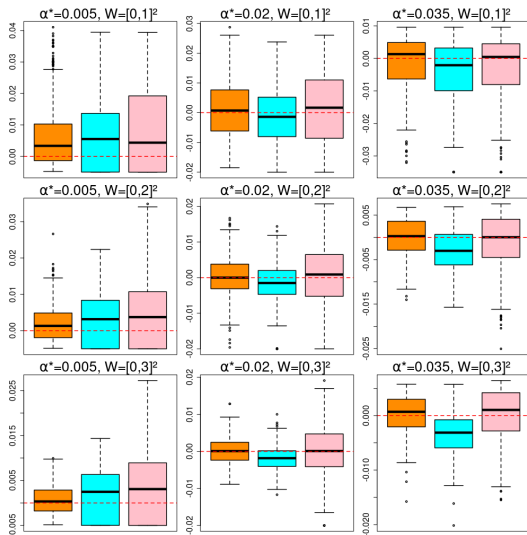
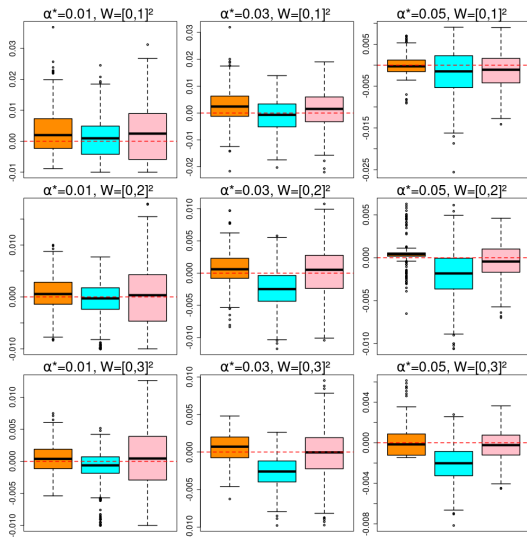


Figure: Boxplots de $\hat{\alpha} - \alpha^*$ pour 500 simulations de DPPs de type **Gaussien**. De gauche à droite: Approximation du MLE avec correction d'effets de bord, Approximation du MLE sans correction d'effets de bord, MCE avec la pcf et MCE avec K.



$$K^{\rho, \alpha}(x, y) = \frac{\rho}{\left(1 + \left\| \frac{y-x}{\alpha} \right\|^2\right)^{3/2}}$$

Figure: Boxplots de $\hat{\alpha} - \alpha^*$ pour 500 simulations de DPPs de type **Cauchy**. De gauche à droite: **Approximation du MLE avec correction d'effets de bord**, **MCE avec la pcf** et **MCE avec K**.



$$K^{\rho, \alpha}(x, y) := \rho \frac{J_1(2 \|\frac{y-x}{\alpha}\|)}{\|\frac{y-x}{\alpha}\|}$$

Figure: Boxplots de $\hat{\alpha} - \alpha^*$ pour 500 simulations de DPPs de type **Bessel**. De gauche à droite: **Approximation du MLE avec correction d'effets de bord**, **MCE avec la pcF** et **MCE avec K**.

Outline

- 1 Processus ponctuels determinantaux
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

Rappel

$$\tilde{l}(\theta|X) := 1 + \int_{\mathbb{R}^d} \log(1 - \hat{K}_0^\theta(x)) dx + \frac{1}{|W|} \log \det((L_0^\theta(y - x))_{x,y \in X \cap W})$$

Rappel

$$\tilde{l}(\theta|X) := 1 + \int_{\mathbb{R}^d} \log(1 - \hat{K}_0^\theta(x)) dx + \frac{1}{|W|} \log \det((L_0^\theta(y-x))_{x,y \in X \cap W})$$

Expression des dérivées secondes de $\tilde{l}(\theta|X)$ pour $\theta = (\theta_1, \dots, \theta_p) \in \Theta \subset \mathbb{R}^p$:

$$\begin{aligned} \partial_{\theta_i} \partial_{\theta_j} \tilde{l}(\theta|X) &= \int_{\mathbb{R}^d} \frac{-\partial_{\theta_i} \partial_{\theta_j} \hat{K}_0^\theta(x) (1 - \hat{K}_0^\theta(x)) - \partial_{\theta_i} \hat{K}_0^\theta(x) \partial_{\theta_j} \hat{K}_0^\theta(x)}{(1 - \hat{K}_0^\theta(x))^2} dx \\ &+ \frac{1}{|W|} \text{Tr}(\partial_{\theta_i} \partial_{\theta_j} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1} \partial_{\theta_i} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1} \partial_{\theta_j} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1}). \end{aligned}$$

Rappel

$$\tilde{l}(\theta|X) := 1 + \int_{\mathbb{R}^d} \log(1 - \hat{K}_0^\theta(x)) dx + \frac{1}{|W|} \log \det((L_0^\theta(y-x))_{x,y \in X \cap W})$$

Expression des dérivées secondes de $\tilde{l}(\theta|X)$ pour $\theta = (\theta_1, \dots, \theta_p) \in \Theta \subset \mathbb{R}^p$:

$$\begin{aligned} \partial_{\theta_i} \partial_{\theta_j} \tilde{l}(\theta|X) &= \int_{\mathbb{R}^d} \frac{-\partial_{\theta_i} \partial_{\theta_j} \hat{K}_0^\theta(x) (1 - \hat{K}_0^\theta(x)) - \partial_{\theta_i} \hat{K}_0^\theta(x) \partial_{\theta_j} \hat{K}_0^\theta(x)}{(1 - \hat{K}_0^\theta(x))^2} dx \\ &+ \frac{1}{|W|} \text{Tr}(\partial_{\theta_i} \partial_{\theta_j} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1} \partial_{\theta_i} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1} \partial_{\theta_j} L_0^\theta[X \cap W] L_0^\theta[X \cap W]^{-1}). \end{aligned}$$

Estimation de l'information de Fisher:

$$\hat{l}(\theta) = -|W| \left(\partial_{\theta_i} \partial_{\theta_j} \tilde{l}(\theta|X) \right)_{1 \leq i, j \leq p}$$

Construction d'intervalle de confiance pour des DPPs
de noyau de la forme $K^{\rho,\alpha}(x, y) = \rho K_0\left(\frac{\|y-x\|}{\alpha}\right)$

Construction d'intervalle de confiance pour des DPPs
de noyau de la forme $K^{\rho, \alpha}(x, y) = \rho K_0\left(\frac{\|y-x\|}{\alpha}\right)$

→ Paramètres estimés par $\hat{\rho} = \text{card}(X \cap W)/|W|$ et $\hat{\alpha} \in \arg \max_{\alpha} \tilde{I}(\hat{\rho}, \alpha | X)$.

Construction d'intervalle de confiance pour des DPPs
de noyau de la forme $K^{\rho, \alpha}(x, y) = \rho K_0\left(\frac{\|y-x\|}{\alpha}\right)$

→ Paramètres estimés par $\hat{\rho} = \text{card}(X \cap W)/|W|$ et $\hat{\alpha} \in \arg \max_{\alpha} \tilde{I}(\hat{\rho}, \alpha|X)$.

→ IC pour α estimé par

$$IC_{95\%}(\alpha) := \left[\hat{\alpha} \pm \frac{1.96}{\sqrt{|W|}} \left(\partial_{\alpha}^2 \tilde{I}(\hat{\rho}, \hat{\alpha}|X) - \frac{(\partial_{\alpha} \partial_{\rho} \tilde{I}(\hat{\rho}, \hat{\alpha}|X))^2}{\partial_{\rho}^2 \tilde{I}(\hat{\rho}, \hat{\alpha}|X)} \right)^{-1/2} \right]$$

Construction d'intervalle de confiance pour des DPPs
de noyau de la forme $K^{\rho, \alpha}(x, y) = \rho K_0\left(\frac{\|y-x\|}{\alpha}\right)$

→ Paramètres estimés par $\hat{\rho} = \text{card}(X \cap W)/|W|$ et $\hat{\alpha} \in \arg \max_{\alpha} \tilde{I}(\hat{\rho}, \alpha|X)$.

→ IC pour α estimé par

$$IC_{95\%}(\alpha) := \left[\hat{\alpha} \pm \frac{1.96}{\sqrt{|W|}} \left(\partial_{\alpha}^2 \tilde{I}(\hat{\rho}, \hat{\alpha}|X) - \frac{(\partial_{\alpha} \partial_{\rho} \tilde{I}(\hat{\rho}, \hat{\alpha}|X))^2}{\partial_{\rho}^2 \tilde{I}(\hat{\rho}, \hat{\alpha}|X)} \right)^{-1/2} \right]$$

⚠ $IC_{95\%}(\alpha)$ n'est pas forcément bien défini!

Window	$[0, 1]^2$			$[0, 2]^2$			$[0, 3]^2$		
α^*	low	mild	high	low	mild	high	low	mild	high
Gauss	88.2 (0)	89.6 (0)	92 (0)	88.6 (0)	94.2 (0)	92.6 (0)	93.2 (0)	93.2 (0)	92.8 (0)
Cauchy	89.8 (2)	88.4 (3.6)	72 (25.4)	92.4 (0)	92.2 (0)	83.6 (0)	91.4 (0)	95 (0)	87.2 (0)
Bessel	66 (0)	76.6 (0.2)	56 (16.2)	77.2 (0)	78.8 (0)	82.2 (7.6)	81 (0)	74.4 (0)	12 (2.2)

Table: Pourcentage de simulation (sur 500) pour lesquelles $IC_{95\%}$ est bien défini et contient α^* . Les valeurs en parenthèse indiquent le pourcentage de simulation pour lesquelles $IC_{95\%}$ n'est pas bien défini.

Outline

- 1 Processus ponctuels determinantaux
 - Jeux de données ponctuelles répulsives
 - Définition des DPPs
 - Construction de la log-vraisemblance
- 2 Approximation asymptotique de la vraisemblance des DPPs
 - Méthode d'approximation
 - Résultats numériques
 - Estimation de la variance
- 3 Conclusion et questions ouvertes

Conclusion

- Les processus déterminant sont utiles pour modéliser des données ponctuelles spatiales possédant un caractère répulsif.
- Leur vraisemblance a une expression explicite mais incalculable numériquement.
- Dans le cas des DPPs stationnaires, on a obtenu une approximation asymptotique de la vraisemblance qui est calculable numériquement.
- Combinée à une correction périodique des effets de bords, cette méthode possède de meilleurs performance que les estimateurs classiques pour une fenêtre rectangulaire (sauf pour le cas Bessel).
- On peut aussi obtenir une estimation de la variance de l'EMV.

Code disponible à l'adresse <https://github.com/APoinas/MLEDPP>.

Questions ouvertes

- Preuve de la consistance de l'EMV.
- Preuve complète de la convergence de l'approximation de la vraisemblance vers la vraisemblance en elle-même.
- Performance d'une approximation de type Nyström de la vraisemblance.

Questions ouvertes

- Preuve de la consistance de l'EMV.
- Preuve complète de la convergence de l'approximation de la vraisemblance vers la vraisemblance en elle-même.
- Performance d'une approximation de type Nyström de la vraisemblance.

MERCI BEAUCOUP

Approximation de $\log\det(Id - \mathcal{K}|_W)$

Proposition

Soit $\{K_0^\theta : \mathbb{R}^d \rightarrow \mathbb{R}, \theta \in \Theta\}$ une famille de fonctions intégrables à transformée de Fourier $\hat{K}_0^\theta$ à valeur dans $[0, M]$ pour $M < 1$ et $(W_n)_{n \geq 0}$ une famille croissante "assez régulière" de sous-ensemble de $\mathbb{R}^d$ telle que $\bigcup_{n \geq 0} W_n = \mathbb{R}^d$.

On suppose:

- $\sup_{\theta \in \Theta} K^\theta(0) < \infty$.
- $x \mapsto \sup_{\theta \in \Theta} |K_0^\theta(x)| \in L^1(\mathbb{R}^d)$.

Alors,

$$\sup_{\theta \in \Theta} \left| \frac{1}{|W_n|} \log\det(\mathcal{I}_{W_n} - \mathcal{K}_{W_n}^\theta) - \int_{\mathbb{R}^d} \log(1 - \hat{K}_0^\theta(x)) dx \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Approximation de L_W

Proposition

Soit $\{K_0^\theta : \mathbb{Z}^d \rightarrow \mathbb{R}, \theta \in \Theta\}$ une famille de fonctions intégrables à transformée de Fourier $\hat{K}_0^\theta$ à valeur dans $[0, M]$ pour $M < 1$ et $(W_n)_{n \geq 0}$ une famille croissante "assez régulière" de sous-ensemble de $\mathbb{Z}^d$ telle que $\bigcup_{n \geq 0} W_n = \mathbb{Z}^d$.

On suppose:

- Θ est un compact de $\mathbb{R}^p$ pour un certain $p \geq 1$
- $(\theta, x) \mapsto \hat{K}^\theta(x)$ est continu sur $\Theta \times \mathbb{T}^d$
- $\exists A, \tau > 0, \forall x \in \mathbb{Z}^d, \sup_{\theta \in \Theta} |K_0^\theta(x)| \leq \frac{A}{1 + \|x\|^{d+\tau}}$.
- $\forall \theta \in \Theta, \forall x \in \mathbb{Z}^d, \hat{K}_0^\theta(x) > 0$

Soit X une réalisation d'un DPP sur $\mathbb{Z}^d$ de noyau K^{θ^*} , $\theta^* \in \Theta$. Alors,

$$\sup_{\theta \in \Theta} \frac{1}{|W_n|} \left| \log \det(L_0^\theta[X \cap W_n]) - \log \det(L_{W_n}^\theta[X \cap W_n]) \right| \xrightarrow{a.s.} 0.$$