Local Linear Estimation of Covariance Matrices via Cholesky Decomposition

Ziqi Chen and Chenlei Leng

Central South University and University of Warwick

Supplementary Material

Lemma 1. *Under conditions (a)–(d), we have*

$$\hat{m}(u) - m(u) = \frac{1}{2}\mu_2 h^2 m^{''}(u) + \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i - u)\{Y_i - m(U_i)\} + o_p(c_n),$$

which holds uniformly in u.

Lemma 2. *If conditions (a)–(d) hold, we have*

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} K_h(U_i - u)K_h(U_j - U_i)\frac{r_{ik}r_{jq}}{f(U_i)} = o_p(c_n),$$

which holds uniformly in u, for $k, q = 1, \cdots, p$.

Proof: We have

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} K_h(U_i - u)K_h(U_j - U_i)\frac{r_{ik}r_{jq}}{f(U_i)} \\
&\quad = \frac{n(n-1)}{n^2}\frac{1}{n(n-1)}\sum_{i\neq j} K_h(U_i - u)K_h(U_j - U_i)\frac{r_{ik}r_{jq}}{f(U_i)} + \frac{K(0)}{nh}\frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u)\frac{r_{ik}r_{iq}}{f(U_i)}.
\end{aligned}$$

Note $\mathcal{F} = \{\frac{1}{n(n-1)}\sum_{i\neq j} K(\frac{U_i-u}{h})K(\frac{U_j-U_i}{h})\frac{r_{ik}r_{jq}}{f(U_i)} : u \in \Omega\}$ is a U-process of order 2 indexed by Ω. We know easily that $K(\frac{U_i-u}{h})K(\frac{U_j-U_i}{h})\frac{r_{ik}r_{jq}}{f(U_i)}$ is P-degenerate for each $u \in \Omega$. From

Lemma 2.13 in Pakes and Pollard (1989), $\mathcal{F}$ is Euclidean for some envelope F satisfying $E(F^2) < \infty$. Thus, by Corollary 4 in Sherman (1994), the first term of the right-hand side of the above equation is $O_p(1/(nh^2))$ uniformly in u. It can be easily seen that the second term of the right-hand side of the above equation is $O_p(1/(nh))$ uniformly in u. Thus, the result follows immediately.

Proof of Theorem 1: Firstly, we have

$$\begin{aligned}
\hat{\Phi}_j(u) - \Phi_j(u) &= (I_{(j-1)}, \mathbf{0}_{(j-1)})\Big\{\sum_{i=1}^n K_h(U_i - u)X_i^{(j)T}X_i^{(j)}\Big\}^{-1} \\
&\quad \times \sum_{i=1}^n K_h(U_i - u)X_i^{(j)T}\{\hat{r}_{ij} - \sum_{l=1}^{j-1}\hat{r}_{il}\phi_{jl}(u) - \sum_{l=1}^{j-1}\hat{r}_{il}\phi_{jl}^{'}(u)(U_i - u)\}. \qquad (1)
\end{aligned}$$

By Lemma 1 and Lemma 2, we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n K_h(U_i - u)\hat{r}_{ik}\hat{r}_{iq} &= \frac{1}{n}\sum_{i=1}^n K_h(U_i - u)[r_{ik} - \{\hat{m}_k(U_i) - m_k(U_i)\}][r_{iq} - \{\hat{m}_q(U_i) - m_q(U_i)\}] \\
&= \frac{1}{n}\sum_{i=1}^n K_h(U_i - u)r_{ik}r_{iq} - \frac{\mu_2 h^2}{2n}\sum_{i=1}^n K_h(U_i - u)[m_k^{''}(U_i)r_{iq} + m_q^{''}(U_i)r_{ik}] \\
&\quad -\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n K_h(U_i - u)K_h(U_j - U_i)\frac{r_{ik}r_{jq}}{f(U_i)} \\
&\quad -\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n K_h(U_i - u)K_h(U_j - U_i)\frac{r_{jk}r_{iq}}{f(U_i)} + o_p(c_n) \\
&= \sigma_{kq}(u)f(u) + o_p(1), \qquad (2)
\end{aligned}$$

which holds uniformly in u, where $\sigma_{kq}(u)$ is the (k,q)-th entry of $\Sigma(u)$. And we have that the following two equations hold uniformly in u:

$$\frac{1}{n}\sum_{i=1}^n K_h(U_i - u)(U_i - u)\hat{r}_{ik}\hat{r}_{iq} = O_p(h^2), \qquad (3)$$

$$\frac{1}{n}\sum_{i=1}^n K_h(U_i - u)(U_i - u)^2\hat{r}_{ik}\hat{r}_{iq} = O_p(h^2) \qquad (4)$$

We can see easily that

$$
\begin{aligned}
&\hat{r}_{ij} - \sum_{l=1}^{j-1} \hat{r}_{il}\phi_{jl}(u) - \sum_{l=1}^{j-1} \hat{r}_{il}\phi_{jl}^{'}(u)(U_i - u) \\
&= \epsilon_{ij} + \frac{1}{2}[(U_i - u)^2 r_{i1}, \cdots, (U_i - u)^2 r_{i(j-1)}]\Phi_j^{''}(u) + O_p((U_i - u)^3) - (\hat{m}_j(U_i) - m_j(U_i)) \\
&\quad - \sum_{l=1}^{j-1} \phi_{jl}(u)(\hat{m}_l(U_i) - m_l(U_i)) - (U_i - u)\sum_{l=1}^{j-1} \phi_{jl}^{'}(u)(\hat{m}_l(U_i) - m_l(U_i)), \qquad (5)
\end{aligned}
$$

and

$$
X_i^{(j)T} = \begin{pmatrix} r_{i1} - (\hat{m}_1(U_i) - m_1(U_i)) \\ \vdots \\ r_{i(j-1)} - (\hat{m}_{j-1}(U_i) - m_{j-1}(U_i)) \\ (U_i - u)\{r_{i1} - (\hat{m}_1(U_i) - m_1(U_i))\} \\ \vdots \\ (U_i - u)\{r_{i(j-1)} - (\hat{m}_{j-1}(U_i) - m_{j-1}(U_i))\} \end{pmatrix}. \qquad (6)
$$

By (5) and (6), using Lemma 1 and Lemma 2, we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u) X_i^{(j)T}\{\hat{r}_{ij} - \sum_{l=1}^{j-1} \hat{r}_{il}\phi_{jl}(u) - \sum_{l=1}^{j-1} \hat{r}_{il}\phi_{jl}^{'}(u)(U_i - u)\} \\
&= \frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u) \begin{pmatrix} r_{i1} \\ \vdots \\ r_{i(j-1)} \\ (U_i - u)r_{i1} \\ \vdots \\ (U_i - u)r_{i(j-1)} \end{pmatrix} \{\epsilon_{ij} + \frac{1}{2}[(U_i - u)^2 r_{i1}, \cdots, (U_i - u)^2 r_{i(j-1)}]\Phi_j^{''}(u)\} \\
&\quad + o_p(c_n), \qquad (7)
\end{aligned}
$$

which holds uniformly in u.

By Equations (1), (2), (3), (4) and (7) and the formula of the inverse of a block matrix, we obtain the result through simple calculations.

Proof of Theorem 2: Note that $\nu(u) = \log(\sigma_j^2(u))$. Using the same arguments in the proof of Theorem 1, we have that

$$\begin{aligned}
\hat{\nu}(u)-\nu(u) &= \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\hat{\epsilon}_{ij}^2\exp\{-\nu(u)-\nu^{'}(u)(U_i-u)\}-1\Big\} \\
&= \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\hat{\epsilon}_{ij}^2\exp\{-\nu(U_i)+\frac{1}{2}\nu^{''}(u)(U_i-u)^2+O_p((U_i-u)^3)\}-1\Big\} \\
&= \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\hat{\epsilon}_{ij}^2\exp\{-\nu(U_i)\}\{1+\frac{1}{2}\nu^{''}(u)(U_i-u)^2\}-1\Big\}+o_p(h^2) \\
&= \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\hat{\epsilon}_{ij}^2\exp\{-\nu(U_i)\}-1\Big\} \\
&\quad + \frac{\nu^{''}(u)}{2nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\hat{\epsilon}_{ij}^2\exp\{-\nu(U_i)\}(U_i-u)^2+o_p(h^2) \\
&= \frac{1}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\epsilon_{ij}^2\exp\{-\nu(U_i)\}-1\Big\} \\
&\quad + \frac{\nu^{''}(u)}{2nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\epsilon_{ij}^2\exp\{-\nu(U_i)\}(U_i-u)^2+o_p(c_n),
\end{aligned}$$

which holds uniformly in u. By Taylor's expansion, it can be easily seen that

$$\begin{aligned}
\hat{\sigma}_j^2(u)-\sigma_j^2(u) &= \exp(\hat{\nu}(u))-\exp(\nu(u)) \\
&= \frac{[\log\{\sigma_j^2(u)\}]^{''}\sigma_j^2(u)}{2nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\frac{\epsilon_{ij}^2}{\sigma_j^2(U_i)}(U_i-u)^2 \\
&\quad + \frac{\sigma_j^2(u)}{nf(u)}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\frac{\epsilon_{ij}^2}{\sigma_j^2(U_i)}-1\Big\}+o_p(c_n) \\
&= \frac{1}{2}\mu_2[\log\{\sigma_j^2(u)\}]^{''}\sigma_j^2(u)h^2+f^{-1}(u)\frac{1}{n}\sum_{i=1}^{n} K_h(U_i-u)\Big\{\epsilon_{ij}^2-\sigma_j^2(U_i)\Big\}+o_p(c_n)
\end{aligned}$$

holds uniformly in u.

Proof of Theorem 3: Let $P_B(u) := -2^{-1}\mu_2 h^2 P^{''}(u)$ and $P_V(u)$ a strictly lower triangular matrix with $0'$s on its diagonal and with the lower triangular vector of its j-th row being

$$f^{-1}(u)\{\Sigma(u)_{(j-1,j-1)}\}^{-1}\frac{1}{n}\sum_{i=1}^{n} K_h(U_i-u)\Big(r_{i1},\cdots,r_{i(j-1)}\Big)\epsilon_{ij}.$$

Let $D_B(u)$ and $D_V(u)$ be both diagonal matrices with the $j-$th diagonal entry being $\sigma_{jB}^2 :=$

$2^{-1}\mu_2 h^2[\log\{\sigma_j^2(u)\}]''\sigma_j^2(u)$ and $\sigma_{jV}^2 := (nf(u))^{-1}\sum_{i=1}^n K_h(U_i-u)(\epsilon_{ij}^2-\sigma_j^2(U_i))$, respectively. We recall that $\Sigma(u) = P^{-1}(u)D(u)P^{-1}(u)^T$ and

$$\begin{aligned}\hat{\Sigma}^{-1}(u) &= \hat{P}(u)^T\hat{D}^{-1}(u)\hat{P}(u)\\ &= \{P(u)+P_B(u)+P_V(u)\}^T\{D(u)+D_B(u)+D_V(u)\}^{-1}\{P(u)+P_B(u)+P_V(u)\}.\end{aligned}$$

Since

$$\begin{aligned}\{D(u)+D_B(u)+D_V(u)\}^{-1} &= D^{-1}(u) - D^{-2}(u)\{D_B(u)+D_V(u)\}\\ &\quad + 2D^{-3}(u)\{D_B(u)+D_V(u)\}^2 + o_p(h^4+\frac{1}{nh}),\end{aligned}$$

we have

$$\begin{aligned}\hat{\Sigma}^{-1}(u) &= \{P(u)+P_B(u)+P_V(u)\}^T\Big[D^{-1}(u) - D^{-2}(u)\{D_B(u)+D_V(u)\}\\ &\quad + 2D^{-3}(u)\{D_B(u)+D_V(u)\}^2\Big]\{P(u)+P_B(u)+P_V(u)\} + o_p(h^4+\frac{1}{nh})\\ &= P(u)^T D^{-1}(u)P(u) + \{P_B(u)+P_V(u)\}^T D^{-1}(u)P(u) + P(u)^T D^{-1}(u)\{P_B(u)+P_V(u)\}\\ &\quad + \{P_B(u)+P_V(u)\}^T D^{-1}(u)\{P_B(u)+P_V(u)\}\\ &\quad - \Big[P(u)^T D^{-2}(u)\{D_B(u)+D_V(u)\}P(u)\\ &\quad + P(u)^T D^{-2}(u)\{D_B(u)+D_V(u)\}\{P_B(u)+P_V(u)\}\\ &\quad + \{P_B(u)+P_V(u)\}^T D^{-2}(u)\{D_B(u)+D_V(u)\}P(u)\\ &\quad + \{P_B(u)+P_V(u)\}^T D^{-2}(u)\{D_B(u)+D_V(u)\}\{P_B(u)+P_V(u)\}\Big]\\ &\quad + 2\Big[P(u)^T D^{-3}(u)\{D_B(u)+D_V(u)\}^2 P(u)\\ &\quad + P(u)^T D^{-3}(u)\{D_B(u)+D_V(u)\}^2\{P_B(u)+P_V(u)\}\\ &\quad + \{P_B(u)+P_V(u)\}^T D^{-3}(u)\{D_B(u)+D_V(u)\}^2 P(u)\\ &\quad + \{P_B(u)+P_V(u)\}^T D^{-3}(u)\{D_B(u)+D_V(u)\}^2\{P_B(u)+P_V(u)\}\Big] + o_p(h^4+\frac{1}{nh}).\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
\mathrm{trace}(\Sigma(u)\hat{\Sigma}^{-1}(u)) &= p + \mathrm{trace}\Big[\{P_B(u)+P_V(u)\}^T\{P_B(u)+P_V(u)\}P^{-1}(u)P^{-1}(u)^T\Big] \\
&\quad - \mathrm{trace}\Big[D^{-1}(u)\{D_B(u)+D_V(u)\}\Big] + 2\mathrm{trace}\Big[D^{-2}(u)\{D_B(u)+D_V(u)\}^2\Big] \\
&\quad + o_p(h^4+\frac{1}{nh}) \\
&= p + \mathrm{trace}\Big[\{P_B(u)+P_V(u)\}^T\{P_B(u)+P_V(u)\}P^{-1}(u)P^{-1}(u)^T\Big] \\
&\quad + \sum_{j=1}^{p}\Big[-\frac{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)}{\sigma_j^2(u)} + 2\frac{\{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)\}^2}{(\sigma_j^2(u))^2}\Big] + o_p(h^4+\frac{1}{nh}) \qquad (8)
\end{aligned}$$

and

$$\begin{aligned}
\log|\Sigma(u)\hat{\Sigma}^{-1}(u)| &= \log|I - D^{-1}(u)\{D_B(u)+D_V(u)\} + 2D^{-2}(u)\{D_B(u)+D_V(u)\}^2| \\
&\quad + o_p(h^4+\frac{1}{nh}) \\
&= \sum_{j=1}^{p}\log\{1-\frac{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)}{\sigma_j^2(u)} + 2\frac{\{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)\}^2}{(\sigma_j^2(u))^2}\} + o_p(h^4+\frac{1}{nh}) \\
&= \sum_{j=1}^{p}\Big[-\frac{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)}{\sigma_j^2(u)} + 2\frac{\{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)\}^2}{(\sigma_j^2(u))^2} - \frac{\{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)\}^2}{2(\sigma_j^2(u))^2}\Big] \\
&\quad + o_p(h^4+\frac{1}{nh}). \qquad (9)
\end{aligned}$$

By Equation (8) and (9), we have

$$\begin{aligned}
L(\Sigma(u),\hat{\Sigma}(u))_{KL} &= E\Big[\mathrm{trace}[\{P_B(u)+P_V(u)\}^T\{P_B(u)+P_V(u)\}P^{-1}(u)P^{-1}(u)^T] \\
&\quad + \sum_{j=1}^{p}\frac{\{\sigma_{jB}^2(u)+\sigma_{jV}^2(u)\}^2}{2(\sigma_j^2(u))^2}\Big] + o(h^4+\frac{1}{nh}) \\
&= \frac{1}{4}\mu_2^2h^4\mathrm{trace}\Big\{P''(u)^TP''(u)P^{-1}(u)P^{-1}(u)^T + \frac{1}{2}[\{\log D(u)\}'']^2\Big\} \\
&\quad + \frac{\int K^2(y)dy}{nhf(u)}\mathrm{trace}\Big\{P^*(u)P^{-1}(u)P^{-1}(u)^T + I\Big\} + o(h^4+\frac{1}{nh}),
\end{aligned}$$

which establishes the first equation of Theorem 3. Since

$$\begin{aligned} L(\Sigma(u),\hat{\Sigma}(u))_F &= E\Big[\text{trace}\Big\{[\{P_B(u)+P_V(u)\}^T D^{-1}(u)P(u)+P(u)^T D(u)^{-1}\{P_B(u)+P_V(u)\} \\ &\quad -P(u)^T D^{-2}(u)\{D_B(u)+D_V(u)\}P(u)]^2\Big\}\Big]+o(h^4+\frac{1}{nh}) \\ &= \frac{1}{4}\mu_2^2 h^4 \text{trace}\Big\{2P''(u)P''(u)P(u)^T P(u)^T D^{-2}(u) \\ &\quad +2P''(u)^T P''(u)\Sigma^{-1}(u)D^{-1}(u)+\Sigma^{-2}(u)D_{B1}^2(u) \\ &\quad +2\Sigma^{-1}(u)P(u)P''(u)D^{-1}(u)D_{B1}(u)\Big\} \\ &\quad +\frac{\int K^2(y)dy}{nhf(u)}\text{trace}\Big\{P^*(u)\Sigma^{-1}(u)D^{-1}(u)+2\Sigma^{-2}(u)\Big\}+o(h^4+\frac{1}{nh}), \end{aligned}$$

we show the second equation of Theorem 3.

References

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica* **57**, 1027–1057.

Sherman, R. P. (1994). Maximal inequalities for degenerate U-processes with applications to optimization estimators. *Ann. Statist.* **22**, 439–459.