NONPARAMETRIC COVARIANCE MODEL

Jianxin Yin[1], Zhi Geng[1], Runze Li[2], and Hansheng Wang[1]

[1]*Peking University* & [2]*Pennsylvania State University*

Supplementary Material

This appendix contains technical proofs of theorems in the main text, and also the details of a simulation study. Before we presenting the detailed theorem proofs, the following lemma is necessarily needed.

Appendix A. Technical lemma and its proof

Lemma. Let $G \subset \{u : f(u) > 0\}$ be a compact subset on the support of U, where $f(u)$ is the density of the *index* variable U. As $n \to \infty$, we then have

$$\hat{m}_j(u) - m_j(u) = \frac{1}{nhf(u)} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \Big\{X_{ij} - m_j(U_i)\Big\} + \frac{h^2\mu_2}{2}\left[2\frac{\dot{f}(u)}{f(u)}\dot{m}_j(u) + \ddot{m}_j(u)\right] + \mathcal{O}_p\{R_{n,1}^{(j)}(u)\} \tag{A.1}$$

$$\hat{\sigma}_{j_1j_2}(u) - \sigma_{j_1j_2}(u) = \frac{1}{nhf(u)} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right)\left[\Big\{X_{ij_1} - \hat{m}_{j_1}(U_i)\Big\}\Big\{X_{ij_2} - \hat{m}_{j_2}(U_i)\Big\} - \sigma_{j_1j_2}(u) - (U_i - u)\dot{\sigma}_{j_1j_2}(u)\right] + h^2\dot{\sigma}_{j_1j_2}(u)\mu_2^{\kappa}\frac{\dot{f}(u)}{f(u)} + \mathcal{O}_p\{R_{n,2}^{(j_1,j_2)}(u)\}, \tag{A.2}$$

where $\mathcal{O}_p\{R_n(u)\}$ with an arbitrary random variable R_n is a random quantity satisfying that $\sup_{u\in G} |\mathcal{O}_p\{R_n(u)\}/R_n(u)| = O_p(1)$. Similarly, $\mathbf{o}_p\{R_n(u)\}$ stands for a random quantity satisfying that $\sup_{u\in G} |\mathbf{o}_p\{R_n(u)\}/R_n(u)| = o_p(1)$. Finally, we have

$$R_{n,1}^{(j)}(u) = \frac{1}{nf(u)}\left\{\left|\sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right)\Big(X_{ij} - m_j(U_i)\Big)\right|\right\} + o(h^2)$$

$$R_{n,2}^{(j_1,j_2)}(u) = \frac{1}{nf(u)}\Bigg\{\Bigg|\sum_{i=1}^{n} K\left(\frac{U_i-u}{h}\right)\Bigg[\Big\{X_{ij_1}-\hat{m}_{j_1}(U_i)\Big\}\Big\{X_{ij_2}-\hat{m}_{j_2}(U_i)\Big\}$$

$$-\sigma_{j_1j_2}(u)-(U_i-u)\dot{\sigma}_{j_1j_2}(u)\Bigg]\Bigg|\Bigg\}+o(h^2).$$

Proof: Denote $k_i(u) = K\{(U_i - u)/h\}$, we have

$$\frac{1}{nh}\sum_{i=1}^{n} k_i(u) = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{U_i-u}{h}\Big) = \hat{f}(u).$$

Then, by Li (2006, pp. 32), we know that $|\hat{f}(u)-f(u)| = \mathcal{O}_p\{\log(n)^{1/2}/(nh)^{1/2}+h^2\}$ under technical conditions (C1) and (C6). Then under condition (C1), we have uniformly for $u \in G$,

$$\left(\frac{1}{nh}\sum_{i=1}^{n} k_i(u)\right)^{-1} = f^{-1}(u)\Bigg\{1+\mathcal{O}_p\bigg(\frac{\log(n)^{1/2}}{(nh)^{1/2}}+h^2\bigg)\Bigg\}.$$

Define for $j = 1, 2$,

$$s_j(u) = \frac{1}{nh}\sum_{i=1}^{n}\Big(\frac{U_i-u}{h}\Big)^{j} K\Big(\frac{U_i-u}{h}\Big). \tag{A.3}$$

Next, by the same argument as in Li (2006, pp. 35), we have

$$s_1(u) - h\dot{f}(u)\mu_2 = \mathcal{O}_p\Big(\sqrt{\frac{\log(n)}{nh}}+o(h)\Big) = \mathbf{o}_p(h). \tag{A.4}$$

Next, by Lemma 2 of Yao and Tong (1998) we know that:

$$\sup_{u\in G}|s_2(u) - f(u)\mu_2| = o_p(1). \tag{A.5}$$

By Taylor's expansion, we then have

$$\begin{aligned} & \hat{m}_j(u) - m_j(u) \\ = \; & \frac{1}{nhf(u)}\Big\{1+\mathcal{O}_p\Big(\sqrt{\frac{\log(n)}{nh}}+h^2\Big)\Big\}\sum_{i=1}^{n} k_i(u)\Big(X_{ij}-m_j(U_i)\Big) \\ & +\frac{1}{f(u)}\Big\{1+\mathcal{O}_p\Big(\sqrt{\frac{\log(n)}{nh}}+h^2\Big)\Big\}\Big\{h\dot{m}_j(u)s_1(u)+\frac{h^2\ddot{m}_j(u)}{2}s_2(u)+\mathbf{o}_p(h^2)\Big\}. \end{aligned}$$

Then it follows by (A.4) and (A.5) that

$$\hat{m}_j(u) - m_j(u) = \frac{1}{nhf(u)} \sum_{i=1}^{n} k_i(u) \Big\{ X_{ij} - m_j(U_i) \Big\}$$

$$+ \frac{h^2 \mu_2}{2} \Big[\frac{2\dot{f}(u)}{f(u)} \dot{m}_j(u) + \ddot{m}_j(u) \Big] + \mathcal{O}_p\{R_{n,1}^{(j)}(u)\}.$$

This proves (A.1). Furthermore, (A.2) follows by similar argument. This completes the proof.

Appendix B. Proof of theorem 1

Let $k_i(u) = K\{(U_i - u)/h\}$ and $\{X_{ij_1} - m_{j_1}(U_i)\}\{X_{ij_2} - m_{j_2}(U_i)\} = \sigma_{j_1j_2}(U_i) + \varepsilon_{j_1j_2}(i)$, where $\varepsilon_{j_1j_2}(i)$ satisfies with $E(\varepsilon_{j_1j_2}(i)|U_i) = 0$ and $\text{var}(\varepsilon_{j_1j_2}(i)\big|U_i) \triangleq \omega_{j_1j_2}(U_i)$. By the Lemma in Appendix A, we have $\hat{\sigma}_{j_1j_2}(u) - \sigma_{j_1j_2}(u) = I_1 + I_2 - I_3 - I_4 + I_5 + \mathcal{O}_p(h)(|I_1^{'} + I_2 - I_3 - I_4 + I_5|)$ where

$$I_1^{'} = \{nhf(u)\}^{-1} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \Big\{ \sigma_{j_1j_2}(U_i) - \sigma_{j_1j_2}(u) - (U_i - u)\dot{\sigma}_{j_1j_2}(u) \Big\},$$

$$I_1 = I_1^{'} + h^2 \dot{\sigma}_{j_1j_2}(u) \mu_2 \dot{f}(u) f^{-1}(u)$$

$$I_2 = \{nhf(u)\}^{-1} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \varepsilon_{j_1j_2}(i)$$

$$I_3 = \{nhf(u)\}^{-1} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \Big\{ \hat{m}_{j_1}(U_i) - m_{j_1}(U_i) \Big\} \Big\{ X_{ij_2} - m_{j_2}(U_i) \Big\}$$

$$I_4 = \{nhf(u)\}^{-1} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \Big\{ X_{ij_1} - m_{j_1}(U_i) \Big\} \Big\{ \hat{m}_{j_2}(U_i) - m_{j_2}(U_i) \Big\}$$

$$I_5 = \{nhf(u)\}^{-1} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) \Big\{ \hat{m}_{j_1}(U_i) - m_{j_1}(U_i) \Big\} \Big\{ \hat{m}_{j_2}(U_i) - m_{j_2}(U_i) \Big\}.$$

It is easy to see that the theorem follows directly from the following statements:

$$I_1^{'} = \frac{h^2}{2} \mu_2 \ddot{\sigma}_{j_1j_2}(u) + o_p(h^2) \tag{A.6}$$

$$I_1 = \frac{h^2}{2} \mu_2 \left\{ \ddot{\sigma}_{j_1j_2}(u) + 2\dot{\sigma}_{j_1j_2}(u) \frac{\dot{f}(u)}{f(u)} \right\} + o_p(h^2) \tag{A.7}$$

$$\sqrt{nh}I_2 \to_d N\left(0, \frac{\nu_0\omega_{j_1j_2}(u)}{f(u)}\right) \quad (A.8)$$

$$I_3 = o_p(h^2),\ I_4 = o_p(h^2),\ \text{and } I_5 = o_p(h^2). \quad (A.9)$$

Next, we will verify (A.6) to (A.9) in a one-by-one manner. Note that (A.6) and (A.7) can be verified by standard Taylor expansion argument. By conditions (C2) and (C6), one can check that $\{nhf(u)\}^{-1}k_i(u)\varepsilon_{j_1j_2}(i)$ satisfies the requirement of Liapunov's Central Limit Theorem (Li, 2006). Furthermore, note that $EI_2 = 0$ and by standard arguments, it follows that

$$\begin{aligned} \text{var}(I_2) &= EI_2^2 = \{nh^2f^2(u)\}^{-1}E\Big[K^2\Big(\frac{U_i-u}{h}\Big)\varepsilon^2_{j_1j_2}(i)\Big] \\ &= \{nhf(u)\}^{-1}\omega_{j_1j_2}(u)\nu_0 + O(h/n) \end{aligned} \quad (A.10)$$

where $\nu_0 = \int K^2(v)dv$. Thus, we have $\sqrt{nh}I_2 \to_d N\{0, \nu_0\omega_{j_1j_2}(u)/f(u)\}$. This proves the statement (A.8).

To prove $I_3 = o_p(h^2)$, we write $X_{ij} = m_j(U_i) + \sigma_j(U_i)\varepsilon_j(i)$ with $\sigma_j(U_i) = \sqrt{\sigma_{jj}(U_i)} > 0$, $E\{\varepsilon_j(i)|U_i\} = 0$, and $\text{var}\{\varepsilon_j(i)|U_i\} = 1$. We then have $I_3 = I_{31} + I_{32} + I_{33}$ with

$$\begin{aligned} I_{31} &= \frac{1}{n^2h^2f(u)}\sum_{i,j=1}^n K\Big(\frac{U_i-u}{h}\Big)K\Big(\frac{u_j-U_i}{h}\Big)f^{-1}(U_i)\sigma_{j_1}(u_j)\sigma_{j_2}(U_i)\varepsilon_{j_1}(j)\varepsilon_{j_2}(i) \\ &= \frac{1}{n^2h^2f(u)}\sum_{i\neq j}\varphi_{ij}^{j_1j_2} + O_p(\frac{1}{nh}) \end{aligned}$$

$$I_{32} = \frac{h\mu_2}{nf(u)}\sum_{i=1}^n K\Big(\frac{U_i-u}{h}\Big)\Big(\dot m_{j1}(U_i)\frac{\dot f(U_i)}{f(U_i)} + \frac{\ddot m_{j_1}(U_i)}{2}\Big)\sigma_{j_2}(U_i)\varepsilon_{j_2}(i) = o_p(h^2)$$

$$|I_{33}| \le \frac{\mathcal{O}_p(1)}{n^2hf(u)}\Bigg|\sum_{i,j=1}^n K\Big(\frac{U_i-u}{h}\Big)K\Big(\frac{u_j-U_i}{h}\Big)\left\{\frac{\sigma_{j_1}(u_j)\sigma_{j_2}(U_i)}{f(U_i)}\right\} \times|\varepsilon_{j_1}(j)|\varepsilon_{j_2}(i)\Bigg| + \mathbf{o}_p(h^2),$$

where $\varphi_{ij}^{j_1j_2} = K\{(U_i-u)/h\}K\{(u_j-U_i)/h\}f^{-1}(U_i)\sigma_{j_1}(u_j)\sigma_{j_2}(U_i)\varepsilon_{j_1}(j)\varepsilon_{j_2}(i)$.

Thus, the statement $I_3 = o_p(h^2)$ follows if we are able to show $I_{3k} =$

$o_p(h^2)$ for $j = 1, 2, 3$. For such a purpose, we first consider I_{31} and its related quantity $E(\sum_{i \neq j} \varphi_{ij}^{j_1 j_2})^2 = \sum E(\varphi_{ij}^{j_1 j_2} \varphi_{kl}^{j_1 j_2})$. Note that if $\{i, j\} \neq \{k, l\}$, then $E\varphi_{ij}^{j_1 j_2} \varphi_{kl}^{j_1 j_2} = 0$. Thus, we only need to concern about the situation $(k, l) = (i, j)$ and $(k, l) = (j, i)$. Without loss of generality, we will only consider the situation $(k, l) = (i, j)$. Note that if $i \neq j$, we have $E\varphi_{ij}^{j_1 j_2} = E_U E_{X|U} \varphi_{ij}^{j_1 j_2} = 0$ and

$$
\begin{aligned}
& E(\varphi_{ij}^{j_1 j_2})^2 \\
= \;& E_U \Big\{ K^2\Big(\frac{u_j - U_i}{h}\Big) K^2\Big(\frac{U_i - u}{h}\Big) \sigma_{j_1}^2(u_j) \sigma_{j_2}^2(U_i) f^{-2}(U_i) E_{X|U} \varepsilon_{j_1}^2(j) \varepsilon_{j_2}^2(i) \Big\} \\
= \;& E_U K^2\Big(\frac{u_j - U_i}{h}\Big) K^2\Big(\frac{U_i - u}{h}\Big) \sigma_{j_1}^2(u_j) \sigma_{j_2}^2(U_i) f^{-2}(U_i) \\
= \;& \int K^2(\frac{U_i - u}{h}) f^{-1}(U_i) \sigma_{j_2}^2(U_i) \Big(h\nu_0 \sigma_{j_1}^2(U_i) f(U_i) + O(h^3) \Big) dU_i \\
= \;& h^2 (\nu_0)^2 \sigma_{j_1}^2(u) \sigma_{j_2}^2(u) + O(h^4). \qquad \text{(A.11)}
\end{aligned}
$$

Since (A.11) is independent of i,j, we have $\max_{1 \le i \neq j \le n} |E(\varphi_{ij}^{j_1 j_2})^2| < O(h^2)$.

As a consequence, we have $E(\sum_{i \neq j} \varphi_{ij}^{j_1 j_2})^2 \le O\{n(n-1)h^2\}$. Then, for an arbitrary $\varepsilon > 0$ and some $\varepsilon_0 > 0$, by Markov's Inequality, we have

$$
\begin{aligned}
& P\Big(n^{-1} (h^2)^{-(\frac{1}{1+\delta} - \varepsilon_0)/2} | \sum_{i \neq j} \varphi_{ij}^{j_1 j_2} | > \varepsilon \Big) \\
\le \;& \frac{(h^2)^{\varepsilon_0}}{n^2 \varepsilon^2} E\Big\{ (h^2)^{-\frac{1}{2(1+\delta)}} \sum_{i \neq j} \varphi_{ij}^{j_1 j_2} \Big\}^2 = o((h^2)^{\varepsilon_0}) \qquad \text{(A.12)}
\end{aligned}
$$

Thus, we have $I_{31} = o_p(n^{-1} h^{-\frac{1+2\delta}{(1+\delta)} - \varepsilon_0}) + O_p(n^{-1} h^{-1})$. We know then $I_{31} = o_p(h^2)$ by choosing $\varepsilon_0 < (1 + \delta)^{-1}$. Similarly, we can also show that $I_{32} = o_p(h^2)$, $I_{33} = o_p(h^2)$, $I_4 = o_p(h^2)$ and $I_5 = o_p(h^2)$. This completes the proof.

Appendix C. Proof of theorem 2

We first prove (2.8). To save space, we denote $\Sigma(u)$, $\hat{\Sigma}(u)$ by Σ and $\hat{\Sigma}$ for short, respectively. Next, we can rewrite the quadratic loss of the $\hat{\Sigma}$ as

$$\begin{aligned}
\Delta_2 &= Etr\Big\{(\hat{\Sigma}-\Sigma)\Sigma^{-1}(\hat{\Sigma}-\Sigma)\Sigma^{-1}\Big\} \\
&= Etr\Big\{\big(\hat{\Sigma}\Sigma^{-1}\big)^2\Big\} - 2Etr\big\{\hat{\Sigma}\Sigma^{-1}\big\} + p \\
&= Etr\Big\{\big(\Sigma^{-1/2}\hat{\Sigma}\Sigma^{-1/2}\big)^2\Big\} - 2Etr\big\{\Sigma^{-1/2}\hat{\Sigma}\Sigma^{-1/2}\big\} + p \\
&= E\sum_{j=1}^{p}(\hat{d}_j-1)^2, \qquad \text{(A.13)}
\end{aligned}$$

where $\hat{d}_j$ is the jth largest eigenvalue of the matrix $\Sigma^{-1/2}\hat{\Sigma}\Sigma^{-1/2}$. Because $\hat{\Sigma}\to_p \Sigma$, we know that $\hat{d}_j \to_p 1$. Furthermore, note that

$$\Delta_1(u) = E\Big[tr\Big\{\Sigma^{-1}\hat{\Sigma}\Big\} - \log\Big|\Sigma^{-1}\hat{\Sigma}\Big|\Big] - p = E\sum_{i=1}^{p}\Big[\hat{d}_i - \log(\hat{d}_i) - 1\Big].$$

By Taylor's expansion, we have $\log(d_i) = \log(1+d_i-1) = d_i - 1 - 0.5(d_i-1)^2 + o\{(d_i-1)^2\}$. We can then write $\Delta_1(u)$ as

$$E\sum_{j=1}^{p}\frac{(\hat{d}_j-1)^2}{2}\{1+o_p(1)\} = \frac{1}{2}E\sum_{j=1}^{p}(\hat{d}_j-1)^2\{1+o_p(1)\}. \qquad \text{(A.14)}$$

Then, by (A.13)and (A.14), we know that the theorem statement is correct. This completes the proof of (2.8).

We next prove (2.9). Let us introduce some notation. Define $\dot{A}(u) = \{\dot{a}_{j_1j_2}(u)\}$ for an arbitrary varying matrix $A(u) = \{a_{j_1j_2}(u)\}$. By (A.1) and (A.2) in Appendix B, we have

$$\begin{aligned}
\hat{m}(u) - m(u) &= \frac{1}{nhf(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Big\{X_i - m(U_i)\Big\} \\
&\quad + \frac{h^2\mu_2}{2}\Big[2\frac{\dot{f}(u)}{f(u)}\dot{m}(u) + \ddot{m}(u)\Big] + \mathcal{O}_p\{R^*_{n,1}(u)\} \qquad \text{(A.15)}
\end{aligned}$$

$$\hat{\Sigma}(u) - \Sigma(u) = \frac{1}{nhf(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Bigg[\Big\{X_i - \hat{m}(U_i)\Big\}\Big\{X_i - \hat{m}(U_i)\Big\}^\top$$

$$-\Sigma(u)-(U_i-u)\dot{\Sigma}(u)\Big]+h^2\mu_2\frac{\dot{f}(u)}{f(u)}\dot{\Sigma}(u)+\mathcal{O}_p\{R^*_{n,2}(u)\} \qquad \text{(A.16)}$$

where

$$R^*_{n,1}(u) = \frac{1}{nf(u)}\left\{\left|\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Big\{X_i-m(U_i)\Big\}\right|\right\}+o(h^2)$$

$$R^*_{n,2}(u) = \frac{1}{nf(u)}\left\{\left|\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Big[\Big\{X_i-\hat{m}(U_i)\Big\}\Big\{X_i-\hat{m}(U_i)\Big\}^\top -\Sigma(u)-(U_i-u)\dot{\Sigma}(u)\Big]\right|\right\}+o(h^2).$$

Similar to the proof of Theorem 1, we write $\{X_i-m(U_i)\}\{X_i-m(U_i)\}^\top = \Sigma(U_i)+\mathcal{E}(i)$, where $\mathcal{E}(i)$ satisfies that $E\Big(\mathcal{E}(i)|U_i\Big)=0$ and

$$\mathrm{var}\Big(\mathcal{E}(i)|U_i\Big) = E_{X|U}\Big[\Big\{X_i-m(U_i)\Big\}\Big\{X_i-m(U_i)\Big\}^\top-\Sigma(U_i)\Big]^2$$

$$= E_{X|U}\Big[\Big\{X_i-m(U_i)\Big\}\Big\{X_i-m(U_i)\Big\}^\top\Big\{X_i-m(U_i)\Big\}\Big\{X_i-m(U_i)\Big\}^\top\Big]-\Sigma(U_i)\Sigma(U_i).$$

Similar to Theorem 1, we can write

$$\hat{\Sigma}(u)-\Sigma(u) = J_1+J_2-J_3-J_4+J_5$$

$$+O_p(h)\Big\{\Big|J_1'+J_2-J_3-J_4+J_5\Big|\Big\}, \qquad \text{(A.17)}$$

where

$$J_1' = \frac{1}{nhf(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Big\{\Sigma(U_i)-\Sigma(u)-(U_i-u)\dot{\Sigma}(u)\Big\}$$

$$J_1 = J_1'+h^2\mu_2\frac{\dot{f}(u)}{f(u)}\dot{\Sigma}(u)$$

$$J_2 = \frac{1}{nhf(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\mathcal{E}(i)$$

$$J_3 = \frac{1}{nhf(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h}\Big)\Big\{\hat{m}(U_i)-m(U_i)\Big\}\Big\{X_i-m(U_i)\Big\}^\top$$

$$J_4 = \frac{1}{nhf(u)}\sum_{i=1}^{n} K\Big(\frac{U_i-u}{h}\Big)\Big\{X_i - m(U_i)\Big\}\Big\{\hat{m}(U_i)-m(U_i)\Big\}^\top$$
$$J_5 = \frac{1}{nhf(u)}\sum_{i=1}^{n} K\Big(\frac{U_i-u}{h}\Big)\Big\{\hat{m}(U_i) - m(U_i)\Big\}\Big\{\hat{m}(U_i)-m(U_i)\Big\}^\top.$$

Similar to the proof of Theorem 1, we can also show that $J_1^{'} = 0.5h^2\mu_2\ddot{\Sigma}(u) + o_p(h^2)$. $J_1 = 0.5h^2\mu_2\{\ddot{\Sigma}(u) + 2\dot{f}(u)f^{-1}(u)\dot{\Sigma}(u)\} + o_p(h^2)$, $J_3 = o_p(h^2)$, $J_4 = o_p(h^2)$, and also $J_5 = o_p(h^2)$. Recalling the definition of the quadratic loss:

$$\Delta_2(u) = Etr\Big\{\Big(\hat{\Sigma}(u)\Sigma^{-1}(u) - I\Big)^2\Big\}$$

$$= trE\Big(\hat{\Sigma}(u) - \Sigma(u)\Big)\Sigma^{-1}(u)\Big(\hat{\Sigma}(u) - \Sigma(u)\Big)\Sigma^{-1}(u). \qquad \text{(A.18)}$$

From (A.17), we then have

$$\begin{aligned}
&\Big\{\hat{\Sigma}(u) - \Sigma(u)\Big\}\Sigma^{-1}(u)\Big\{\hat{\Sigma}(u) - \Sigma(u)\Big\}\Sigma^{-1}(u)\\
= &\Bigg[\frac{h^2}{2}\mu_2\Big\{\ddot{\Sigma}(u)\Sigma^{-1}(u) + 2\frac{\dot{f}(u)}{f(u)}\dot{\Sigma}(u)\Sigma^{-1}(u)\Big\} + J_2\Sigma^{-1}(u) + o_p(h^2)\\
&+O_p(h)\Big|\Big(J_1^{'} + J_2 - J_3 - J_4 + J_5\Big)\Sigma^{-1}(u)\Big|\Bigg]^2 \qquad \text{(A.19)}
\end{aligned}$$

Because $J_1^{'}$, J_3, J_4 and J_5 are all of the $O_p(h^2)$ order, (A.19) equals

$$\Bigg[\frac{h^2}{2}\mu_2\Big\{\ddot{\Sigma}(u)\Sigma^{-1}(u)+2\frac{\dot{f}(u)}{f(u)}\dot{\Sigma}(u)\Sigma^{-1}(u)\Big\}+J_2\Sigma^{-1}(u)+o_p(h^2)+O_p(h)|J_2\Sigma^{-1}(u)|\Bigg]^2.$$

According to the point-wise results presented in Appendix C, we know that $J_2 = O_p((nh)^{-1/2})$. Consequently, the above quantity can be further written as

$$\Bigg[\frac{h^2}{2}\mu_2\Big\{\ddot{\Sigma}(u)\Sigma^{-1}(u) + 2\frac{\dot{f}(u)}{f(u)}\dot{\Sigma}(u)\Sigma^{-1}(u)\Big\} + J_2\Sigma^{-1}(u) + o_p(h^2 + (nh)^{-1/2})\Bigg]^2.$$

Consequently, we have

$$\Delta_2(u) = \frac{h^4}{4}\Big(\mu_2\Big)^2 C_1(u) + trE\Big\{J_2\Sigma^{-1}(u)J_2\Sigma^{-1}(u)\Big\} + o\Big(h^4 + (nh)^{-1}\Big) \quad (A.20)$$

where $C_1(u) = tr\{[\ddot{\Sigma}(u)\Sigma^{-1}(u) + 2\dot{f}(u)/f(u)\dot{\Sigma}(u)\Sigma^{-1}(u)]^2\}$. Next, note that

$$J_2\Sigma^{-1}(u)J_2\Sigma^{-1}(u) = \frac{1}{n^2h^2f^2(u)}\Bigg[\sum_{i=1}^{n} K^2\Big(\frac{U_i - u}{h}\Big)\mathcal{E}(i)\Sigma^{-1}(u)\mathcal{E}(i)\Sigma^{-1}(u)$$

$$+\sum_{i\neq j} K\Big(\frac{U_i - u}{h}\Big)K\Big(\frac{u_j - u}{h}\Big)\mathcal{E}(i)\Sigma^{-1}(u)\mathcal{E}(j)\Sigma^{-1}(u)\Bigg]. \quad (A.21)$$

It is easy to verify that whenever $i \neq j$, we must have

$$E_{X|U}\Big\{K\Big(\frac{U_i - u}{h}\Big)K\Big(\frac{u_j - u}{h}\Big)\mathcal{E}(i)\Sigma^{-1}(u)\mathcal{E}(j)\Sigma^{-1}(u)\Big\} = 0. \quad (A.22)$$

In addition to that, we also have

$$\begin{aligned} & trE_{X|U}\Big\{\mathcal{E}(i)\Sigma^{-1}(u)\mathcal{E}(i)\Sigma^{-1}(u)\Big\} \\ = \;& trE_{X|U}\Big\{\Big[\Big(X_i - m(U_i)\Big)\Big(X_i - m(U_i)\Big)^\top\Sigma^{-1}(u) - \Sigma(U_i)\Sigma^{-1}(u)\Big]^2\Big\} \\ = \;& trE_{X|U}\Big\{\Big(X_i - m(U_i)\Big)\Big(X_i - m(U_i)\Big)^\top\Sigma^{-1}(u) \\ & \times\Big(X_i - m(U_i)\Big)\Big(X_i - m(U_i)\Big)^\top\Sigma^{-1}(u)\Big\} \\ & -tr\Big\{\Sigma(U_i)\Sigma^{-1}(u)\Sigma(U_i)\Sigma^{-1}(u)\Big\} \triangleq \Phi(U_i, u) - \Psi(U_i, u), \end{aligned} \quad (A.23)$$

where

$$\begin{aligned} \Phi(U_i, u) &= E_{X|U}\Big\{\Big[\Big(X_i - m(U_i)\Big)^\top\Sigma^{-1}(u)\Big(X_i - m(U_i)\Big)\Big]^2\Big\} \\ \Psi(U_i, u) &= tr\Big\{\Sigma(U_i)\Sigma^{-1}(u)\Sigma(U_i)\Sigma^{-1}(u)\Big\}. \end{aligned}$$

In particular, we have $\Psi(u, u) = p$. By (A.21)-(A.23) and conditions (C1), (C4) in Appendix A, we have

$$trE\Big\{J_2\Sigma^{-1}(u)J_2\Sigma^{-1}(u)\Big\} \tag{A.24}$$

$$= \frac{1}{n^2h^2f^2(u)}\Big[n\int_{-\infty}^{\infty}K^2\Big(\frac{U_i-u}{h}\Big)\Big[\Phi(U_i,u)-\Psi(U_i,u)\Big]f(U_i)dU_i\Big]$$

$$= \frac{1}{nhf(u)}\nu_0\Big[\Phi(u,u)-p\Big]+o\Big((nh)^{-1}\Big). \tag{A.25}$$

Consequently, by (A.20)-(A.25), we have:

$$\Delta_2(u) = \frac{h^4}{4}\Big(\mu_2\Big)^2C_1(u)+\frac{1}{nhf(u)}\nu_0\Big[\Phi(u,u)-p\Big]+o\Big(h^4+(nh)^{-1}\Big).$$

This completes the proof of Theorem 2.

Appendix D. Simulation Studies

To corroborate our theoretical findings, two simulation examples are presented in this appendix. For these two examples, we consider the following 5-dimensional nonparametric covariance model:

Case I: The regression function $m(u) = (6u, 10\cos(u), 25\sin(2u), 20\exp(u), 30u-10)^\top \in \mathbb{R}^5$ and $\Sigma(u) = T(u)T^\top(u)$ with

$$T(u) = \begin{pmatrix} 2\cos(u) & 0 & 0 & 0 & 0 \\ 3/2\sin(u) & 4\cos(u) & 0 & 0 & 0 \\ 2\sin(u) & 5/2\sin(u) & 6\cos(u) & 0 & 0 \\ 5/2\sin(u) & 3\sin(u) & 7/2\sin(u) & 8\cos(u) & 0 \\ 3\sin(u) & 7/2\sin(u) & 4\sin(u) & 9/2\sin(u) & 10\cos(u) \end{pmatrix}.$$

Case II: $m(u) = (5u, 4u+1, 7u, 3u, 6\exp(u))^\top \in \mathbb{R}^5$ and

$$T(u) = \begin{pmatrix} \exp(u) & 0 & 0 & 0 & 0 \\ -u & \exp(-u) & 0 & 0 & 0 \\ u & 0 & 1/(1+u) & 0 & 0 \\ -u & 2u-1 & -1.5u & \cos(2\pi u)+1 & 0 \\ u & -2u & 1.5u & \sin(u) & \log(u+5) \end{pmatrix}.$$

In each simulation iteration, we firstly simulate U from a uniform distribution on $[0, 1]$ and then simulate $X \in \mathbb{R}^5$ from a 5-dimensional normal distribution with mean $m(u)$ and covariance $\Sigma(u)$. We then apply the CV method to determine the optimal bandwidths needed for the estimation. With the estimated optimal bandwidth, we then estimate the covariance matrix $\Sigma(u)$ as (2.3). The estimation error of such an estimate is then evaluated by both the Stein and quadratic losses. Specifically, for each simulation data set, the estimate $\hat{\Sigma}(u)$ is evaluated at $u = U_i$ and calculate

$$\text{median}(\Delta_j) = \text{median}\{\Delta_j(U_i) : i = 1, \cdots, n\}$$

for $j = 1$ and 2. Here we summarize the simulation results using the sample median instead of the sample mean to minimize the impact of the boundary effect of the kernel estimate on our summary. Figures 3.2 and 3.3 depict the boxplot of 200 median(Δ_j)'s over 200 simulations. It can be seen from Figures 3.2 and 3.3 that as sample size increases from $n = 100$ till $n = 800$, the estimation error of our estimate steadily decreases, which numerically confirms the consistency of our estimate (2.3).

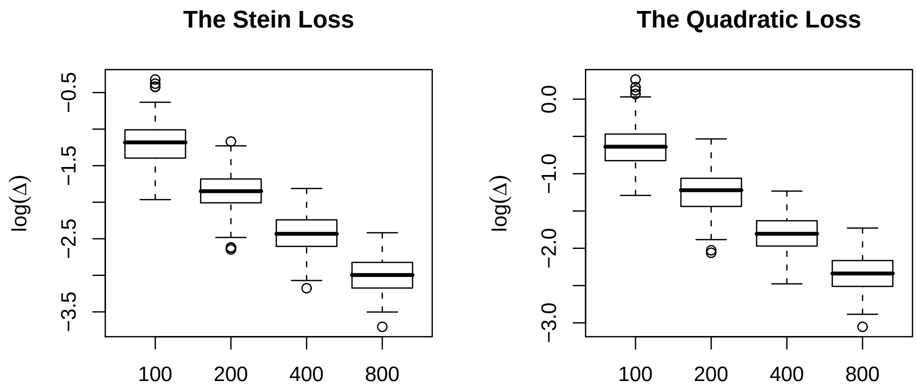

Figure 3.2: Box-plot of Estimation Errors for Case I

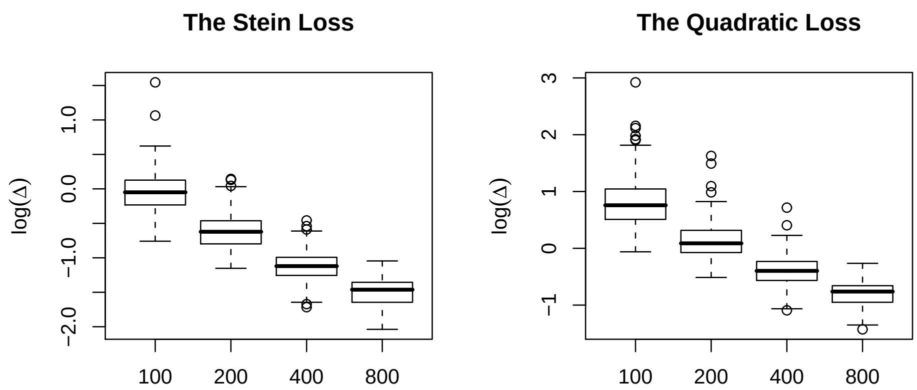

Figure 3.3: Box-plot of Estimation Errors for Case II.