Supplementary file of Asymptotic theory for maximum deviations of sample covariance matrix estimates

Han Xiao[a], Wei Biao Wu[b]

[a]*Department of Statistics and Biostatistics, Rutgers University, 501 Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854, United States*

[b]*Department of Statistics, The University of Chicago, 5734 S. University Ave, Chicago, IL 60637, United States*

Abstract

In this document we give the proofs of Lemma 3, Lemma 4 and Lemma 7 of the main article.

In this document we give the proofs of Lemma 3, Lemma 4 and Lemma 7 of the main article. The lemmas and equations introduced in this document are numbered with a "S"-prefix.

Proof of Lemma 3. Assume X_i has mean zero and variance one. Let $\gamma_k = \mathbb{E}(X_0 X_k)$ be the autocovariance of lag k. Then by Proposition 8, Eq. (34) of [1], we know

$$|\gamma_k| \le \Psi_2 \cdot \Psi_2(|k|). \tag{S.1}$$

(i) Since $\Psi_4 < \infty$, we know for any $\eta > 0$, there exists a $N_1 > 0$ such that $|\gamma_k| < \eta$ when $k \ge N_1$. For $j \le k$, define $\tilde{X}_{k,j} = g(\epsilon_k, \ldots, \epsilon_{j+1}, \epsilon'_j, \epsilon'_{j-1}, \ldots)$, where $(\epsilon'_i)_{i\in\mathbb{Z}}$ is an i.i.d. copy of $(\epsilon_i)_{i\in\mathbb{Z}}$. By Eq. (38) of [1], we know there exists a $N_2 > 0$ such that when $k \ge N_2$, $\|X_k - \tilde{X}_k\|_4 \le \eta$. Set $N = \max\{N_1, N_2\}$, when $k \ge N$, we have

$$\begin{aligned}\mathrm{Var}(X_0X_k) = \mathbb{E}(X_0^2X_k^2) - \gamma_k^2 &= \mathbb{E}\left(X_k^2 X_{k,j}^2\right) + \mathbb{E}\left[X_0^2(X_k^2 - X_{k,j}^2)\right] - \gamma_k^2 \\ &\ge 1 - \eta^2 - 2\|X_0\|_4^3 \cdot \eta.\end{aligned}$$

Therefore, (A1) holds because η can be arbitrarily small.

(ii) We need to show that

$$\sup_{j\ge 0,\, 0\le k\le l,\, (0,j)\ne(k,l)} \mathrm{Cor}(X_0X_j, X_kX_l) < 1.$$

It suffices to show that for some $N > 0$

$$\sup_{j\ge 0,\, 0\le k\le l,\, (0,j)\ne(k,l),\, j+k+l\ge N} \mathrm{Cor}(X_0X_j, X_kX_l) < 1.$$

If $j+k+l \ge N$, then the set $\{0, j, k, l\}$ can be partitioned into two non-empty subsets $\mathcal{B}_1$ and $\mathcal{B}_2$ whose distance is no less than $N/6$. We only consider this type of partitions. If there is a partition such that one of $\mathcal{B}_1$ and $\mathcal{B}_2$ has cardinality one, then similarly as (i), we know for any $\eta > 0$, when N is large enough,

$$|\mathrm{Cov}(X_0X_j, X_kX_l)| = |\mathbb{E}(X_0X_jX_kX_l) - \gamma_j\gamma_{l-k}| \le \eta.$$

If for any partition both $\mathcal{B}_1$ and $\mathcal{B}_2$ has cardinality two, there are two sub-cases. (a) $j < k \le l$ and $k - j \ge N/6$. For any $\eta > 0$, when N is large enough, we have

$$|\mathrm{Cov}(X_0 X_j, X_k X_l)| = |\mathbb{E}\left[X_0 X_j (X_k X_l - X_{k,j} X_{l,j})\right]| \le \eta.$$

(b) $\min\{j, l\} - k \ge N/6$. As in (i), for any $\eta > 0$, when N is large enough, we have $\mathrm{Var}(X_0 X_j) \ge 1 - \eta$, $\mathrm{Var}(X_k X_l) \ge 1 - \eta$, and $|\gamma_j \gamma_{l-k}| < \eta$. On the other hand, the condition $\Psi_4 > 0$ guarantees that the process is non-deterministic, and hence $\gamma := \sup_{t \ge 1} |\gamma_t| < 1$. It follows that when N is large enough

$$\begin{aligned}|\mathbb{E}(X_0 X_j X_k X_l)| &= |\mathbb{E}(X_0 X_{j,k} X_k X_{l,k}) + \mathbb{E}[X_0 X_k (X_j X_l - X_{j,k} X_{l,k})]| \\ &\le \gamma + \eta.\end{aligned}$$

Therefore,

$$|\mathrm{Cor}(X_0 X_j, X_k X_l)| \le (\gamma + 2\eta)/(1 - \eta) < 1$$

when η is small enough. The proof of (ii) is now complete.

(iii) We first consider (A3). Note that

$$\mathrm{Cov}(X_i X_j, X_k X_l) = \mathrm{Cum}(X_i, X_j, X_k, X_l) + \gamma_{i-k}\gamma_{j-l} + \gamma_{i-l}\gamma_{j-k},$$

where $\mathrm{Cum}(X_i, X_j, X_k, X_l)$ is the fourth order joint cumulant of $(X_i, X_j, X_k, X_l)^\top$. Fix a subset $\{i, j\}$, for any integer $b > 0$, there are at most $8b^2$ subsets $\{k, l\}$ such that $\{k.l\} \subset B(i; b) \cup B(j; b)$, where $B(x; r)$ is the open ball $\{y : |x - y| < r\}$. For all other subsets $\{k, l\}$, by (S.1), we have

$$|\gamma_{i-k}\gamma_{j-l} + \gamma_{i-l}\gamma_{j-k}| \le C\Psi_4(b).$$

On the other hand, using similar arguments as Theorem 21 of [1], we can show that

$$|\mathrm{Cum}(X_i, X_j, X_k, X_l)| \le C\Psi_4(\lfloor b/2 \rfloor).$$

Therefore, if $\Psi_4(k) = o(1/\log k)$ as $k \to \infty$, then (A3) holds.
Now we turn to (A3′). Write

$$\mathrm{Cov}(X_i X_j, X_k X_l) = \mathbb{E}(X_i X_j X_k X_l) - \gamma_{i-j}\gamma_{k-l}.$$

By (S.1), it is easily seen that

$$\sum_{1 \le i,j,k.l \le m} \gamma_{i-j}^2 \gamma_{k-l}^2 = O(m^{4-2\delta}).$$

It then suffices to show

$$\sum_{1 \le i \le j \le k \le l \le m} [\mathbb{E}(X_i X_j X_k X_l)]^2 = O(m^{4-\delta}),$$

which is true because by Eq. (38) of [1]

$$[\mathbb{E}(X_i X_j X_k X_l)]^2 = [\mathbb{E}(X_i X_j X_k (X_l - X_{l,k}))]^2 \le 12 \|X_0\|_4^6 [\Psi_4(l - k)]^2.$$

The proof of Lemma 3 is now complete. $\square$

We now give the proof of Lemma 4.

Proof of Lemma 4. Suppose (Y_1, Y_2, Y_3, Y_4) has a joint normal distribution. We can write $Y_i = \alpha_i^\top \boldsymbol{Z}$, where $\boldsymbol{Z}$ is a four dimensional standard Gaussian random vector. For any $0 < \nu < 1$, define the subset of $\mathbb{R}^{16}$,

$$D_\nu = \left\{ (\alpha_1^\top, \alpha_2^\top, \alpha_3^\top, \alpha_4^\top) : |\alpha_i|^2 = 1 \text{ and } |\alpha_i^\top \alpha_j| \le 1 - \nu \text{ for } 1 \le i \ne j \le 4 \right\}.$$

Since $|\mathrm{Cor}(Y_1Y_2, Y_3Y_4)|$ is a continuous function on D_ν, and D_ν is compact, the maximum correlation is attained at some point in D_ν.

On the other hand, elementary calculation shows that $\mathrm{Cor}(Y_1Y_2, Y_3Y_4) = 1$ if and only if Y_1, Y_2, Y_3, Y_4 are all perfectly correlated. The proof is now complete. $\square$

The proof of Lemma 7 is a refined version of that of Lemma 20 in [1]. We need the following bounds on normal tail probabilities, which are taken from Lemma 19 of [1].

Denote by $\varphi_d((r_{ij}); x_1, \ldots, x_d)$ the density of a d-dimensional multivariate normal random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ with mean zero and covariance matrix (r_{ij}), where we always assume $r_{ii} = 1$ for $1 \le i \le d$ and (r_{ij}) is nonsingular. Let

$$Q_d\left((r_{ij}); z\right) = \int_z^\infty \cdots \int_z^\infty \varphi_d\left((r_{ij}), x_1, \ldots, x_d\right) \, \mathrm{d}x_d \cdots \mathrm{d}x_1.$$

Lemma S.1. *For every $z > 0$, $0 < s < 1$, $d \ge 1$ and $\epsilon > 0$, there exists positive constants C_d and ϵ_d such that for $0 < \epsilon < \epsilon_d$*

1. *if $|r_{ij}| < \epsilon$ for all $1 \le i < j \le d$, then*

$$Q_d\left((r_{ij}); z\right) \le C_d \, f_d(\epsilon, 1/z) \exp\left\{ -\left(\frac{d}{2} - C_d \epsilon \right) z^2 \right\} \tag{S.2}$$

where $f_{2k}(x, y) = \sum_{l=0}^{k} x^l y^{2(k-l)}$ and $f_{2k-1}(x, y) = \sum_{l=0}^{k-1} x^l y^{2(k-l)-1}$ for $k \ge 1$;

2. *if for all $1 \le i < j \le d+1$ such that $(i, j) \ne (1, 2)$, $|r_{ij}| \le \epsilon$, then*

$$Q_{d+1}\left((r_{ij}); z\right) \le C_d \exp\left\{ -\left(\frac{(1 - |r_{12}|)^2 + d}{2} - C_d \epsilon \right) z^2 \right\}. \tag{S.3}$$

We first give a one-sided version of Lemma 7 and its proof, then we show how it implies Lemma 7.

Lemma S.2. *Assume either (B1) or (B2). For a positive real number z_n, define the event $A_{n,i}$ and $Q_{n,d}$ as*

$$A_{n,i} = \{X_{n,i} > z_n\} \quad \text{and} \quad Q_{n,d} = \sum_{\mathcal{A} \subset \mathcal{I}_n, |\mathcal{A}| = d} P\left(\bigcap_{i \in \mathcal{A}} A_{n,i} \right).$$

If z_n satisfies that $z_n^2 = 2 \log s_n - \log\log s_n - \log(4\pi) + 2z + o(1)$, then for all $d \ge 1$

$$\lim_{n \to \infty} Q_{n,d} = \frac{e^{-dz}}{d!}.$$

Proof. The following facts about normal tail probabilities are well-known:

$$P(X_1 \geq x) \leq \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \text{ for } x > 0 \quad \text{and} \quad \lim_{x\to\infty} \frac{P(X_1 \geq x)}{(1/x)(2\pi)^{-1/2}\exp\{-x^2/2\}} = 1. \tag{S.4}$$

By the assumption on z_n, if for each n, $X_{n,i}$, $i \in \mathcal{I}_n$ are i.i.d., then by (S.4),

$$\begin{aligned} \lim_{n\to\infty} Q_{n,d} &= \lim_{n\to\infty} \binom{n}{d} Q_d(I_d, z_n) \\ &= \lim_{n\to\infty} \binom{n}{d} \frac{1}{(2\pi)^{d/2} z_n^d} \exp\left\{-\frac{dz_n^2}{2}\right\} = \frac{e^{-dz}}{d!}. \end{aligned}$$

When the $X_{n,i}$'s are dependent, the result is still trivially true when $d = 1$. Now we deal with the $d \geq 2$ case. Suppose (b_n) is a sequence of positive numbers which approaches infinity. For each subset J of $\mathcal{I}_n$ with cardinality $|J| = d$, we define an undirected graph $\mathscr{G}(J)$ by identifying each $i \in J$ with a node and saying i and j are adjacent if $|r_{n,i,j}| > \gamma(n, b_n)$. Suppose the graph $\mathscr{G}(J)$ has $d - s$ connected components $\mathcal{B}_1, \ldots, \mathcal{B}_{d-s}$. If $s \geq 1$, assume w.l.o.g. that $|\mathcal{B}_1| \geq 2$. Pick $k_0, k_1 \in \mathcal{B}_1$, and $k_p \in \mathcal{B}_p$ for $2 \leq p \leq d - s$, and set $K = \{k_0, k_1, k_2, \ldots, k_{d-s}\}$. Define $Q_J = P(\cap_{k\in J} A_k)$ and Q_K similarly, then $Q_J \leq Q_K$. By (S.3) of Lemma S.1, there exists a number $M > 1$ depending on d and the sequences (γ_n) and (b_n), such that when $n \geq M$,

$$\begin{aligned} Q_K &\leq C_{d-s} \exp\left\{-\left(\frac{(1-\gamma_n)^2 + d - s}{2} - C_{d-s}\gamma(n, b_n)\right) z_n^2\right\} \\ &\leq C_{d-s} \exp\left\{-\left(\frac{d-s}{2} + \frac{(1-\gamma_n)^2}{3}\right) z_n^2\right\}. \end{aligned}$$

Note that $z_n^2 = 2\log s_n - \log\log s_n + O(1)$. Pick $b_n = \lfloor s_n^\alpha \rfloor$ for some $\alpha < (1-\gamma_n)^2/3d$. For any $1 \leq a \leq d-1$, since there are at most $O\left(b_n^a s_n^{d-a}\right)$ subsets $J \subset \mathcal{I}_n$ such that $|J| = d$ and the graph $\mathscr{G}(L)$ has $d - a$ connected components, we know the sum of Q_J over these J is dominated by

$$C_{d-a} \exp\left\{\log s_n \left((d-a) + \frac{2(d-1)(1-\gamma_n)^2}{3d} - (d-a) - \frac{2(1-\gamma_n)^2}{3}\right)\right\}$$

when n is large enough, which converges to zero. Therefore, it remains to consider all the subsets $J \subset \mathcal{I}_n$ such that the graph $\mathscr{G}(J)$ has no edges.

Let $J \subset \mathcal{I}_n$ be a subset such that $|J| = d$, and $|r_{n,i,j}| < \gamma(n, b_n)$ for all pairs i, j such that $i, j \in J$ and $i \neq j$, and $\mathscr{J}(d, b_n)$ be the collection of all such subsets. Let $(r_{ij})_{i,j\in J}$ be the d-dimensional covariance matrix of $\boldsymbol{X}_J := (X_{n,i})_{i\in J}$. There exists a matrix $R_J = \theta(r_{ij})_{i,j\in J} + (1-\theta)I_d$ for some $0 < \theta < 1$ such that

$$Q_J - Q_d(I_d, z_n) = \sum_{h,l\in J, h<l} \frac{\partial Q_d}{\partial r_{hl}}[R_J; z_n] r_{hl}.$$

Let R_H, $H = J \setminus \{h, l\}$, be the correlation matrix of the conditional distribution of $\boldsymbol{X}_H$ given X_h and X_l. By (S.2) of Lemma S.1, for n large enough

$$\begin{aligned} \frac{\partial Q_d}{\partial r_{hl}}[R_J; z_n] &\leq C \exp\left\{-\frac{z_n^2}{1 + |r_{n,h,l}|}\right\} \cdot Q_{d-2}\left(R_K; (1 - 3\gamma(n, b_n))z_n\right) \\ &\leq CC_{d-2} f_{d-2}(\gamma(n, b_n), 1/z_n) \exp\left\{-\frac{z_n^2}{1 + |r_{n,h,l}|}\right\} \end{aligned}$$

$$\times \exp\left\{-\left(\frac{d-2}{2}-2C_{d-2}\gamma(n,b_n)\right)(1-3\gamma(n,b_n))^2 z_n^2\right\}$$
$$\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)$$
$$\times \exp\left\{-\left(\frac{d}{2}-(2C_{d-2}+3(d-2))\gamma(n,b_n)-|r_{n,h,l}|\right) z_n^2\right\}$$
$$\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)\exp\left\{-\left(\frac{d}{2}-C_d\gamma(n,b_n)\right) z_n^2\right\}.$$

It follows that

$$\begin{aligned}
&\sum_{J\in\mathcal{J}(d,b_n)}|Q_J-Q_d(I_d;z_n)| \\
&\le C_d f_{d-2}(\gamma(n,b_n),1/z_n) \\
&\quad\times \sum_{J\in\mathcal{J}(d,b_n)}\sum_{i,j\in J;\, i\neq j}\exp\left\{-\left(\frac{d}{2}-C_d\gamma(n,b_n)\right) z_n^2\right\}|r_{n,i,j}| \\
&\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)s_n^{d-2} \\
&\quad\times \sideset{}{^*}\sum_{i,j\in\mathcal{I}_n}\exp\left\{-\left(\frac{d}{2}-C_d\gamma(n,b_n)\right) z_n^2\right\}|r_{n,i,j}|,
\end{aligned} \tag{S.5}$$

where the sum $\sum^*_{i,j\in\mathcal{I}_n}$ is over all the pair (i,j) such that $|r_{n,i,j}|\le\gamma(n,b_n)$. Under the assumption (B1), we have

$$\begin{aligned}
&\sum_{J\in\mathcal{J}(d,b_n)}|Q_J-Q_d(I_d;z_n)| \\
&\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)(\log s_n)^{d/2}\gamma(n,b_n)\exp\left\{C_d\gamma(n,b_n)(\log s_n)\right\}.
\end{aligned} \tag{S.6}$$

Since $\lim_{n\to\infty}\gamma(n,b_n)\log b_n=0$, it holds that $\lim_{n\to\infty}\gamma(n,b_n)\log s_n=0$. Using the fact that $\lim_{n\to\infty}(\log s_n)^{1/2}/z_n=2^{-1/2}$, we have $\lim_{n\to\infty}f_{d-2}(\gamma(n,b_n),1/z_n)(\log s_n)^{d/2-1}=2^{-d/2+1}$. Therefore, the term in (S.6) converges to zero, and the theorem holds under (B1).

Alternatively, if (B2) is true, from (S.5) we have

$$\begin{aligned}
&\sum_{J\in\mathcal{J}(d,b_n)}|Q_J-Q_d(I_d;z_n)| \\
&\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)s_n^{-2}(\log s_n)^{d/2}\sideset{}{^*}\sum_{i,j\in\mathcal{I}_n}\exp\left\{C_d\gamma(n,b_n)(\log s_n)\right\}|r_{n,i,j}| \\
&\le C_d f_{d-2}(\gamma(n,b_n),1/z_n)s_n^{-1}(\log s_n)^{d/2}\exp\left\{C_d\gamma(n,b_n)(\log s_n)\right\}\left(\sum_{i,j\in\mathcal{I}_n}r_{n,i,j}^2\right)^{1/2} \\
&\le C_d s_n^{-\delta/2}(\log s_n)\exp\left\{C_d\gamma(n,b_n)(\log s_n)\right\}=o(1),
\end{aligned}$$

and the proof is complete. □

Now we give the proof of Lemma 7.

Proof of Lemma 7. In the proof of Theorem S.2, the upper bounds on Q_J and $|Q_J - Q(I_d; z_n)|$ are expressed through the absolute values of the covariances, so we can obtain the same bounds for probabilities of the form $P(\cap_{1\le i\le d}\{(-1)^{a_i} X_{t_i} \ge z_n\})$ for any $(a_1, \ldots, a_d) \in \{0, 1\}^d$. Based on this observation, Lemma 7 is an immediate consequence of Lemma S.2.

$\square$

[1] H. Xiao, W. B. Wu, Asymptotic inference of autocovariances of stationary processes, *preprint*, available at `http://arxiv.org/abs/1105.3423` (2011).