Errata for Proof of Eq. (S2.42)

Jiaxin Qiu, Zeng Li, Jianfeng Yao

(Latest Update: July 30, 2025)

Abstract

This document provides the errata for the proof of Eq. (S2.42) in Qiu et al. [2023]: for any $\varepsilon > 0$ and any $\ell > 0$,

$$\Pr(\|\mathbf{B}_n\| \geqslant \eta + \varepsilon) = o(n^{-\ell}), \tag{1}$$

where $\eta = 2\limsup \|\Sigma\|/\sqrt{\theta}$, $\theta = \lim_p p^{-1}\mathrm{tr}(\Sigma^2)$. To correct the proof, we impose an additional assumption that $p = \Theta(n^t)$ for some $t > 1$.

1 Counterexample for Eq. (S2.45)

In the proof of Eq. (S2.42) in Qiu et al. [2023], we used an intermediate inequality (S2.45) as follows:

$$\max_{\|\mathbf{z}\|=1} \sum_{i=1}^{n} \left\{ \sum_{j \neq i} (\mathbf{x}_i' \Sigma \mathbf{x}_j) z_j \right\}^2 \leqslant \|\Sigma\|^2 \max_{\|\mathbf{z}\|=1} \sum_{i=1}^{n} \left\{ \sum_{j \neq i} (\mathbf{x}_i' \mathbf{x}_j) z_j \right\}^2.$$

Unfortunately, this inequality turns out to be incorrect. The following provides a counterexample:

$$p = n = 2, \quad \mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{pmatrix}, \quad 0 < \varepsilon < 1, \quad \Longrightarrow \quad \text{LHS} = \varepsilon^2 > \text{RHS} = 0.$$

2 Corrected Proof of Eq. (S2.42)

Sketch of the proof:

- **Step 1**: Establish the result for the Gaussian case. Under the Gaussian assumption, Σ can be diagonalized, then we apply the method of Chen and Pan [2012] to prove the result.
- **Step 2**: Use the Lindeberg replacement method and the mathematical induction to generalize the result to the non-Gaussian case.
  - To apply the Lindeberg replacement method, we impose an additional assumption that $p = \Theta(n^t)$ for some $t > 1$. This growth condition ensures that the total error introduced by the replacement steps remains negligible in the asymptotic regime.

2.1 Step 1: Proof under Gaussian case

Recall that

$$\mathbf{B}_n = \frac{1}{\sqrt{npb_p}} \underbrace{\left\{ \mathbf{X}'\Sigma\mathbf{X} - \mathrm{diag}(\mathbf{X}'\Sigma\mathbf{X}) \right\}}_{=:\mathbf{B}_n^\circ}.$$

We take $\delta = \delta_n$ even $k = k_n$ and satisfying, as $n, p \to \infty$, $\delta \to 0$, $\delta (np)^{1/4} \to 0$,

$$\frac{k}{\log n} \to \infty, \quad \frac{\delta^{1/3} k}{\log n} \to 0, \quad \frac{\delta^2 \sqrt[4]{n}}{k^3} \geqslant 1, \quad \frac{\delta^2 \sqrt{p/n}}{k} > 1, \quad \frac{k}{\log(\delta^2 \sqrt{p/n}/k)} > 1, \quad \frac{\delta^2 k^3}{\log(2\delta^2 \sqrt{p/n}/k)} \leqslant 1. \tag{2}$$

Remark 1.

- We assume that k is even such that the inequality holds: $\|\mathbf{B}_n\|^k \leqslant \sum_{i=1}^n \lambda_i^k(\mathbf{B}_n) = \mathrm{tr}(\mathbf{B}_n^k)$.
- The first three conditions for k in (2) are taken from Chen and Pan [2012] to obtain (3) below. The 4th and 5th conditions are used in (8), and the last condition is used in (11).
- Recall that $p = \Theta(n^t)$ for some $t > 1$. Let $\delta = n^{-\alpha}$ and $k = k_n = (\log n)^2$. For any $t > 1$, we can choose $\alpha = \min\{\frac{1}{16}, \frac{t-1}{8}\}$ to make all the conditions in (2) hold.

We assume that the elements of $\mathbf{X}$ are i.i.d. standard Gaussian, and denote x_{ij} by z_{ij}. Moreover, we denote the corresponding $\mathbf{B}_n$ as $\mathbf{B}_\mathrm{g}$. Without loss of generality, we assume that Σ is diagonal with eigenvalues $\sigma_1 \geqslant \cdots \geqslant \sigma_p > 0$.

Using the bound of the case where $\Sigma = \mathbf{I}$ shown in Chen and Pan [2012, Page 1418], we have

$$\begin{aligned} 0 \leqslant \mathbb{E}\mathrm{tr}(\mathbf{B}_\mathrm{g}^k) &= \frac{1}{(\sqrt{npb_p})^k} \sum \mathbb{E}\Big\{\Big(\mathbf{Z}'\Sigma\mathbf{Z}\Big)_{j_1j_2}\Big(\mathbf{Z}'\Sigma\mathbf{Z}\Big)_{j_2j_3}\cdots\Big(\mathbf{Z}'\Sigma\mathbf{Z}\Big)_{j_kj_1}\Big\} \\ &\leqslant \frac{1}{(\sqrt{npb_p})^k} \sum \mathbb{E}\Big|\Big(z_{i_1j_1}\sigma_{i_1}z_{i_1j_2}\Big)\Big(z_{i_2j_2}\sigma_{i_2}z_{i_2j_3}\Big)\cdots\Big(z_{i_kj_k}\sigma_{i_k}z_{i_kj_1}\Big)\Big| \\ &= \left(\frac{\|\Sigma\|}{\sqrt{b_p}}\right)^k \cdot \underbrace{\frac{1}{(np)^{k/2}} \sum \mathbb{E}\Big|z_{i_1j_1}z_{i_1j_2}z_{i_2j_2}z_{i_2j_3}\cdots z_{i_kj_k}z_{i_kj_1}\Big|}_{\leqslant (2+\zeta)^k \text{ by Chen and Pan [2012, Page 1418]}} \leqslant \left\{\frac{\|\Sigma\|(2+\zeta)}{\sqrt{b_p}}\right\}^k, \end{aligned} \tag{3}$$

where $0 < \zeta < \varepsilon/\eta$, and the summation is over all $i_1, \ldots, i_k \in [p]$ and $j_1, \ldots, j_k \in [n]$ such that $j_1, \ldots, j_k$ are distinct. Thus, we have

$$\mathbb{P}(\|\mathbf{B}_\mathrm{g}\| \geqslant \eta + \varepsilon) \leqslant \frac{\mathbb{E}\|\mathbf{B}_\mathrm{g}\|^k}{(\eta+\varepsilon)^k} \leqslant \frac{\mathbb{E}\mathrm{tr}(\mathbf{B}_\mathrm{g}^k)}{(\eta+\varepsilon)^k} \leqslant \left(\frac{2+\zeta}{\eta+\varepsilon}\cdot\frac{\|\Sigma\|}{\sqrt{b_p}}\right)^k = o(n^{-\ell}),$$

since $k/\log n \to \infty$ and $\frac{2+\zeta}{\eta+\varepsilon}\cdot\frac{\|\Sigma\|}{\sqrt{b_p}} < 1$.

2.2 Step 2: Extension to non-Gaussian case

As shown in the last section, the key input is the upper bound of $\mathbb{E}\mathrm{tr}(\mathbf{B}_n^k)$ as in (3). We can apply the Lindeberg replacement trick to extend the result to the non-Gaussian case.

We define $\mathbf{B}_{i,j}$ as the matrix obtained by replacing all entries $z_{k\ell}$ in $\mathbf{B}_\mathrm{g}$ with $x_{k\ell}$ for all (k,ℓ) such that $k < i$ or $k = i$ and $\ell \leq j$. Then, we have $\mathbf{B}_n = \mathbf{B}_{p,n}$ and $\mathbf{B}_\mathrm{g} = \mathbf{B}_{1,0}$, and we can write

$$\mathbb{E}\mathrm{tr}(\mathbf{B}_n^k) - \mathbb{E}\mathrm{tr}(\mathbf{B}_\mathrm{g}^k) = \sum_{i=1}^p\sum_{j=1}^n \Big\{\mathbb{E}\mathrm{tr}(\mathbf{B}_{i,j}^k) - \mathbb{E}\mathrm{tr}(\mathbf{B}_{i,j-1}^k)\Big\}.$$

We claim that (will be proved later, see Page 3) for any i and j,

$$\mathbb{E}\mathrm{tr}(\mathbf{B}_{i,j}^k) - \mathbb{E}\mathrm{tr}(\mathbf{B}_{i,j-1}^k) \lesssim \left\{\frac{\|\Sigma\|(2+2\zeta)}{\sqrt{b_p}}\right\}^k. \tag{4}$$

Starting from $\mathbf{B}_\mathrm{g} = \mathbf{B}_{1,0}$, after pn replacements we obtain $\mathbf{B}_n = \mathbf{B}_{p,n}$, and we have

$$\mathbb{E}\mathrm{tr}(\mathbf{B}_n^k) \lesssim (pn)\left\{\frac{\|\Sigma\|(2+2\zeta)}{\sqrt{b_p}}\right\}^k,$$

and therefore

$$\mathbb{P}(\|\mathbf{B}_n\| \geqslant \eta + \varepsilon) \leqslant \frac{\mathbb{E}\mathrm{tr}(\mathbf{B}_n^k)}{(\eta+\varepsilon)^k} \lesssim (pn)\left(\frac{2+2\zeta}{\eta+\varepsilon}\cdot\frac{\|\Sigma\|}{\sqrt{b_p}}\right)^k = o(n^{-\ell}), \tag{5}$$

the last equality follows from (2), $p = O(n^\alpha)$, and the fact that $\frac{2+2\zeta}{\eta+\varepsilon}\cdot\frac{\|\Sigma\|}{\sqrt{b_p}} < 1$.

Remark 2. The additional assumption $p = \Theta(n^t)$ is necessary to derive the last equality in (3). Let $\gamma = \frac{2+2\zeta}{\eta+\varepsilon} \cdot \frac{\|\Sigma\|}{\sqrt{b_p}} < 1$. Then, we have

$$\frac{\log(pn\gamma^k n^\ell)}{\log n} \lesssim \frac{k}{\log n}\log\gamma + (t+\ell+1) \xrightarrow{(2)} -\infty \quad \Longrightarrow \quad pn\gamma^k = o(n^{-\ell}).$$

Now, we prove (4). We only prove the case $i = j = 1$ for simplicity, and the other cases can be proved similarly. Let $\mathbf{Z}_{(1)} = \mathbf{Z} - \mathbf{z}_1\mathbf{e}_1'$, we have

$$\begin{aligned}
\mathbf{B}_{1,0}^\circ = \mathbf{B}_{\mathrm{g}}^\circ &= \{\mathbf{Z}_{(1)} + \mathbf{z}_1\mathbf{e}_1'\}'\Sigma\{\mathbf{Z}_{(1)} + \mathbf{z}_1\mathbf{e}_1'\} - \mathrm{diag}(\mathbf{Z}'\Sigma\mathbf{Z}) \\
&= \mathbf{Z}_{(1)}'\Sigma\mathbf{z}_1\mathbf{e}_1' + \mathbf{e}_1\mathbf{z}_1'\Sigma\mathbf{Z}_{(1)} + \mathbf{Z}_{(1)}'\Sigma\mathbf{Z}_{(1)} + \mathbf{e}_1\mathbf{z}_1'\Sigma\mathbf{z}_1\mathbf{e}_1' - \mathrm{diag}(\mathbf{Z}'\Sigma\mathbf{Z}) \\
&= z_{11}\Big(\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\mathbf{e}_1' + \mathbf{e}_1\boldsymbol{\sigma}_1'\mathbf{Z}_{(1)}\Big) + \boldsymbol{\Delta}_{1,1},
\end{aligned}$$

where $\boldsymbol{\sigma}_1$ is the first column of Σ, and $\boldsymbol{\Delta}_{1,1}$ is obtained by replacing z_{11} in $\mathbf{B}_{1,0}^\circ$ with zero. Similarly, we have

$$\mathbf{B}_{1,1}^\circ = x_{11}\Big(\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\mathbf{e}_1' + \mathbf{e}_1\boldsymbol{\sigma}_1'\mathbf{Z}_{(1)}\Big) + \boldsymbol{\Delta}_{1,1}.$$

Using these notations, we have

$$\begin{aligned}
&\Big|\mathbb{E}\mathrm{tr}(\mathbf{B}_{1,1}^k) - \mathbb{E}\mathrm{tr}(\mathbf{B}_{1,0}^k)\Big| \\
&= \frac{1}{(npb_p)^{k/2}}\Big|\mathbb{E}\mathrm{tr}\Big\{x_{11}\Big(\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\mathbf{e}_1' + \mathbf{e}_1\boldsymbol{\sigma}_1'\mathbf{Z}_{(1)}\Big) + \boldsymbol{\Delta}_{1,1}\Big\}^k - \mathbb{E}\mathrm{tr}\Big\{z_{11}\Big(\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\mathbf{e}_1' + \mathbf{e}_1\boldsymbol{\sigma}_1'\mathbf{Z}_{(1)}\Big) + \boldsymbol{\Delta}_{1,1}\Big\}^k\Big| \\
&\lesssim \frac{1}{(npb_p)^{k/2}}\left[\mathbb{E}\sum_{\ell=1}^k\binom{k}{\ell}\Big\{2|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|\Big\}^\ell\|\boldsymbol{\Delta}_{1,1}\|^{k-\ell} + \mathbb{E}\sum_{\ell=1}^k\binom{k}{\ell}\Big\{2|z_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|\Big\}^\ell\|\boldsymbol{\Delta}_{1,1}\|^{k-\ell}\right] \\
&\leqslant \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big(2|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\| + \|\boldsymbol{\Delta}_{1,1}\|\Big)^k + \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big(2|z_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\| + \|\boldsymbol{\Delta}_{1,1}\|\Big)^k,
\end{aligned}$$

where in the first inequality, we use the fact that $\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\mathbf{e}_1' + \mathbf{e}_1\boldsymbol{\sigma}_1'\mathbf{Z}_{(1)}$ is of rank 2, and $\mathrm{tr}(\mathbf{A}) \leqslant \mathrm{r}(\mathbf{A})\|\mathbf{A}\|$. It suffices to estimate the first term on the RHS, since the second term have the same bound. We have

$$\begin{aligned}
&\frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big(2|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\| + \|\boldsymbol{\Delta}_{1,1}\|\Big)^k \\
&\leqslant \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big(2|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\| + \|\boldsymbol{\Delta}_{1,1}\|\Big)^k\mathbb{1}\Big\{\|\boldsymbol{\Delta}_{1,1}\| \geqslant 2\delta^{-1}|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|\Big\} \\
&\qquad + \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big(2|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\| + \|\boldsymbol{\Delta}_{1,1}\|\Big)^k\mathbb{1}\Big\{\|\boldsymbol{\Delta}_{1,1}\| < 2\delta^{-1}|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|\Big\} \\
&\leqslant \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big\{(1+\delta)\|\boldsymbol{\Delta}_{1,1}\|\Big\}^k + \frac{1}{(npb_p)^{k/2}}\mathbb{E}\Big\{2(1+\delta^{-1})|x_{11}|\cdot\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|\Big\}^k \\
&=: \mathcal{I}_1 + \mathcal{I}_2.
\end{aligned}$$

We claim that (will be proved later, see Page 5) for any even k,

$$\mathbb{E}\mathrm{tr}(\boldsymbol{\Delta}_{1,1}^k) \leqslant \mathbb{E}\mathrm{tr}(\mathbf{B}_{1,0}^{\circ k}). \tag{6}$$

Using this fact, we have

$$\frac{\mathbb{E}\|\boldsymbol{\Delta}_{1,1}\|^k}{(npb_p)^{k/2}} \leqslant \frac{\mathbb{E}\mathrm{tr}(\boldsymbol{\Delta}_{1,1}^k)}{(npb_p)^{k/2}} \leqslant \mathrm{tr}(\mathbf{B}_{1,0}^k) \leqslant \left\{\frac{\|\Sigma\|(2+\zeta)}{\sqrt{b_p}}\right\}^k, \tag{7}$$

and thus

$$\mathcal{I}_1 \leqslant \left\{\frac{\|\Sigma\|(1+\delta)(2+\zeta)}{\sqrt{b_p}}\right\}^k \leqslant \left\{\frac{\|\Sigma\|(2+2\zeta)}{\sqrt{b_p}}\right\}^k.$$

Using multinomial expansion, we have

$$\begin{aligned}
\mathbb{E}\|\mathbf{Z}_{(1)}'\boldsymbol{\sigma}_1\|^{2k} &= \mathbb{E}\left\{\sum_{j=2}^{n}(\mathbf{z}_j'\boldsymbol{\sigma}_1)^2\right\}^k \\
&= \sum_{\ell=1}^{k}\sum_{2\leqslant j_1<\ldots<j_\ell\leqslant n}\sum_{\substack{k_1+\cdots+k_\ell=k\\ k_1,\ldots,k_\ell\geqslant 1}}\binom{k}{k_1,\ldots,k_\ell}\mathbb{E}(\mathbf{z}_{j_1}'\boldsymbol{\sigma}_1)^{2k_1}\cdots\mathbb{E}(\mathbf{z}_{j_\ell}'\boldsymbol{\sigma}_1)^{2k_\ell} \\
&\overset{(9)}{\leqslant} \sum_{\ell=1}^{k} n^\ell \ell^k \frac{\{k\delta(np)^{1/4}\|\Sigma\|\}^{2k}k^\ell}{\{\delta^2(np)^{1/2}\}^\ell} \\
&= \{k\delta(np)^{1/4}\|\Sigma\|\}^{2k}\sum_{\ell=1}^{k}\frac{\ell^k}{\{\delta^2(p/n)^{1/2}/k\}^\ell} \\
&\overset{(2),(10)}{\leqslant} \{k\delta(np)^{1/4}\|\Sigma\|\}^{2k}k\left(\frac{k}{\log\{\delta^2(p/n)^{1/2}/k\}}\right)^k,
\end{aligned} \tag{8}$$

where

- the first inequality follows from, for any $d \leqslant 2k$,

$$\begin{aligned}
\mathbb{E}|\mathbf{z}_2'\boldsymbol{\sigma}_1|^d &= \mathbb{E}\left|\sum_{i=1}^{p}\sigma_{i1}z_{i2}\right|^d \\
&= \sum_{\ell=1}^{d/2}\sum_{1\leqslant i_1<\ldots<i_\ell\leqslant p}\sum_{\substack{d_1+\cdots+d_\ell=d\\ d_1,\ldots,d_\ell\geqslant 2}}\binom{d}{d_1,\cdots,d_\ell}\mathbb{E}|\sigma_{i_1 1}z_{i_1 2}|^{d_1}\cdots\mathbb{E}|\sigma_{i_\ell 1}z_{i_\ell 2}|^{d_\ell} \\
&\leqslant \sum_{\ell=1}^{d/2}\sum_{1\leqslant i_1<\ldots<i_\ell\leqslant p}|\sigma_{i_1 1}|^2\cdots|\sigma_{i_\ell 1}|^2\sum_{\substack{d_1+\cdots+d_\ell=d\\ d_1,\ldots,d_\ell\geqslant 2}}\binom{d}{d_1,\cdots,d_\ell}\|\Sigma\|^{d-2\ell}\{\delta(np)^{1/4}\}^{d-2\ell} \\
&\leqslant \sum_{\ell=1}^{d/2}\|\boldsymbol{\sigma}_1\|^{2\ell}\ell^d\|\Sigma\|^{d-2\ell}\{\delta(np)^{1/4}\}^{d-2\ell} \\
&\leqslant \frac{\{\delta(np)^{1/4}\|\Sigma\|k\}^d k}{\delta^2(np)^{1/2}}.
\end{aligned} \tag{9}$$

- the last inequality follows from the (2) and the inequality [Yin et al., 1988, Page 519]

$$a^{-t}t^b \leqslant \left(\frac{b}{\log a}\right)^b, \quad a>1, b>0, t\geqslant 1, \frac{b}{\log a}>1. \tag{10}$$

The above estimates yield

$$\begin{aligned}
\mathcal{I}_2 &\leqslant \frac{1}{(npb_p)^{k/2}}\{2(1+\delta^{-1})\}^k\cdot\{\delta(np)^{1/4}\}^k\cdot\left\{\frac{k}{2}\delta(np)^{1/4}\|\Sigma\|\right\}^k\frac{k}{2}\left(\frac{k/2}{\log\{2\delta^2(p/n)^{1/2}/k\}}\right)^{k/2} \\
&\leqslant \left(\frac{\|\Sigma\|}{\sqrt{b_p}}\right)^k\underbrace{\{(k/2)^{1/k}\}^k}_{\to 1}\underbrace{\left(\frac{2\delta^2k^3}{\log\{2\delta^2(p/n)^{1/2}/k\}}\right)^{k/2}}_{<2} \\
&\overset{(2)}{\leqslant} \left\{\frac{\|\Sigma\|(2+\varepsilon_0)}{\sqrt{b_p}}\right\}^k,
\end{aligned} \tag{11}$$

for some small constant $0<\varepsilon_0<1$. The above estimates complete the proof of (4) for the case $(i,j)=(1,1)$.

Assume that (4) holds for $(i,j)=(i_0,j_0)$, and we consider the case $(i,j)=(i_0,j_0+1)$. Similar to (6) – (7), we have

$$\frac{\mathbb{E}\mathrm{tr}(\boldsymbol{\Delta}_{i_0,j_0+1}^k)}{(npb_p)^{k/2}} \leqslant \mathbb{E}\mathrm{tr}(\mathbf{B}_{i_0,j_0}^k) \lesssim \left\{\frac{\|\Sigma\|(2+2\zeta)}{\sqrt{b_p}}\right\}^k,$$

where the last inequality is proven in the previous case $(i,j)=(i_0,j_0)$. All the other estimates are similar to the proof of the case $(i,j)=(1,1)$. Thus, we complete the proof of (4) by using the mathematical induction.

Proof of (6). Recall that $\mathbf{B}_{1,0}^{\circ}=\boldsymbol{\Delta}_{1,1}+z_{11}\mathbf{H}$, where $\mathbf{H}=\mathbf{Z}_{(1)}'\sigma_1\mathbf{e}_1'+\mathbf{e}_1\sigma_1'\mathbf{Z}_{(1)}$ is a rank-2 symmetry matrix. For notation simplicity, we denote $\mathbf{B}_{1,0}^{\circ}$ as $\mathbf{B}(z_{11})$ and let $\psi(z_{11})=\mathrm{tr}(\mathbf{B}(z_{11})^k)$, where k is any even integer.

We claim that $\psi(z_{11})$ is convex in z_{11} (will be proved later). Using the conditional Jensen's inequality, we have

$$\mathbb{E}\{\psi(z_{11})\mid z_{ij},(i,j)\neq(1,1)\}\geqslant\psi(\mathbb{E}[z_{11}])=\psi(0)=\mathrm{tr}(\boldsymbol{\Delta}_{1,1}^k),$$

that is,

$$\mathbb{E}\Big\{\mathrm{tr}(\mathbf{B}_{1,0}^{\circ k})-\mathrm{tr}(\boldsymbol{\Delta}_{1,1}^k)\mid z_{ij},(i,j)\neq(1,1)\Big\}\geqslant 0.$$

Taking the outer expectation, we obtain (6).

Finally, we prove the convexity of $\psi(z_{11})$. Using the non-commutative chain rule for matrix powers, we have $\partial\mathbf{B}(z_{11})^k/\partial z_{11}=\sum_{j=0}^{k-1}\mathbf{B}(z_{11})^j\mathbf{H}\mathbf{B}(z_{11})^{k-1-j}$, so that $\psi'(z_{11})=k\mathrm{tr}(\mathbf{H}\mathbf{B}(z_{11})^{k-1})$ and

$$\psi''(z_{11})=k\sum_{j=0}^{k-2}\mathrm{tr}\Big(\mathbf{H}\mathbf{B}(z_{11})^j\mathbf{H}\mathbf{B}(z_{11})^{k-2-j}\Big).$$

Diagonalizing the symmetric matrix $\mathbf{B}(z_{11})=\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{T}}$ and writing $\widetilde{\mathbf{H}}=\mathbf{U}^{\mathsf{T}}\mathbf{H}\mathbf{U}$, we obtain

$$\mathrm{tr}\Big(\mathbf{H}\mathbf{B}(z_{11})^j\mathbf{H}\mathbf{B}(z_{11})^{k-2-j}\Big)=\mathrm{tr}\Big(\widetilde{\mathbf{H}}\boldsymbol{\Lambda}^j\widetilde{\mathbf{H}}\boldsymbol{\Lambda}^{k-2-j}\Big)=\sum_{i=1}^n\sum_{\ell=1}^n\Lambda_{\ell\ell}^j\Lambda_{ii}^{k-2-j}\widetilde{H}_{i\ell}^2,$$

and thus

$$\psi''(z_{11})=k\sum_{i,\ell}\widetilde{H}_{i\ell}^2\sum_{j=0}^{k-2}\lambda_\ell^j\lambda_i^{k-2-j}=k\left(\sum_{i\neq\ell}\frac{\Lambda_{\ell\ell}^{k-1}-\Lambda_{ii}^{k-1}}{\Lambda_{\ell\ell}-\Lambda_{ii}}\widetilde{H}_{i\ell}^2+\sum_{i=1}^n\Lambda_{ii}^{k-2}\widetilde{H}_{ii}^2\right)\geqslant 0\quad(k\text{ is even}).$$

Hence, ψ is convex in z_{11}, and we complete the proof of (6).

□

References

Binbin Chen and Guangming Pan. Convergence of the largest eigenvalue of normalized sample covariance matrices when p and n both tend to infinity with their ratio converging to zero. *Bernoulli*, 18(4), 2012. doi: 10.3150/11-bej381. (cited on pages 1 and 2)

Jiaxin Qiu, Zeng Li, and Jianfeng Yao. Asymptotic normality for eigenvalue statistics of a general sample covariance matrix when $p/n\to\infty$ and applications. *The Annals of Statistics*, 51(3):1427–1451, 2023. doi: 10.1214/23-AOS2300. (cited on page 1)

Yongquan Yin, Zhidong Bai, and Pathak R Krishnaiah. On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory and Related Fields*, 78:509–521, 1988. doi: 10.1007/BF00353874. (cited on page 4)