Single-Index Model for Inhomogeneous Spatial Point Processes

Yixin Fang and Ji Meng Loh

New York University and New Jersey Institute of Technology

Supplementary Material

Appendix

A.1 Lemmas

Lemmas A.1-A.3 are from Ichimura (1993). We include them here for the reader's easy reference. Lemmas from A.4 onwards are specific to our paper.

Lemma A.1. *Let f be the density of a random variable U and function g be a function $\mathbb{R} \to \mathbb{R}$. Assume that $E\{g(U)/h_n K[(u-U)/h_n]\}$ exists. If function gf is twice continuously differentiable, the second derivative satisfies the Lipschitz condition, K satisfies Assumption A7, and u is an interior point of the support of U, then for $h_n > 0$ and $h_n \to 0$,*

$$\left|E\{g(U)/h_n K[(u-U)/h_n]\} - g(u)f(u)\right| = O(h_n^2).$$

Lemma A.2. *Let f be the density of a random variable U and function g be a function $\mathbb{R} \to \mathbb{R}$. Assume that $E\{g(U)/h_n^2 K'[(u-U)/h_n]\}$ exists. If function gf is twice continuously differentiable, the second derivative satisfies the Lipschitz condition, K satisfies Assumption A7, and u is an interior point of the support of U, then for $h_n > 0$ and $h_n \to 0$,*

$$\left|E\{g(U)/h_n^2 K'[(u-U)/h_n]\} - [g(u)f(u)]'\right| = O(h_n^2).$$

Lemma A.3. *Let f be the density of a random variable U and function g be a function $\mathbb{R} \to \mathbb{R}$. Assume that $E\{g(U)/h_n^3 K''[(u-U)/h_n]\}$ exists. If function gf is three times continuously differentiable, the third derivative satisfies the Lipschitz condition, K satisfies Assumption A7, and u is an interior point of the support of U, then for $h_n > 0$ and $h_n \to 0$,*

$$\left|E\{g(U)/h_n^3 K''[(u-U)/h_n]\} - [g(u)f(u)]''\right| = O(h_n^2).$$

Lemma A.4. *Under Assumptions A1-A7, as $n \to \infty$, $h_n \to 0$ and $nh_n^2 \to \infty$,*

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} \frac{1}{|W_n| h_n} \Big| \sum_{t \in X \cap W_n} K\{|\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}|/h_n\} - \int_{W_n} K\{|\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}|/h_n\} \times \rho(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0)\, dt \Big| = o_p(1).$$

Proof of Lemma A.4. Without loss of generality, assume that $W_n = [-n, n) \times [-n, n)$ and partition it into $4n^2$ small windows $W_{ij}^n = [i, i+1) \times [j, j+1)$, $i, j = -n, \cdots, n-1$. Let $N_{ij}^n = X(W_{ij}^n)$. Partition $\{\boldsymbol{\beta} : \|\boldsymbol{\beta}\| = 1\}$ into n^{p-1} non-overlapped small regions I_k^n, $k = 1, \cdots, n^{p-1}$ such that the length of each region is C/n. Select one point β_k^n from I_k^n and denote $U_k^n = \{\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_k^n : s \in W_n\}$. Further partition U_k^n into n small, non-overlapping intervals U_{kl}^n, $l = 1, \cdots n$, such that the length of each interval is C/n. Select one point u_{kl}^n from U_{kl}^n. Then there are n^p small regions $I_k^n \times U_{kl}^n$ and point $(\boldsymbol{\beta}_k^n, u_{kl}^n)$ is selected from each region.

Denote $S_n(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta}) = \frac{1}{|W_n| h_n} \sum_{t \in X \cap W_n} K\{|\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}|/h_n\}$ and it suffices to show that, for any $\epsilon > 0$ and $\eta > 0$, when n is large,

$$Pr\{\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} \Big| S_n(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta}) - E_Z\{S_n(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})\} \Big| > \epsilon\} < \eta,$$

where E_Z is expectation given $\mathbf{Z}(\cdot)$. Let $M_n = n^{1/3}$. The left-hand side of the above is less than

$$Pr\{\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} |S_n(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta}) - E_Z\{S_n(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})\}| > \epsilon, \max_{ij} N_{ij}^n \le M_n\} \quad (A.1)$$

$$+ Pr\{\max_{ij} N_{ij}^n > M_n\}. \quad (A.2)$$

The term (A.2) is less than $\sum_{ij} Pr\{N_{ij}^n > M_n\}$, which is less than $\eta/2$ when n is large, by Markov inequality and Assumption A2. Now we show (A.1) is less than $\eta/2$ when n is large. Note that for each $\boldsymbol{\beta}$ and s, $(\boldsymbol{\beta}, \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})$ belongs to a small region, say $I_k^n \times U_{kl}^n$. Let

$$h(\boldsymbol{\beta}, u) = |S_n(u; \boldsymbol{\beta}) - E_Z\{S_n(u; \boldsymbol{\beta})\}|.$$

Then the supremum of the term (A.1) is less than

$$\max_{k,l} \sup_{\boldsymbol{\beta} \in I_k^n} \sup_{u \in U_{kl}^n} |h(\boldsymbol{\beta}, u) - h(\boldsymbol{\beta}_k^n, u_{kl}^n)| + \max_{k,l} |h(\boldsymbol{\beta}_k^n, u_{kl}^n)|.$$

Because the size of these small regions is C/n, $\max_{k,l}\sup_{\boldsymbol{\beta}\in I_k^n}\sup_{u\in U_{kl}^n}|h(\boldsymbol{\beta},u)-h(\boldsymbol{\beta}_k^n,u_{kl}^n)|=O_p((nh_n^2)^{-1})$, which is $o_p(1)$ if $nh_n^2\to\infty$. Hence, it suffices to show that, when n is large enough,

$$Pr\{\max_{k,l}|h(\boldsymbol{\beta}_k^n,u_{kl}^n)|>\epsilon,\max_{i,j}N_{ij}^n\le M_n\}<\eta/2. \tag{A.3}$$

Note that $h(\boldsymbol{\beta}_k^n,u_{kl}^n)=|\sum_{ij}Z_{ij}|/(|W_n|h_n)$, where

$$Z_{ij}=\sum_{s\in X\cap W_{ij}^n}K(|\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_k-u_{kl}|/h_n)-E\{\sum_{s\in X\cap W_{ij}^n}K(|\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_k-u_{kl}|/h_n)\}.$$

Also note that under $\max_{i,j}N_{ij}^n\le M_n$, Z_{ij} is bounded by CM_n. Therefore, under the mixing condition stated in Assumption A5, by the Bernstein's inequality developed in Lemma 4.7 of Zhu and Lahiri (2007), we have

$$\begin{aligned}P\{|\sum_{i,j}Z_{ij}|>\xi_n|W_n|\big|\max_{i,j}N_{ij}^n\le M_n\}&\le C_1\exp\left(-\frac{C_2(\lambda_n/b_n)^4\xi_n^2}{M_n^2+(\lambda_n/b_n)M_n\xi_n}\right)\\&\quad+C_1(\lambda_n/b_n)^2(M_n/\xi_n)^{1/2}\alpha(C_2b_n;\lambda_n^2).\end{aligned}$$

If M_n is chosen as $n^{1/3}$, λ_n as n, b_n as $n^{1/3}$, and $\xi_n=h_n\epsilon$, the above upper-bound becomes $C_1\exp(-C_2n^{1/2})+Cn^2n^{-\tau/3}n^{2\delta}$. Then we have

$$Pr\{|h(\boldsymbol{\beta}_k^n,u_{kl}^n)|>\epsilon\big|\max_{i,j}N_{ij}^n\le M_n\}Cn^2n^{-\tau/3}n^{2\delta},\ \text{and thus}$$

$$Pr\{\max_{k,l}|h(\boldsymbol{\beta}_k^n,u_{kl}^n)|>\epsilon\big|\max_{i,j}N_{ij}^n\le M_n\}\le Cn^{2+p+2\delta}n^{-\tau/3}.$$

Hence, (A.3) holds for $\tau>6+6\delta+3p$ and thus Lemma A.4 is proved. ■

Lemma A.5. *Under Assumptions A1, A3, A6 and A7, as $n\to\infty$, $h_n\to 0$ and $nh_n^2\to\infty$,*

$$\sup_{\|\boldsymbol{\beta}\|=1}\sup_{s\in W_n}\left|\frac{1}{|W_n|h_n}\int_{W_n}K(|\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}|/h_n)\,dt-f(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta})\right|=o_p(1).$$

Proof of Lemma A.5. By Assumption A1 and Lemma A.1, we have

$$\left|E\left\{\frac{1}{|W_n|h_n}\int_{W_n}K(|\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}-u|/h_n)\,dt\right\}-f(u;\boldsymbol{\beta})\right|=O(h_n^2).$$

Under the strong mixing assumption A6 of $\mathbf{Z}(\cdot)$, Lemma A.5 can be proved following arguments similar to those in the proof of Lemma A.4. ■

Lemma A.6. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^2 \to \infty$,

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} \Bigg| \frac{1}{|W_n| h_n} \int_{W_n} K(|\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}|/h_n) \rho(\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta}_0)\, dt - E\{\rho(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}_0) | \mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}\} f(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta}) \Bigg| = o_p(1).$$

Proof of Lemma A.6. By Assumption A1 and Lemma A.1, we have

$$\left| E\left\{ \frac{1}{|W_n| h_n} \int_{W_n} K(|\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta} - u|/h_n) \rho(\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta}_0)\, dt \right\} - \rho^*(u; \boldsymbol{\beta}) f(u; \boldsymbol{\beta}) \right| = O(h_n^2).$$

Under the strong mixing Assumption A6 of $\mathbf{Z}(\cdot)$, Lemma A.6 can be proved following arguments similar to those in the proof of Lemma A.4. ■

Lemma A.7. *Under Assumptions A1-A7, if* $n \to \infty$, $h_n \to 0$ *and* $nh_n^2 \to \infty$,

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} |\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}; \boldsymbol{\beta}) - \rho^*(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta})| \to 0.$$

Proof of Lemma A.7. It suffices to show that

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} |\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta}) - E_Z\{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta})\}| \to 0,$$
$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} |E_Z\{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta})\} - \rho^*(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta})| \to 0.$$

The first result follows from Lemmas A.4 and A.5, noting $\inf_{\|\boldsymbol{\beta}\|=1} f(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}; \boldsymbol{\beta}) \geq c > 0$ in Assumption A3. The second result is equivalent to

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} \left| \frac{\int_{W_n} K(|\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}|/h_n) \rho(\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta}_0)\, dt}{\int_{W_n} K(|\mathbf{Z}(t)^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}|/h_n)\, dt} - E\{\rho(\mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}_0) | \mathbf{Z}(s)^{\mathrm{T}} \boldsymbol{\beta}\} \right|$$

converging to 0, which follows from Lemmas A.5 and A.6. ■

Lemma A.8. *Under Assumptions A1-A7,*

$$Pr\Big\{ \sup_{\|\boldsymbol{\beta}\|=1} |l_n(\boldsymbol{\beta}) - E\{l_n(\boldsymbol{\beta})\}| \geq \varepsilon \Big\} \to 0.$$

Proof of Lemma A.8. Let

$$\begin{aligned} Q_n(\boldsymbol{\beta}) &= \frac{1}{|W_n|}\sum_{s\in X\cap W_n} \log\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta}), \\ Q_n^*(\boldsymbol{\beta},b) &= \frac{1}{|W_n|}\sum_{s\in X\cap W_n} \sup_{\widetilde{\boldsymbol{\beta}}\in B(\boldsymbol{\beta},b)} \log\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\widetilde{\boldsymbol{\beta}};\widetilde{\boldsymbol{\beta}}), \quad \text{and} \\ Q_{n*}(\boldsymbol{\beta},b) &= \frac{1}{|W_n|}\sum_{s\in X\cap W_n} \inf_{\widetilde{\boldsymbol{\beta}}\in B(\boldsymbol{\beta},b)} \log\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\widetilde{\boldsymbol{\beta}};\widetilde{\boldsymbol{\beta}}), \end{aligned}$$

where $B(\boldsymbol{\beta},b)=\{\widetilde{\boldsymbol{\beta}}:\|\widetilde{\boldsymbol{\beta}}\|=1,\|\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}\|\le b\}$. By Assumptions A1 and A4, we have

$$\overline{\lim}_{b\to 0}\sup_{n\ge 1}\left|E\{Q_n^*(\boldsymbol{\beta},b)\}-E\{Q_n(\boldsymbol{\beta})\}\right|=0,\ \text{for any}\ \boldsymbol{\beta},$$

noting that $\left|E\{Q_n^*(\boldsymbol{\beta},b)\}-E\{Q_n(\boldsymbol{\beta})\}\right|\le E\{\left|Q_n^*(\boldsymbol{\beta},b)-Q_n(\boldsymbol{\beta})\right|\}$, which is controlled by

$$\frac{1}{|W_n|}E\Big\{\sum_{s\in X\cap W_n}\sup_{\widetilde{\boldsymbol{\beta}}\in B(\boldsymbol{\beta},b)}\left|\log\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\widetilde{\boldsymbol{\beta}};\widetilde{\boldsymbol{\beta}})-\log\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta})\right|\Big\}\le C.$$

Likewise, $\overline{\lim}_{b\to 0}\sup_{n\ge 1}\left|E\{Q_{n*}(\boldsymbol{\beta},b)\}-E\{Q_n(\boldsymbol{\beta})\}\right|=0$.

Given $\epsilon>0$, for any $\boldsymbol{\beta}$, there exists $b(\boldsymbol{\beta})>0$ such that for $n\ge 1$,

$$E\{Q_n(\boldsymbol{\beta})\}-\epsilon\le E\{Q_{n*}(\beta,b(\boldsymbol{\beta}))\}\le E\{Q_n^*(\beta,b(\boldsymbol{\beta}))\}\le E\{Q_n(\boldsymbol{\beta})\}+\epsilon.$$

The collection of balls $\{B(\beta,b(\boldsymbol{\beta})):\|\boldsymbol{\beta}\|=1\}$ is an open cover of the compact set $\{\boldsymbol{\beta}:\|\boldsymbol{\beta}\|=1\}$, and hence, has a finite subcover $\{B(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l)):l=1,\cdots,L\}$.

For any $\boldsymbol{\beta}\in B(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))$, we have

$$\begin{aligned} Q_n(\boldsymbol{\beta})-E\{Q_n(\boldsymbol{\beta})\} &\le Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\} \\ &\le Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}+2\epsilon, \end{aligned}$$

and likewise, $Q_n(\boldsymbol{\beta})-E\{Q_n(\boldsymbol{\beta})\}\ge Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}-2\epsilon$. Then, for any $\boldsymbol{\beta}$,

$$\begin{aligned} &\min_{l\le L}\left[Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}\right]-2\epsilon\le Q_n(\boldsymbol{\beta})-E\{Q_n(\boldsymbol{\beta})\} \\ &\qquad\le \min_{l\le L}\left[Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}\right]+2\epsilon. \end{aligned}$$

Lemma A.8 results if we prove, for each l, that

$$Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_{n*}(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}=o_p(1)\ \text{and}\ Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))-E\{Q_n^*(\boldsymbol{\beta}_l,b(\boldsymbol{\beta}_l))\}=o_p(1).$$

To prove this, it suffices to prove $Var(Q_{n*}(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))) \to 0$ and $Var(Q_n^*(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))) \to 0$. If we let $h(s) = \sup_{\widetilde{\boldsymbol{\beta}} \in B(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))} \log \rho^*(\mathbf{Z}(s)^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}; \widetilde{\boldsymbol{\beta}})$, where $|h(s)| \le \log(C)$, then, by Campbell's Theorem,

$$Var(Q_n^*(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))) = \frac{1}{|W_n|^2} \Big\{ \int_{W_n \times W_n} h(s)h(t)\lambda(s)\lambda(t)[g(s,t)-1]\, ds\, dt + \int_{W_n} [h(s)]^2 \lambda(s)\, ds \Big\},$$

where $\lambda(s) = E\{\lambda(s|\,\mathbf{Z}(s))\}$ and $g(s,t)$ is the pair correlation function of X. Assumption A5 implies that

$$\sup_{s,t\in\mathbb{R}^2} |g(s,t)| \le C \text{ and } \sup_{s\in\mathbb{R}^2} \int_{\mathbb{R}^2} |g(s,t)-1|\, dt \le C.$$

These assure that $Var(Q_n^*(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))) \to 0$. Likewise, $Var(Q_{n*}(\boldsymbol{\beta}_l, b(\boldsymbol{\beta}_l))) \to 0$. ■

Lemma A.9. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^3 \to \infty$,

$$\sup_{\|\boldsymbol{\beta}\|=1} \sup_{s \in W_n} \left| \frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})}{d\boldsymbol{\beta}} - \frac{d\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})}{d\boldsymbol{\beta}} \right| = o_p(1),$$

where

$$\frac{d\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})}{d\boldsymbol{\beta}} = \frac{\partial \rho^*(u; \boldsymbol{\beta})}{\partial u}\Big|_{\{u = \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}\}} \Big[\mathbf{Z}(s) - E\{\mathbf{Z}(s)|\, \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}\} \Big].$$

Proof of Lemma A.9. Note that

$$\frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})}{d\boldsymbol{\beta}} = \frac{\sum_{t\in X\cap W_n} \frac{1}{|W_n|h_n} dK((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)/d\boldsymbol{\beta}}{\frac{1}{|W_n|h_n}\int_{W_n} K((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)\, dt} \quad (A.4)$$

$$- \frac{\sum_{t\in X\cap W_n} \frac{1}{|W_n|h_n} K((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)}{\frac{1}{|W_n|h_n}\int_{W_n} K((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)\, dt} \quad (A.5)$$

$$\times \frac{\frac{1}{|W_n|h_n}\int_{W_n} dK((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)/d\boldsymbol{\beta}\, dt}{\frac{1}{|W_n|h_n}\int_{W_n} K((\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta})/h_n)\, dt}. \quad (A.6)$$

By Lemma A.5, the denominator of (A.4)-(A.6) converges uniformly to $f(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}; \boldsymbol{\beta})$. The numerate of (A.4) equals

$$\sum_{t\in X\cap W_n} \frac{1}{|W_n|h_n^2} K'((\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta})/h_n)[\mathbf{Z}(s) - \mathbf{Z}(t)],$$

which, by Lemma A.2 and the arguments in the proof of Lemma A.4, converges uniformly to

$$\frac{\partial\big\{f(u;\boldsymbol{\beta})\rho^*(u;\boldsymbol{\beta})[\mathbf{Z}(s)-E\{\mathbf{Z}(s)|\,\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}=u\}]\big\}}{\partial u}\Big|_{\{u=\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}\}}.$$

Similarly, the numerator of (A.6) converges uniformly to

$$\frac{\partial\big\{f(u;\boldsymbol{\beta})[\mathbf{Z}(s)-E\{\mathbf{Z}(s)|\,\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}=u\}]\big\}}{\partial u}\Big|_{\{u=\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}\}}.$$

By Lemma A.6 and the arguments in the proof of Lemma A.4, (A.5) converges uniformly to $\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta})$. Combining these three results completes the proof of Lemma A.9. ■

Lemma A.10. *Under Assumptions A1, A3, A6 and A7, as* $n\to\infty$, $h_n\to 0$ *and* $nh_n^4\to\infty$,

$$\sup_{\|\boldsymbol{\beta}\|=1}\sup_{s\in W_n}\left|\frac{d^2\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta})}{d\boldsymbol{\beta}\,d\boldsymbol{\beta}^{\mathrm{T}}}-\frac{d^2\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta};\boldsymbol{\beta})}{d\boldsymbol{\beta}\,d\boldsymbol{\beta}^{\mathrm{T}}}\right|=o_p(1).$$

Proof of Lemma A.10. By Lemma A.3, following similar arguments as in the proof Lemma A.9, we can show Lemma A.10.

Lemma A.11. *Under Assumptions A1, A3, A6 and A7, as* $n\to\infty$, $h_n\to 0$ *and* $nh_n^4\to\infty$,

$$\frac{1}{\sqrt{|W_n|}}\sum_{s\in X\cap W_n}\frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\boldsymbol{\beta}}{[\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)]^2}\Big[\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)-\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)\Big]=o_p(1).$$

Lemma A.12. *Let* $\Delta_n(s)=d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\boldsymbol{\beta}-d\rho^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\boldsymbol{\beta}$. *Under Assumptions A1, A3, A6 and A7, as* $n\to\infty$, $h_n\to 0$ *and* $nh_n^4\to\infty$,

$$\frac{1}{\sqrt{|W_n|}}\sum_{s\in X\cap W_n}\left[\frac{\Delta_n(s)}{\rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)}-\int_{W_n}\Delta_n(t)\,dt\right]=o_p(1).$$

Proofs of Lemmas A.11 and A.12 The proofs of Lemmas A.11 and A.12 are tedious, but not difficult since the convergence is at the true parameter $\boldsymbol{\beta}_0$ instead of uniformly over the parameter space. They can be shown following arguments similar to those in the proofs of Lemmas 5.8 and 5.9 in Ichimura (1993). ■

Lemma A.13. *Let* $\widetilde{\mathbf{Z}}(s) = Z(s) - E\{\mathbf{Z}(s)|\,\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0\}$. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^4 \to \infty$,

$$\frac{1}{\sqrt{W_n}}\Sigma_n^{-1/2}\Bigg[\sum_{s\in X\cap W_n}\frac{\rho'(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)\widetilde{\mathbf{Z}}(s)}{\rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)} - \int_{W_n}\rho'(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0)\widetilde{\mathbf{Z}}(t)\,dt\Bigg] \xrightarrow{D} \mathrm{MVN}(\mathbf{0}, I).$$

Proof of Lemma A.13. Note that this result is stated for the case where ρ is known. This result was proved in Waagepetersen and Guan (2009), where ρ was assumed to be known. ■

Lemma A.14. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^4 \to \infty$,

$$\sqrt{|W_n|}\Sigma_n^{-1/2}\frac{d\widehat{l}_n(\boldsymbol{\beta}_0)}{d\boldsymbol{\beta}} \xrightarrow{D} \mathrm{MVN}(\mathbf{0}, I).$$

Proof of Lemma A.14. Note that

$$\frac{d\widehat{l}_n(\boldsymbol{\beta}_0)}{d\boldsymbol{\beta}} = \frac{1}{|W_n|}\Bigg[\sum_{s\in X\cap W_n}\frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\,\boldsymbol{\beta}}{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)} - \int_{W_n} d\widehat{\rho}^*(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\,\boldsymbol{\beta}\,dt\Bigg].$$

Then the lemma follows from Lemmas A.11 and A.12. ■

Lemma A.15. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^4 \to \infty$,

$$-\frac{d^2\widehat{l}_n(\boldsymbol{\beta}_0)}{d\boldsymbol{\beta}d\boldsymbol{\beta}^{\mathrm{T}}} = V_n + o_p(1).$$

Proof of Lemma A.15. Note that

$$\begin{aligned}\frac{d^2\widehat{l}_n(\boldsymbol{\beta}_0)}{d\boldsymbol{\beta}d\boldsymbol{\beta}^{\mathrm{T}}} = \frac{1}{|W_n|}\Bigg[\sum_{s\in X\cap W_n}&\frac{d^2\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\,\boldsymbol{\beta}\,d\,\boldsymbol{\beta}^{\mathrm{T}}}{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)} - \int_{W_n}\frac{d^2\widehat{\rho}^*(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)}{d\,\boldsymbol{\beta}\,d\,\boldsymbol{\beta}^{\mathrm{T}}}\,dt\Bigg] \\ &-\frac{1}{|W_n|}\sum_{s\in X\cap W_n}\left(\frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\,\boldsymbol{\beta}}{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)}\right)\left(\frac{d\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)/d\,\boldsymbol{\beta}}{\widehat{\rho}^*(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0;\boldsymbol{\beta}_0)}\right)^{\mathrm{T}}.\end{aligned}$$

The first term on the right-hand-side converges to zero following Lemma A.10 and the second term converges to V_n following Lemma A.9. ■

Lemma A.16. *Under Assumptions A1, A3, A6 and A7, as* $n \to \infty$, $h_n \to 0$ *and* $nh_n^4 \to \infty$, *for any* $\varepsilon > 0$ *there exists a neighborhood of* $\boldsymbol{\beta}_0$, $\mathbf{B}_0$, *such that*

$$Pr\Big\{\sup_{\boldsymbol{\beta}\in\mathbf{B}_0}\Big|\frac{d^2\widehat{l}_n(\boldsymbol{\beta})}{d\boldsymbol{\beta}d\boldsymbol{\beta}^{\mathrm{T}}} - \frac{d^2\widehat{l}_n(\boldsymbol{\beta}_0)}{d\boldsymbol{\beta}d\boldsymbol{\beta}^{\mathrm{T}}}\Big| > \varepsilon\Big\} \to 0.$$

Proof of Lemma A.16. The lemma follows from Lemmas A.9 and A.10 and the continuity stated in Assumption A4. ■

A.2 Proof that D_n is non-negative

For an arbitrary set of locations in W_n, $\mathcal{J} = \{s_1, \cdots, J\}$, let $I(ds_j)$ be the indicator that there is an event in region ds_j, $j = 1, \cdots, J$. Denote vector $(I(ds_1), \cdots, I(ds_J))^{\mathrm{T}}$ as $\mathbf{Y}(\mathcal{J})$. Then given covariate process $\mathbf{Z}$, the covariance matrix of $\mathbf{Y}(\mathcal{J})$ is $\mathrm{Cov}(\mathbf{Y}(\mathcal{J})) = (S_{s,t})_{s,t\in\mathcal{J}}$, which is positive semi-definite, where $S_{s,t} = \rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)\rho(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0)[g(s,t) - 1]$. Hence, for any function function $h(s)$, we have $\sum_{s\in\mathcal{J}}\sum_{t\in\mathcal{J}} h(s)h(t)\rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)\rho(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0)[g(s,t) - 1] \geq 0$. Further,

$$\frac{1}{|W_n|}\int_{W_n\times W_n} h(s)h(t)\rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)\rho(\mathbf{Z}(t)^{\mathrm{T}}\boldsymbol{\beta}_0)[g(s,t) - 1]\, ds\, dt \geq 0.$$

Letting $h(s) = \rho'(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)\widetilde{\mathbf{Z}}(s)/\rho(\mathbf{Z}(s)^{\mathrm{T}}\boldsymbol{\beta}_0)$, it shows that $D_n \geq 0$. ■

References

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics* **58**, 71–120.

Waagepetersen, R. and Guan, Y. (2009). Two-step estimation for inhomogeneous spatial point process. *Journal of the Royal Statistical Society, Series B* **71**, 685–702.

Zhu, J. and Lahiri, S. N. (2007). Bootstrappoing the empirical distribution function of a spatial process. *Statistical Inference for Stochastic Processes* **10**, 107-145.