On local polynomial estimation of hazard rates and their derivatives under left truncation and right censoring models

Abstract

Estimating hazard rate function is an important problem in survival analysis. There are some estimation approaches based on kernel smoothing. However, they suffer from the boundary effects or need high order kernels, which increases the mean squared error. We introduce local polynomial estimators of hazard rates and their derivatives for the left truncation and right censoring models. The estimators have favorable properties similar to those of local polynomial regression estimators. Asymptotic expressions for the mean squared errors (AMSE’s) are obtained. Consistency and joint asymptotic normality of the local polynomial estimators are established. A data-based local bandwidth selection rule is proposed.

Keywords: censoring, data-driven local bandwidth selection, hazard rate, local polynomial, truncation

Introduction

Consider a subject in survival studies. Only if its onset time, i.e. the time origin of its lifetime, passes the beginning of the study, the subject can enter into the study. For the entered individuals, each of them is then followed for a fixed time point. Such subjects are so-called left truncated and right censored. To be specific, let \( X \) be the lifetime with distribution function \( F(x) \), \( T \) the random left truncation time with arbitrary distribution \( G(x) \), and \( C \) the random censoring time with arbitrary distribution \( L(x) \), where \( s \) is independent of \( (T, C) \). Then the cumulative hazard function of \( F(x) \) is \( \Lambda(x) = \int_0^x dF(t)/(1 - F(t)) \). Under the left truncation and right censoring model, one observes only those i.i.d. \((Y, C)\), where and throughout the paper for any \( dF \). \( E \), \( E = 1 - E \) is the corresponding survival function.

The left truncation and right censoring model has been investigated by many authors. Interesting work along the field can be found in Gross & Lai and Gijbels & Wang among others. Several authors have considered the estimation of hazard functions under the left truncation and right censoring model. For examples, Uzunogullari & Wang, and Wu & Wells.

In the present investigation, we study local polynomial (LP for short) estimators of hazard functions and their derivatives based on the i.i.d sample \( \{Y_i, T_i, \delta_i\}_{i=1}^{n} \) from \((Y, T, \delta)\). Under the left truncation and right censoring model, one observes only those i.i.d pairs \((Y_i, T_i, \delta_i)\) for which \( Y_i \geq T_i \). It is worth pointing out that our estimators inherit some favorable properties from local polynomial regression estimators, in particular, our estimators can reduce the bias according to the degree of the polynomial without increasing the variance and automatically correct the left boundary effect. The point wise asymptotic normality of our estimators enables one to find the asymptotically optimal variable bandwidth choice, and thereafter allows one to develop a data-driven optimal local bandwidth selector by using the ideas of Fan & Gijbels. We here present a simpler data-driven method for choosing the local bandwidth.

The outline of this paper is as follows. In Section 2, we introduce the LP estimators. In Section 3 we concentrate on the asymptotic properties of the proposed estimators, including point wise strong consistency and joint asymptotic normality. In Section 4, we propose the data-driven local bandwidth selection rule. Technical proofs are given in the Appendix.

Estimation

In order to introduce the LP estimators, we use the following notation from Gijbels & Wang\(^1\)

\[
C(z) = P(T \leq z \leq Y | Y \geq T) = \alpha^{-1}P(T \leq z \leq C)\left[1 - F(z)\right]
\]

\[C_n(z) = (n + 1)^{-1}\sum_{i=1}^{n}I(T_i \leq z \leq Y_i),\]

the modified empirical estimator of \(C(z)\).

\[F^*_n(z) = 1 - \prod_{i=1}^{n} \left(1 - \left(\frac{C_n(Y_i)}{n}\right)^{1/n}\right),\] the product-limit estimator of \(F(z)\) from Tsai\(^1\).

Let \( a_d = \inf\{t : E(t) > 0\} \) and \( b_d = \inf\{t : E(t) = 1\} \) denote the left and right endpoints of the support for any d.f. \(E(x)\), respectively. Then \(F(x)\) is identifiable if \(a_d \leq a_\delta \) and \( b_d \leq b_\delta\) (Woodroofe\(^6\) and Gijbels & Wang\(^1\)).

Therefore, we assume this condition holds. As in Gijbels & Wang\(^1\), for estimating the density function \(f(x)\) of \(X\), we also assume \(a_d \leq a_\delta\).

Following Gijbels & Wang\(^1\), we define \(W_{\alpha}(y) = P(Y \leq y, \delta = 1 | Y \geq T)\) and \(W_{\alpha}(y) = n^{-1}\sum_{i=1}^{n}I(Y_i \leq y, \delta_i = 1)\). Then \(dW_{\alpha}(y) = \alpha^{-1}P(T \leq Y \leq C)dF(y)\) and

\[f_{\alpha}(y) = \frac{dW_{\alpha}(y)}{\alpha^{-1}P(T \leq Y \leq C)}\]
On local polynomial estimation of hazard rates and their derivatives under left truncation and right censoring models

The method can be also adapted to the case of estimating the hazard function \( \lambda(x) \) and its derivatives if one uses \( \Lambda_n \) instead of \( \hat{F}_n \): \( \hat{\lambda}(x) = \frac{1}{h_n} \int K_n \left[ \frac{x - u}{h_n} \right] d\Lambda_n(u). \)

Following Jiang and Doksum, the estimator is an extension to that of Müller & Wang where right censoring model is considered. However, for the estimation of derivatives or reduction of bias, the estimator needs higher order kernels, which can lead to a negative hazard rate estimator. The practical advantages of using higher order kernels can be quite small for moderate sample sizes as demonstrated in Marron & Wand. When estimating at point \( x \) near \( a_0 \) or \( b_0 \), the effective support [\( x - b_0, x + b_0 \)] of the kernel is not contained in \( [a_0, b_0] \), most kernel estimators in density estimation and regression settings will encounter boundary effects. The estimator (3) suffers from boundary effects near the endpoints of the support of the hazard rates. In the presence of censoring for estimating hazard rate function, Müller & Wang solved the problem by employing boundary kernels and a data-adaptive varying bandwidth selection procedure. Hall & Wehrly used a geometrical method for removing edge effects from kernel-type nonparametric regression estimators. These boundary correction methods may also be adapted to the estimation approach in (3). Here we introduce a simple and intuitive approach to the problem. Our approach does not need higher order kernels or boundary kernels while automatically correcting the boundary effects. Our idea is similar to that of Jiang & Doksum, but their procedure cannot be directly applied to the current setting.

Following Jiang and Doksum, we consider the following optimization problem:

\[
\min_{\Lambda_n} \frac{1}{b_n} \int K_n \left[ \frac{x - Y_j}{b_n} \right] \left( \lambda(u) - \frac{1}{n} \sum_{j=1}^{n} g(u) \right)^2 \, du.
\]

By Taylor expansion, the solution of the optimization problem, \( \mathbf{a}^* = (a_{1}^*, \ldots, a_{p}^*)^T \), will estimate \( \mathbf{a}(x) = (\lambda(x), \ldots, \lambda^{(p)}(x)/p)^T \). Since \( \Lambda_n \) in (1) is the empirical estimator of \( \Lambda(x) \), we define the following generalized empirical hazard rate as the generalized derivative of \( \Lambda_n(x) \):

\[
\hat{\lambda}_n(x) = \frac{1}{m} D(x - Y_j) \delta_j,
\]

where \( D(x) \) is the Dirac function with the following property:

\[
\int g(u)D(u-x)\,du = g(x)
\]

For any integral function \( g(x) \). Then \( \int_0^1 \hat{\lambda}_n(t)\,dt = \Lambda_n(x) \), which is why we call \( \hat{\lambda}_n \) the generalized empirical hazard rate. Replacing \( \hat{\lambda}(x) \) in (4) by \( \hat{\lambda}_n(x) \), we obtain that

\[
\hat{\mathbf{a}}(x) = (\hat{\lambda}_{n1}, \ldots, \hat{\lambda}_{np})^T = \arg\min_{\mathbf{a}} \left[ \frac{1}{b_n} \int K_n \left[ \frac{x - Y_j}{b_n} \right] \lambda(u) - \frac{1}{n} \sum_{j=1}^{n} a_j (u-x)^{j-1} \, du \right].
\]

Then \( \hat{\mathbf{a}}(x) \) is the LP estimator of \( \mathbf{a}(x) \). Jones considered a locally linear estimator and established its link with the generalized jackknife boundary correction for \( p = 1 \) by using the ideas of Lejeune & Sarda in local linear fitting to the empirical distribution \( F_n \). Here we study the local polynomial estimation of hazard functions and their derivatives, \( \mathbf{a}(x) \), under the left truncation and right censoring model. Obviously, our method can be used for the complete data case, which corresponds to “no truncation and no censoring”.

Taking the derivative with respect to \( a_i \) of the integral in (6), we obtain the LP estimator \( \hat{\mathbf{a}}(x) \) as the solution to the linear equations: for \( \ell = 0, \ldots, p \),

\[
\sum_{j=1}^{n} \frac{1}{b_n} K_n \left[ \frac{Y_j - x}{b_n} \right] (Y_j - x) \delta_j = \frac{p}{n} \sum_{j=1}^{n} a_j (u-x)^{j-1} \frac{1}{b_n} K_n \left[ \frac{Y_j - x}{b_n} \right] \, du.
\]

It follows that the LP estimator at any point \( x_0 \in (a_0, b_0) \) has the following closed form:

\[
\hat{\mathbf{a}}(x_0) = S^{-1} \mathbf{S}_n(x_0),
\]

where \( \mathbf{S}_n(x_0) = (S_{n1}(x_0), \ldots, S_{np}(x_0)) \), and

\[
S_{n\ell}(x_0) = \sum_{j=1}^{n} \frac{1}{b_n} K_n \left[ \frac{Y_j - x_0}{b_n} \right] (Y_j - x_0) \delta_j = \frac{1}{b_n} \int K_n \left[ \frac{Y_j - x_0}{b_n} \right] \, du.
\]

When \( p = 0 \) and \( s_0 = 1 \),

\[
\hat{\lambda}_0 = \frac{1}{b_n} \int K_n \left[ \frac{Y_j - x_0}{b_n} \right] \delta_j = \frac{1}{b_n} \int K_n \left[ \frac{Y_j - x_0}{b_n} \right] \, du,
\]

which is the same as the kernel estimator of \( \lambda(x) \) in (3). However, this equivalence does not hold for boundary points.

We will show in next section that the LP estimator shares nice properties with the local polynomial regression estimator, in particular, the estimator will keep its convergence rate up to the left boundary point, i.e. the estimator automatically corrects the left boundary effect, which contrasts with the results for other hazard rate estimators.

Asymptotic properties

In this section, we will establish the consistency and joint asymptotic normality of the local polynomial estimators. To this end, we introduce some regularity conditions. For a given point \( x_0 \in (a_0, b_0) \), the following notations and assumptions are needed.

The hazard rate function \( \lambda(x) = \Lambda'(x) \) has a continuous \( (p+1)th \) derivative at the point \( x_0 \).

The sequence of bandwidths \( b_n \) tends to zero such that \( nb_n \to +\infty \) as \( n \to +\infty \). Let \( B = \text{diag}(b_0, \ldots, b_p) \).

\( C(x) \) is continuous at the point \( x_0 \).
The kernel function $K(x)$ is a continuous function of bounded variation and with bounded support $[-1, 1]$, say. Let $s_j = \int_0^1 K(u) du$ , $v_i = \int_{-1}^1 K^2(u) du$, for $\ell \geq 0$, $c_i = (s_{i+\ell}, \ldots, s_{2i+\ell})^T$, $S = (s_{ij})$ and $V^* = (v_{ij})$ (for $0 \leq i \leq p$, $0 \leq j \leq p$) be $(p+1) \times (p+1)$ matrices.

**Theorem 3.1** Under conditions (A1) – (A4),

$$B(a(x_0) - a(x_0)) \to 0, \quad n \to \infty.$$  

**Theorem 3.2** Under conditions (A1) – (A4),

$$\sqrt{nb_n}[B(a(x_0) - a(x_0))]_{p+1} \to N(0, S^{-1}V^*S^{-1}2(x_0)C^{-1}(x_0)).$$  

**Remark 3.1** When estimating a hazard rate which is a polynomial of order $p$ on an interval, the finite sample bias of the LP estimators on the interval is zero (see the proof of Theorem 3.2). This contrasts the methods of Müller and Wang, based on higher order kernels, for which the respective zero bias only holds true asymptotically.

**Remark 3.2** When there is no truncation, we take $T = 0$ and $G(x) = 1$ over the support of $F(x)$, then $C(x) = a^T P(G \geq x) \hat{F}(x) = \hat{W}(x)$, and the asymptotic normality for interior point $x_0$ is the same as in Müller & Wang. When there is no censoring, we take the censoring variable $C = a^T x$ and $L(x) = 0$ over the support of $F(x)$, then $C(x) = P(T \leq x \leq X \mid Y \geq T)$, and the asymptotic normality for interior point $x_0$ is the same as in Gürler & Wang.

$$AMSE_k(b_n - x_0) = \frac{1}{\hat{b}^-_{a, x_0}}[c_i^T S^{-1} e_i^T A(x_0) C(x_0)],$$

$$b_{kopt}(x_0) = n^{-\frac{1}{2(p+3)}} \left( \frac{(p+1)!}{2(p+1-k)!} \right)^{\frac{1}{2(p+3)}} [c_i^T S^{-1} e_i^T A(x_0) C(x_0)]^{1/2(p+3)}.$$  

**Theorem 3.3** Consider the left edge effect on the estimator. Assume that we estimate $a(x)$ at $x = d b_n$ in the left boundary region for some positive constant $d \in [0, 1]$. Then similar to (8), $\tilde{a}(x_0)$ in (6) has the following closed form, for $p = 0, 1, 2, \ldots$,

$$\tilde{B}_d(x_0) = S_d^T S_d(x_0),$$  

Where $S_d$ defined as $S$ but with $t_j$ replaced by $s_{jd} = \int_0^1 K(u) du$. Let $V_{jd} = \int_{-1}^1 K^2(u) du$. Then the joint asymptotic normality (10) continues to hold with $c_{ip}$, $S$ and $V$ replaced by $c_{ipd}, S_d$ and $V_d$, respectively, where $c_{ipd} = (s_{ip+1}, \ldots, s_{2ip+1})^T$ and $V_{jd} = (v_{ij})$ is $(p+1) \times (p+1)$ matrices.

This property of our estimator in Theorem 3.3 is similar to that of local polynomial regression estimation, which is not shared by other kernel estimators of hazard rates (Hess et al. 14). The LP estimators are automatically boundary adaptive in the sense of Fan & Gijbels. Note that the above property holds even for $p = 0$, which contrasts with the cases of local polynomial regression estimation.

**Remark 3.3** For a finite sample, one may encounter right boundary effects when estimating near $T$. A good method for dealing with the problem is to use the following estimator similar to the estimator in (13):

$$\tilde{B}_d(x_0) = S_d^T S_d(x_0),$$  

Where $x_0 = T - q b_n$ for $q \in [0, 1]$ and $S_d$ defined as $S$ but with $s_j$ replaced by $s_{jd} = \int_0^1 K(u) du$.

**Data-driven local bandwidth choice**

The proposed estimators depend on the band width $b_n$. It is important to develop a local bandwidth choice for estimating hazard functions and their derivatives, especially when one would like to have a data-driven approach to bandwidth selection. Here we extend the data-driven local bandwidth choice of Jiang & Doksum11 to the left truncation and right censoring model.

From the proof of Theorem 3.2, we see that the exact bias of the estimator $\hat{a}_k(x_0)$ is

$$B_k(b_n, x_0) = e_i^T (S^{-1}_k - 9 S^{-1}) \hat{a}(x_0) = \hat{B}_d(b_n, x_0),$$  

Where $\hat{B}_d(b_n, x_0) = (\hat{b}_d(b_n, x_0) \ldots \hat{b}_d(b_n, x_0) )^T$ and $\hat{b}_d(x_0) = \int [K(t) d\lambda(x_0 + b_n t)] dt$. The asymptotic variance of $\hat{a}_k(x_0)$ is

$$V_k(b_n, x_0) = \frac{1}{nh^2_k} e_i^T (S^{-1}_k - 9 S^{-1}) e_i,$$  

Where $\lambda = (\hat{\lambda}(x_0), \hat{\lambda}(x_0 + b_n))$. Then we propose to estimate the AMSE of $\hat{a}_k(x_0)$ via

$$0.362 \cdot AMSE_k(b_n, x_0) = \hat{B}_d^2(b_n, x_0) + \hat{V}_d(b_n, x_0),$$  

Where $\hat{B}_d(b_n, x_0)$ and $\hat{V}_d(b_n, x_0)$ are defined similarly to $B_d(b_n, x_0)$ and $V_d(b_n, x_0)$ but with $\lambda(x)$ replaced by a pilot estimator and $C(x)$ replaced by its empirical estimator $C_n(x)$. We define

$$B_{k, opt}(x_0) = \arg \min \{ 0.362 \cdot AMSE_k(b_n, x_0) \}.$$  

Then, for estimating $\lambda(x)$ when $p=1$, the following algorithm similar to Jiang and Doksum$^1$ can be used for estimating $\lambda(x)$.

**Algorithm for estimating $\lambda(x)$:**

1. **Pilot estimators of $\lambda(x)$:** Choose a kernel, such as the Epanechnikov kernel, and an initial global bandwidth $b_0$. The choice of the initial bandwidth depends on the specific case. Assume the data are available on $[0,T]$, then a possible value for $b_0$ is $T/(8n_{ux})$, where $n_{ux}$ is the number of uncensored observations. The pilot estimators $\hat{\lambda}(x)$ of $\lambda(x)$ are obtained by using $b_0(x)=b_0$ and our estimators (8), (13) and (14).

2. **Minimizing of $^{0362}AMSE(b_0,x)$:** Choose an equispaced grid of $m_1$ points $\hat{x}_i$, $i=1,$ $\ldots$, $m_1$ between 0 and $T$. For each of the grid points $\hat{x}_i$ compute $^{0362}AMSE(b_0,\hat{x}_i)$ in (17) with $k=0$ and obtain its minimizing $b_0(\hat{x}_i)$ on the interval $[\hat{x}_i/4, 4\hat{x}_i]$, say. The minimization of $^{0362}AMSE(b_0,\hat{x}_i)$ may be computed via Discretisation.

3. **Smoothing bandwidths:** Choose another equispaced grid of $m_2$ points $x_r$, $r=1,$ $\ldots$, $m_2$, over the interval $[0,T]$ on which the final hazard estimator is desired. Running local linear smoother by employing global bandwidth $b_0=0$ or $\hat{b}_0$;

$$\hat{\beta}(x_i) = \sum_{j=1}^{m_1} w_j \hat{b}(\hat{x}_j)/\sum_{j=1}^{m_1} w_j,$$

where $w_j = K(\hat{x}_j-x_i)/b_0$ and $M_{x,y} = \sum_{i=1}^{m_1} K((\hat{x}_i-x_i)/b_0)(\hat{x}_i-x_i)^T$, for $j=1,2$.

4. **Final hazard function estimators:** Using (8), (13) and (14), obtain the estimators $\hat{\lambda}(x)$ by employing the bandwidth $b_r(x)$, for $r=1,$ $\ldots$, $m_2$.

The above algorithm can be repeated by using the estimators $\hat{\lambda}(x)$ in Step 4 as pilot estimators in Step 1 and running Step 2–Step 4 again. The pilot estimators of $\hat{\lambda}(x)$ in Step 1 can also be obtained via maximum likelihood if one has a plausible parametric model in mind.$^{18}$

**Appendix proofs of theorems**

**Proofs of theorems 3.1 and 3.2**

i. Using (1) and (9), we obtain that

$$S_\omega(x_0) = \frac{1}{b_0} \sum_{i=1}^{m_1} K(\hat{X}_i-x_0)/b_0 d\Lambda_\omega(u) \Lambda_\omega(u)$$

$$= \frac{1}{b_0} \sum_{i=1}^{m_1} K(\hat{X}_i-x_0)/b_0 d\Lambda(u) + \frac{1}{b_0} \sum_{i=1}^{m_1} K(\hat{X}_i-x_0)/b_0 d\Lambda(u)-\Lambda(u)$$

$$= \beta_\omega(x_0) + \gamma_\omega.$$

Let $\beta_\omega(x_0) = (\beta_{\omega,0}, \ldots, \beta_{\omega,p})^T$ and $\gamma_\omega(x_0) = (\gamma_{\omega,0}, \ldots, \gamma_{\omega,p})^T$. Then

$$S_r(x_0) = \beta_r(x_0) + \gamma_r(x_0).$$

By (8), we have

$$\mathbb{B}([a(x_0)-a(x_0)]) = [S^{-1}_r \beta_r(x_0) - S^{-1}_a \beta_a(x_0)] + S^{-1}_r \gamma_r(x_0)$$

$$= \alpha_r(x_0) + S^{-1}_r \gamma_r(x_0).$$

We will show that $\alpha_r(x_0)$ contributes to the bias term, and $S^{-1}_r \gamma_r(x_0)$ the variance term of our estimator.

(a) By (18), a change of variable for integration, and the Taylor expansion, we get

$$\beta_r(x_0) = \int K(t)^T \Delta(x_0+b_0 t) dt$$

$$= \int \frac{b_0^i}{(p+1)!} \Delta^{(i)}(x_0) b_{r,i} + \frac{b_0^{p+1}}{(p+1)!} \Delta^{(p+1)}(x_0) b_{r,p+1} + p_4[1+a_r(x_0)].$$

Then

$$\alpha_r(x_0) = \frac{b_0^{p+1}}{(p+1)!} \Delta^{(p+1)}(x_0) S^{-1} \mathbb{E} P^{1+a_r(x_0)}.$$

(b) For $a_\omega < a_\omega$, by Gijbels & Wang, we have

$$\Lambda_\omega(x)-\Lambda(x) = \frac{1}{n \beta_\omega} \sum_{i=1}^{m_1} \int_{C(u)} \Delta_i(x) \Delta_i(t) dt$$

$$= O\left(\frac{n}{\beta_\omega}\right),$$

where for $x \geq 0$, any $b < b_\omega$ and $\delta = 1$ or 0, $\sup_{0 \leq t \leq b} |t(x)| = O\left(\frac{n}{\beta_\omega}\right)$, a.s. and

$$\epsilon(x,y,t,\delta,x) = I(y \leq x, \delta = 1) / C(y) - \int_y^t I(t \leq u \leq y) [C(u)^{-1}]^{2} dW_f(u),$$

which satisfies that $E\epsilon(x,Y_t,\delta,t) = 0$, and $\text{Cov}(\sum_{i=1}^{m_1} \epsilon(X_i,Y_t,\delta,t)) = g(m_1(x,t))$, where $g(x) = [C(u)^{-1}]^{2} dW_f(u)$. Let $K_r(t) = K(t)^T$. Using the definition of $\epsilon(x,t)$, (21) and integration by parts, we obtain the following almost surely representation of $\epsilon_r(x)$:

$$\epsilon_r(x_0) = \sigma_r(x_0) + \epsilon_r(x_0),$$

where $\sigma_r(x_0) = -n^{-\frac{1}{2}} \int_{b_0}^{b_0} K(u) b_{r,0} d\Lambda(u)-\Lambda(u)$$

$$= (nh_{r})^{-\frac{1}{2}} \sum_{i=1}^{m_1} \int_{c_0}^{c_0} \epsilon(Y_i,T_i,\delta_i,x_0+b_0 t) dt |K_r(t)|$$

is the stochastic component of $S_r(x_0)$ and contributes to the variance of our estimator, and $\epsilon_r(x_0)$ is the negligible error of the approximation which satisfies

$$\sup_{0 \leq t \leq b} |\epsilon_r(x)| = O\left(\frac{n}{\beta_\omega}\right),$$

for $0 \leq t \leq p$. Note that $E(\sigma_r(x_0)) = 0$ and


On local polynomial estimation of hazard rates and their derivatives under left truncation and right censoring models.
On local polynomial estimation of hazard rates and their derivatives under left truncation and right censoring models.

\[ \text{Cov}(\sigma_m(x_0), \sigma_n(x_0)) = \frac{1}{nb_x} \int_{-\infty}^{\infty} (g'(u) + b_m g(u)) dK_m(u) dK_m(v) \]

\[ = \frac{1}{nb_x} \int_{-\infty}^{\infty} \frac{1}{b_n} g(\min(x_0 + b_m u, x_0 + b_m v)^2) dK_m(u) dK_m(v) \]

\[ = \frac{1}{nb_x} \frac{1}{b_n} \frac{\lambda(x_0)}{C(x_0)} v_n(1 + o(1)) \]

(24)

Let \( \sigma_m(x_0) = (\sigma_{m1}, \ldots, \sigma_{mp})^T \) and \( e_n(x_0) = (e_{n1}, \ldots, e_{np})^T \). Then by (22) and (24)

\[ \sqrt{n} \varepsilon_n(x_0) = \sqrt{n} \sigma_m(x_0) + \sqrt{n} \varepsilon_n(x_0) \]

and

\[ \text{Cov}(\sigma_m(x_0), \sigma_n(x_0)) = \frac{1}{nb_n} \frac{\lambda(x_0)}{C(x_0)} v_n(1 + o(1)) \]

(25)

By the central limit theorem, we get

\[ \sqrt{n} \varepsilon_n(x_0) \rightarrow N(0, \frac{\lambda(x_0)}{C(x_0)} v_n) \]

(26)

By (23), we know

\[ \sup_{\varepsilon_n(x_0)} |S \varepsilon_n(x_0)| = O(\frac{\log n}{n}) \text{ a.s.} \]

(27)

Then by (25), (26) and (27)

\[ \sqrt{n} \varepsilon_n(x_0) \rightarrow N(0, \frac{\lambda(x_0)}{C(x_0)} V_n^{-1} V_n) \]

(28)

Combination of (19), (20) and (28) completes the proof of Theorem 3.2.

ii. Note that, \( g(u) \) are i.i.d.'s sums. By the SLN, we know Then . This combined with (22) and (23) yields . Therefore, by (19) and (20),

\[ \sqrt{n} e_n(x_0) \rightarrow 0 \]

Proofs of Theorem 3.3 The result follows by the same argument as in part (i) of the proof of Theorem 3.2.

Acknowledgements

None.

Conflicts of interest

Author declares that there are no conflicts of interest.

References
