Efficient, Cross-Fitting Estimation of Semiparametric Spatial Point Processes

Xindi Lin, Hyunseung Kang

Main Ideas

- **Setup:** In ecology, criminology, and molecular biology, an important goal is to study *spatial treatment effects*, i.e. the relationship between a spatial treatment variable and a spatial point pattern outcome, after adjusting for spatial confounders.

Figure 1: point pattern of cells in tumour tissue [1]

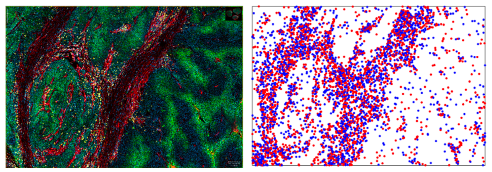
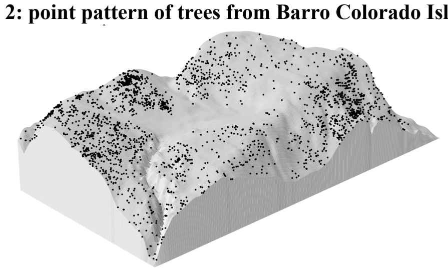


Figure 2: point pattern of trees from Barro Colorado Island [2]



- **Problem**: All existing work assumes parametric models to estimate the treatment effect and none have established optimality in terms of semi-parametric efficiency.
- **Our Contribution**: We propose a *semiparametric spatial point process model* where the outcomes are potentially dependent spatial point processes, the treatment effect is parametric and the confounder effect is nonparametric.
  1. We construct a spatial, cross-fitting estimator of the treatment effect.
  2. The proposed estimator is *consistent* and *asymptotically Normal*.
  3. The proposed estimator is *semiparametrically efficient* for Poisson processes.
  4. We propose a *computationally efficient* algorithm to compute the estimator.

Model: Semiparametric Spatial Point Processes

- Let $u \in \mathbb{R}^2$ be points, $x(u)$ be treatment; $z(u)$ be confounders, $A \subset \mathbb{R}^2$ be the observational window
- Y is a spatial point process parametrized by the *semiparametric intensity function*

$$\lambda(u;\theta,\eta) = \Psi(\theta^\top x(u) + \eta(z)), \quad \theta \in \mathbb{R}^k, \eta \in H$$

- θ is the target parameter, $\eta(\cdot)$ is nuisance parameter in a linear space H
- $g(u,v)$: pair correlation function (PCF) measures dependence between points.

V-Fold Cross-Fitting for Spatial Point Processes

- The pseudo-log-likelihood [4] of the model is

$$\ell(\theta,\eta) = \sum_{u \in Y \cap A} \log \lambda(u;\theta,\eta) - \int_A \lambda(u;\theta,\eta)du$$

- **Takeaways:** we extend semiparametric efficiency bounds [5] to spatial point processes and propose cross-fitting estimator [3] that can achieve this bound.

Assumptions

- **Smoothness** of $\lambda(u;\theta,\eta)$
- **Boundedness** of $\lambda(u;\theta,\eta)$, $g(u,v)$, $x(u)$, $z(u)$
- **Weak Dependence** that $g(u,v) \approx \|u-v\|^{-\alpha}$

Estimation

1. **Random Thinning** (i.e. common in Bayesian MCMC for point processes): For every $u \in Y \cap A$, sample v from $\{1,\ldots,V\}$, assign u to Y_v. Then we obtain subprocesses $Y_1,\ldots,Y_V$
2. **Estimation:** For every $v \in \{1,\ldots,V\}$,

   (i) **Nuisance Estimation**: Fix θ, $Y_v^c = \cup_{j \neq v} Y_j$

   $$\hat\eta_\theta^{(v)}(z) = \arg\max_{\gamma \in \mathbb{R}} \hat{E}\left[\ell(\theta,\gamma;Y_v^c) \mid z\right]$$

   We modify the classic *Nadaraya-Watson estimator* for point processes to achieve the desired rates (see below).

   (ii) **Target Estimation**: Plug in $\hat\eta_\theta^{(v)}$ and solve

   $$\hat\theta^{(v)} = \arg\max_\theta \ell(\theta,\hat\eta_\theta^{(v)};Y_v)$$

   (iii)**Aggregate estimator** $\hat\theta = \frac{1}{V}\sum_{i=1}^{V}\hat\theta^{(v)}$

Computation

- For computation, we use a **quadrature approximation** of the likelihood where

$$\ell(\theta,\eta) \approx \sum_{u_i \in Y \cap A} \log \lambda(u_i;\theta,\eta) - \sum_{j=1}^{m} \lambda(u_j;\theta,\eta)w_j$$
$$= \sum_{j=1}^{m} (y_j \log \lambda(u_j;\theta,\eta) - \lambda(u_j;\theta,\eta))w_j$$

- $\{u_j\}_{j=1}^m$ are quadrature points containing Y, $\{w_j\}_{j=1}^m$ are quadrature weights, $y_j = I(u_j \in Y \cap A)$.
- RHS is the log-likelihood of Poisson regression and thus, can be solved with existing R packages: gplm, mgcv, SemiPar.

Variance Estimation

- We propose a spatial, sandwich variance estimator: $\hat{S}^{-1}\hat\Sigma\hat{S}^{-1}$ where

$$\hat{S} = \sum_{j=1}^{m} w_j \tilde{x}(u_j)\tilde{x}(u_j)^\top \lambda(u_j)$$

$$\hat\Sigma = \hat{S} + \sum_{i,j=1, i\neq j}^{m} w_i w_j \tilde{x}(u_i)\tilde{x}(u_j)^\top \lambda(u_i)\lambda(u_j)(g(u_i,u_j)-1)$$

- $\tilde{x}(u)$ are projection residuals (talk to me for more details)

Asymptotic Properties

- Consider $A_1 \subset A_2 \cdots \subset A_n \in \mathbb{R}^2$ where $|A_n| \to \infty$,

Theorem 1: $\hat\theta \to_p \theta$ if nuisance estimator is uniformly consistent

Theorem 2: $\hat\theta$ is asymptotically Normal if nuisance estimator converges at rate $O_p(|A_n|^{-1/4})$

Theorem 3: $\hat\theta$ is semiparametrically efficient if Y is a Poisson process.

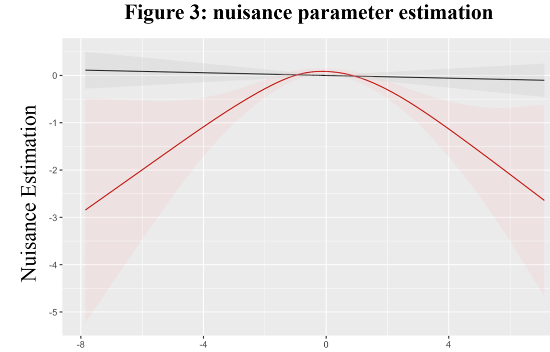
Simulation: Poisson & LGCP

- $y(u), z(u)$ be dependent Gaussian random fields
- $[0,2] \times [0,2]$ be the observational window
- $\lambda(u) = \exp(400 + \theta y(u) + \eta(z(u)))$
- $\theta = 0.3$, $\eta(z) = -0.09z^2$
- Poisson process, log-Gaussian Cox process (LGCP)
- Parametric v.s. Ours v.s. Oracle (nuisance is known)

Table 1: simulation result

Model	Method	$\text{Bias}_{\times 100}$	SE	CP95 (%)
Poisson	Ours	-0.0219	0.0275	96.2
Poisson	Oracle	-0.3648	0.0265	96.0
Poisson	Parametric	-2.5624	0.0257	85.8
LGCP	Ours	-0.7064	0.0326	95.1
LGCP	Oracle	-0.6360	0.0327	94.6
LGCP	Parametric	-2.6446	0.0258	77.6

Figure 3: nuisance parameter estimation



Application: Elevation Effect of *Beilschmiedia pendula* Tree

- Y: tree locations, $x(u)$: elevation, $z(u)$:gradient

Figure 4: fitted intensity of parametric (Top), our method (Bottom)

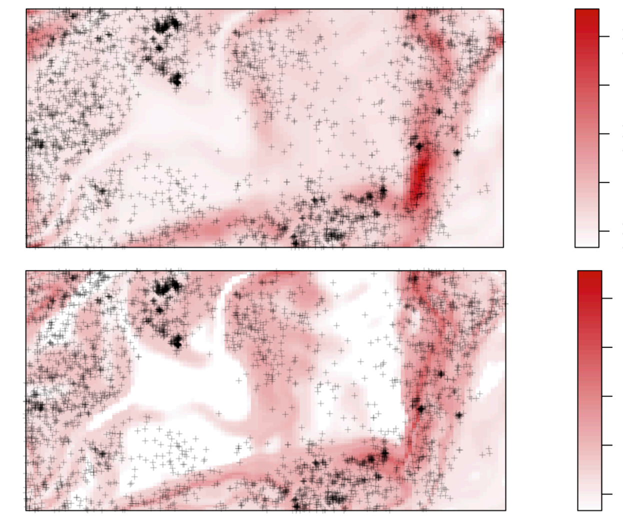


Figure 5: nuisance parameter estimation

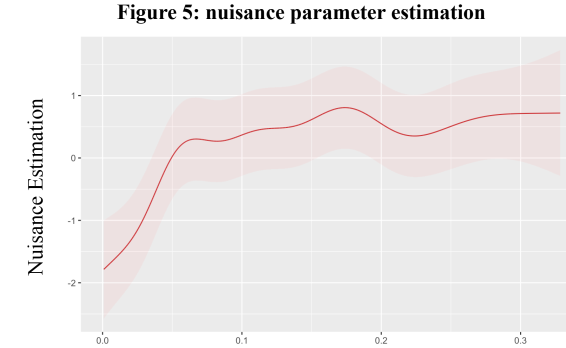


Table2: estimation of elevation effect

Method	$\hat\theta$	SE($\hat\theta$)	95%CI
Ours	2.983	2.268	(-1.462,7.428)
Parametric	2.144	2.346	(-2.454,6.742)

- Our method is more robust to non-linear trends in the gradient (i.e. z(u))
- Our method captures the spatial pattern of the trees better than the parametric method, which is prone to model misspecification.

References

[1] Kristian Bjørn Hessellund, Ganggang Xu, Yongtao Guan, and Rasmus Waagepetersen. *Journal of the Royal Statistical Society Series C: Applied Statistics, 71(1):244–268, 2022b*

[2] Adrian Baddeley, Ege Rubak, and Rolf Turner. *Spatial Point Patterns: Methodology and Applications with R*. CRC press,2015

[3] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen,Whitney Newey, and James Robins. *The Econometrics Journal, 21(1):C1–C68, 2018*

[4] Rasmus Waagepetersen and Yongtao Guan. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 71(3):685–702, 2009*

[5] Thomas A Severini and Wing Hung Wong. *The Annals of Statistics, pages 1768–1802, 1992*