

Elsevier L^AT_EX template[☆]

Elsevier¹

Radarweg 29, Amsterdam

Elsevier Inc^{a,b}, Global Customer Service^{b,}*

^a1600 John F Kennedy Boulevard, Philadelphia

^b360 Park Avenue South, New York

Abstract

This template helps you to create a properly formatted L^AT_EX manuscript.

Keywords: `elsarticle.cls`, L^AT_EX, Elsevier, template

2010 MSC: 00-01, 99-00

1. model

Consider i.i.d. random sample $X_1, \dots, X_n \in \mathbb{R}^p$ which has means $\mu = (\mu_1, \dots, \mu_p)^T$ and covariance matrix Σ . We consider testing the following high-dimensional hypothesis:

$$H_0 : \mu = 0_p \quad \text{versus} \quad H_1 : \mu \neq 0_p. \quad (1)$$

We assume, like chen qin and bai, the following multivariate model:

$$X_i = \mu + \Gamma Z_i \quad \text{for } i = 1, \dots, n, \quad (2)$$

where Γ is a $p \times m$ matrix for some $m \geq p$ such that $\Gamma_i \Gamma_i^T = \Sigma_i$ and $\{Z_i\}_{i=1}^n$ are m -variate i.i.d. random vectors satisfying $E(Z_i) = 0$ and $\text{Var}(Z_i) = I_m$, the $m \times m$ identity matrix. Write $Z_i = (z_{i1}, \dots, z_{im})^T$, we assume $E(z_{ij}^4) = 3 + \Delta <$

[☆]Fully documented templates are available in the elsarticle package on CTAN.

^{*}Corresponding author

Email address: support@elsevier.com (Global Customer Service)

URL: www.elsevier.com (Elsevier Inc)

¹Since 1880.

∞ and

$$\mathbb{E}(z_{il_1}^{\alpha_1} z_{il_2}^{\alpha_2} \cdots z_{il_q}^{\alpha_q}) = \mathbb{E}(z_{il_1}^{\alpha_1}) \mathbb{E}(z_{il_2}^{\alpha_2}) \cdots \mathbb{E}(z_{il_q}^{\alpha_q}) \quad (3)$$

for a positive integer q such that $\sum_{l=1}^q \alpha_l \leq 8$ and $l_1 \neq l_2 \neq \cdots \neq l_q$.

Consider the test statistic

$$T = \|\bar{X}\|^2. \quad (4)$$

Statistics like this are studied in some high dimensional mean test literature. Most of existing papers determined the critical value by asymptotic distribution, which is not exact. Randomization method have the advantages that the test is exact. However, as far as we know, there's no existing work give a theoretical justification of the power behavior of randomization method.

Randomization. For a test statistic $T(X_1, \dots, X_n)$. Suppose $\epsilon_1, \dots, \epsilon_n$ are i.i.d. Rademacher variables ($\Pr(\epsilon_i = 1) = \Pr(\epsilon_i = -1) = 1/2$) which are independent of the data. The critical value is defined as the $1 - \alpha$ quantile of the conditional distribution

$$\mathcal{L}(T(\epsilon_1 X_1, \dots, \epsilon_i X_i, \dots, \epsilon_n X_n) | X_1, \dots, X_n). \quad (5)$$

It's easy to see that the randomization test based on $\|\bar{X}\|^2$, $\|\bar{X}\|^2 - \frac{1}{n}S$ (S is the sample covariance matrix) and $\sum_{j < i} X_i^T X_j$ are equivalent. In what follows, we will consider

$$T_2(X_1, \dots, X_n) = \sum_{j < i} X_i^T X_j. \quad (6)$$

Then $T(\epsilon_1 X_1, \dots, \epsilon_i X_i, \dots, \epsilon_n X_n) = \sum_{j < i} X_i^T X_j \epsilon_i \epsilon_j$.

CLT for quadratic form of Rademacher variables. We will study the quadratic form of Rademacher variables. Let $\epsilon_1, \dots, \epsilon_n$ be indepent Rademacher variables. Consider quadratic form $W_n = \sum_{1 \leq j < i \leq n} a_{ij} \epsilon_i \epsilon_j$, where $\{a_{ij}\}$ are nonrandom numbers. Here $\{\epsilon_i\}$ and $\{a_{ij}\}$ may depend on n , a parameter we suppress. Obviously, $\mathbb{E}(W_n) = 0$ and $\text{Var}(W_n) = \sum_{1 \leq j < i \leq n} a_{ij}^2$.

Proposition 1. *A sufficient condition for*

$$\frac{W_n}{\sqrt{\sum_{1 \leq j < i \leq n} a_{ij}^2}} \xrightarrow{\mathcal{L}} N(0, 1) \quad (7)$$

is that

$$\sum_{j < k} \left(\sum_{i: i > k} a_{ij} a_{ik} \right)^2 + \sum_{j < i} a_{ij}^4 + \sum_{j < k < i} a_{ij}^2 a_{ik}^2 = o\left(\left(\sum_{j < i} a_{ij}^2\right)^2\right). \quad (8)$$

Proof. Define $U_{in} = \epsilon_i \sum_{j=1}^{i-1} a_{ij} \epsilon_j$, $i = 2, \dots, n$, and $\mathcal{F}_{in} = \sigma\{\epsilon_1, \dots, \epsilon_i\}$, $i = 1, \dots, n$. Now $W_n = \sum_{i=2}^n U_{in}$ and $\{U_{in}\}$ is a martingale difference array with respect to $\{\mathcal{F}_{in}\}$. To prove the proposition, we shall verify two conditions (See David Pollard's book):

$$\frac{\sum_{i=2}^n \mathbb{E}(U_{in}^2 | \mathcal{F}_{i-1,n})}{\sum_{1 \leq j < i \leq n} a_{ij}^2} \xrightarrow{P} 1, \quad (9)$$

and

$$\frac{\sum_{i=2}^n \mathbb{E}(U_{in}^2 \{U_{in}^2 > \epsilon \sum_{1 \leq j < i \leq n} a_{ij}^2\} | \mathcal{F}_{i-1,n})}{\sum_{1 \leq j < i \leq n} a_{ij}^2} \xrightarrow{P} 0, \quad (10)$$

for every $\epsilon > 0$.

Proof of (9). Since $\mathbb{E}(U_{in}^2 | \mathcal{F}_{i-1,n}) = \left(\sum_{j=1}^{i-1} a_{ij} \epsilon_j\right)^2$, we have

$$\begin{aligned} \sum_{i=2}^n \mathbb{E}(U_{in}^2 | \mathcal{F}_{i-1,n}) &= \sum_{i=2}^n \left(\sum_{j=1}^{i-1} a_{ij} \epsilon_j \right)^2 \\ &= \sum_{i=2}^n \left(\sum_{j=1}^{i-1} a_{ij}^2 + 2 \sum_{j,k: j < k < i} a_{ij} a_{ik} \epsilon_j \epsilon_k \right) \\ &= \sum_{i=2}^n \sum_{j=1}^{i-1} a_{ij}^2 + 2 \sum_{j < k < i} a_{ij} a_{ik} \epsilon_j \epsilon_k. \end{aligned}$$

But

$$\begin{aligned} \mathbb{E}\left(\sum_{j < k < i} a_{ij} a_{ik} \epsilon_j \epsilon_k\right)^2 &= \mathbb{E}\left(\sum_{j < k} \left(\sum_{i: i > k} a_{ij} a_{ik}\right) \epsilon_j \epsilon_k\right)^2 \\ &= \sum_{j < k} \left(\sum_{i: i > k} a_{ij} a_{ik}\right)^2 \\ &= o\left(\left(\sum_{j < i} a_{ij}^2\right)^2\right), \end{aligned}$$

15 where the last equality holds by assumption. Hence (9) holds.

Proof of (10). By Markov's inequality, we only need to prove

$$\frac{\sum_{i=2}^n \mathbb{E}(U_{in}^4 | \mathcal{F}_{i-1,n})}{\left(\sum_{1 \leq j < i \leq n} a_{ij}^2\right)^2} \xrightarrow{P} 0. \quad (11)$$

Since the relevant random variables are all positive, we only need to prove (11) converges to 0 in mean. But

$$\begin{aligned}
\sum_{i=2}^n \mathbb{E} U_{in}^4 &= \sum_{i=2}^n \mathbb{E} \left(\sum_{j:j < i} a_{ij} \epsilon_j \right)^4 \\
&= \sum_{i=2}^n \mathbb{E} \left(\sum_{j:j < i} a_{ij}^2 + 2 \sum_{j,k:j < k < i} a_{ij} a_{ik} \epsilon_j \epsilon_k \right)^2 \\
&= \sum_{i=2}^n \left(\left(\sum_{j:j < i} a_{ij}^2 \right)^2 + 4 \mathbb{E} \left(\sum_{j,k:j < k < i} a_{ij} a_{ik} \epsilon_j \epsilon_k \right)^2 \right) \\
&= \sum_{i=2}^n \left(\sum_{j:j < i} a_{ij}^4 + 6 \sum_{j,k:j < k < i} a_{ij}^2 a_{ik}^2 \right) \\
&= \sum_{j < i} a_{ij}^4 + 6 \sum_{j < k < i} a_{ij}^2 a_{ik}^2 \\
&= o \left(\left(\sum_{j < i} a_{ij}^2 \right)^2 \right),
\end{aligned}$$

where the last equality holds by assumption. Hence (10) holds. \square

2. Asymptotic normality

Lemma 1. Suppose $A = (a_{ij})$ is an $m \times m$ positive semi-definite matrix, then

$$\mathbb{E} (Z_i^T A Z_i)^2 \asymp (\text{tr}(A))^2 + \text{tr} A^2 \quad (12)$$

Proof. Notice that

$$\begin{aligned}
(Z_i^T A Z_i)^2 &= \left(\sum_{j=1}^m a_{jj} z_{ij}^2 + 2 \sum_{k < j} a_{jk} z_{ij} z_{ik} \right)^2 \\
&= \left(\sum_{j=1}^m a_{jj} z_{ij}^2 \right)^2 + 4 \left(\sum_{j=1}^m a_{jj} z_{ij}^2 \right) \left(\sum_{k < j} a_{jk} z_{ij} z_{ik} \right) + 4 \left(\sum_{k < j} a_{jk} z_{ij} z_{ik} \right)^2 \\
&= \sum_{j=1}^m a_{jj}^2 z_{ij}^4 + 2 \sum_{k < j} a_{jj} a_{kk} z_{ij}^2 z_{ik}^2 + 4 \left(\sum_{j=1}^m a_{jj} z_{ij}^2 \right) \left(\sum_{k < j} a_{jk} z_{ij} z_{ik} \right) \\
&\quad + 4 \left(\sum_{k < j} a_{jk}^2 z_{ij}^2 z_{ik}^2 + \sum_{k < j, l < \alpha: \text{card}(\{i,j\} \cap \{l,\alpha\}) < 2} a_{jk} a_{\alpha l} z_{ij} z_{ik} z_{i\alpha} z_{il} \right)
\end{aligned} \quad (13)$$

Hence

$$\begin{aligned}
\mathbb{E}(Z_i^T A Z_i)^2 &= \sum_{j=1}^n a_{jj}^2 \mathbb{E} z_{ij}^4 + 2 \sum_{k < j} a_{jj} a_{kk} \mathbb{E}(z_{ij}^2 z_{ik}^2) + 4 \sum_{k < j} a_{jk}^2 \mathbb{E}(z_{ij}^2 z_{ik}^2) \\
&\asymp \sum_{j=1}^n \sum_{k=1}^n a_{jj} a_{kk} + \sum_{j=1}^n \sum_{k=1}^n a_{jk}^2 = (\text{tr}(A))^2 + \text{tr} A^2
\end{aligned} \tag{14}$$

□

Lemma 2. *Suppose condition (32) holds, then*

$$\frac{\sum_{j < i} (X_i^T X_j)^2}{\frac{n(n-1)}{2} \text{tr}(\Sigma + \mu \mu^T)^2} \xrightarrow{P} 1. \tag{15}$$

Proof.

$$\begin{aligned}
\mathbb{E}(X_i^T X_j)^2 &= \mathbb{E}(X_i^T X_j X_j^T X_i) = \mathbb{E}(X_i^T (\Sigma + \mu \mu^T) X_i) \\
&= \mathbb{E} \text{tr}((\Sigma + \mu \mu^T) X_i X_i^T) = \text{tr}(\Sigma + \mu \mu^T)^2
\end{aligned} \tag{16}$$

Hence

$$\mathbb{E} \sum_{j < i} (X_i^T X_j)^2 = \frac{n(n-1)}{2} \text{tr}(\Sigma + \mu \mu^T)^2 \tag{17}$$

So we only need to consider the variance. According to $\text{card}(\{i, j\} \cap \{k, l\}) = 0, 1, 2$, we have

$$\begin{aligned}
\left(\sum_{j < i} (X_i^T X_j)^2 \right)^2 &= \sum_{j < i} (X_i^T X_j)^4 + \sum_{j < i, k < l: \{i, j\} \cap \{k, l\} = \emptyset} (X_i^T X_j)^2 (X_k^T X_l)^2 \\
&+ 2 \sum_{j < i < k} ((X_i^T X_j)^2 (X_k^T X_i)^2 + (X_i^T X_j)^2 (X_k^T X_j)^2 + (X_k^T X_j)^2 (X_k^T X_i)^2)
\end{aligned} \tag{18}$$

$$\begin{aligned}
(X_i^T X_j)^4 &= (Z_i^T \Gamma^T \Gamma Z_j + \mu^T \Gamma Z_i + \mu^T \Gamma Z_j + \mu^T \mu)^4 \\
&\leq 64 ((Z_i^T \Gamma^T \Gamma Z_j)^4 + (\mu^T \Gamma Z_i)^4 + (\mu^T \Gamma Z_j)^4 + (\mu^T \mu)^4)
\end{aligned} \tag{19}$$

Note that by lemma 1, we have

$$\mathbb{E}(\mu^T \Gamma Z_i)^4 = \mathbb{E}(Z_i^T \Gamma^T \mu \mu^T \Gamma Z_i)^2 \asymp (\mu^T \Sigma \mu)^2 \leq \lambda_{\max}^2(\Sigma) (\mu^T \mu)^2. \tag{20}$$

and apply lemma 1, we get

$$\begin{aligned}
\mathbb{E}(Z_i^T \Gamma^T \Gamma Z_j)^4 &= \mathbb{E}(Z_i^T \Gamma^T \Gamma Z_j Z_j^T \Gamma^T \Gamma Z_i)^2 \\
&= \mathbb{E} \mathbb{E}((Z_i^T \Gamma^T \Gamma Z_j Z_j^T \Gamma^T \Gamma Z_i)^2 | Z_j) \\
&\asymp \mathbb{E}(Z_j^T \Gamma^T \Sigma \Gamma Z_j)^2 \\
&\asymp (\text{tr} \Sigma^2)^2 + \text{tr} \Sigma^4 \\
&\leq (\text{tr} \Sigma^2)^2 + \lambda_{\max}^2(\Sigma) \text{tr} \Sigma^2.
\end{aligned} \tag{21}$$

By condition (32), we have

$$\mathbb{E}(X_i^T X_j)^4 = O\left((\text{tr}(\Sigma + \mu \mu^T))^2\right) \tag{22}$$

Similarly, we have

$$\begin{aligned}
&\mathbb{E}(X_i^T X_j)^2 (X_k^T X_i)^2 \\
&= \mathbb{E} \mathbb{E}((X_i^T X_j)^2 (X_k^T X_i)^2 | X_i) \\
&= \mathbb{E}(X_i^T (\Sigma + \mu \mu^T) X_i)^2 \\
&= \mathbb{E}(Z_i^T \Gamma^T (\Sigma + \mu \mu^T) \Gamma Z_i + 2\mu^T (\Sigma + \mu \mu^T) \Gamma Z_i + \mu \Sigma \mu + (\mu^T \mu)^2)^2 \\
&\leq 4\mathbb{E}(Z_i^T \Gamma^T (\Sigma + \mu \mu^T) \Gamma Z_i)^2 + 16\mathbb{E}(\mu^T (\Sigma + \mu \mu^T) \Gamma Z_i)^2 + 4(\mu \Sigma \mu)^2 + 4(\mu^T \mu)^4 \\
&\asymp (\text{tr}(\Gamma^T (\Sigma + \mu \mu^T) \Gamma))^2 + \text{tr}(\Gamma^T (\Sigma + \mu \mu^T) \Sigma (\Sigma + \mu \mu^T) \Gamma) \\
&\quad + \mu^T (\Sigma + \mu \mu^T) \Sigma (\Sigma + \mu \mu^T) \mu + (\mu \Sigma \mu)^2 + (\mu^T \mu)^4 \\
&= \text{tr} \Sigma^4 + (\mu^T \mu)^4 + (\text{tr} \Sigma^2)^2 + 2\text{tr} \Sigma^2 (\mu^T \Sigma \mu) + 3(\mu^T \Sigma \mu)^2 \\
&\quad + 3\mu^T \Sigma^3 \mu + 2(\mu^T \Sigma^2 \mu)(\mu^T \mu) + (\mu^T \Sigma \mu)(\mu^T \mu)^2 \\
&= O(1)((\text{tr} \Sigma^2)^2 + (\mu^T \mu)^4) \\
&= O\left((\text{tr}(\Sigma + \mu \mu^T))^2\right)
\end{aligned} \tag{23}$$

The last 2 equality follows by the fact $\mu^T \Sigma^i \mu \leq \lambda_{\max}^i(\Sigma) \mu^T \mu$ ($i = 1, 2, \dots$),
the condition (32) and the inequality $ab \leq a^p/p + b^q/p$ for $1/p + 1/q = 1$ and
 $p, q > 0$.

In (18), there are $n(n-1)/2$, $n(n-1)(n-2)(n-3)/4$ and $n(n-1)(n-2)/6$

terms in each summation. Hence

$$\begin{aligned} \mathbb{E}\left(\sum_{j < i} (X_i^T X_j)^2\right)^2 &= \frac{n(n-1)(n-2)(n-3)}{4} (\text{tr}(\Sigma + \mu\mu^T)^2)^2 \\ &\quad + O(1)\left(\frac{n(n-1)}{2} + n(n-1)(n-2)\right) (\text{tr}(\Sigma + \mu\mu^T)^2)^2 \end{aligned} \quad (24)$$

Hence $\text{Var}(\sum_{j < i} (X_i^T X_j)^2) = o\left(\left(\mathbb{E}(\sum_{j < i} (X_i^T X_j)^2)\right)^2\right)$, which prove the lemma. \square

Lemma 3. *Suppose (32) holds. Assume*

$$\mu^T \mu = o(\sqrt{\text{tr} \Sigma^2}), \quad (25)$$

then

$$\sum_{j < k} \left(\sum_{i: i > k} X_i^T X_j X_i^T X_k \right)^2 = o_P \left(\left(\frac{n(n-1)}{2} \text{tr}(\Sigma + \mu\mu^T)^2 \right)^2 \right) \quad (26)$$

$$\sum_{j < k} (X_i^T X_j)^4 = o_P \left(\left(\frac{n(n-1)}{2} \text{tr}(\Sigma + \mu\mu^T)^2 \right)^2 \right) \quad (27)$$

$$\sum_{j < k < i} (X_i^T X_j)^2 (X_i^T X_k)^2 = o_P \left(\left(\frac{n(n-1)}{2} \text{tr}(\Sigma + \mu\mu^T)^2 \right)^2 \right) \quad (28)$$

Proof.

$$\begin{aligned} &\mathbb{E} \sum_{j < k} \left(\sum_{i: i > k} X_i^T X_j X_i^T X_k \right)^2 \\ &= \mathbb{E} \sum_{j < k} \left(\sum_{i: i > k} (X_i^T X_j)^2 (X_i^T X_k)^2 + 2 \sum_{i_1, i_2: i_1 > i_2 > k} X_{i_1}^T X_j X_{i_1}^T X_k X_{i_2}^T X_j X_{i_2}^T X_k \right). \end{aligned} \quad (29)$$

By (23), we have

$$\mathbb{E} \sum_{j < k < i} (X_i^T X_j)^2 (X_i^T X_k)^2 = O(n^3) (\text{tr}(\Sigma + \mu\mu^T)^2)^2 \quad (30)$$

But

$$\begin{aligned}
& \mathbb{E} \sum_{j < k < i_2 < i_1} X_{i_1}^T X_j X_{i_1}^T X_k X_{i_2}^T X_j X_{i_2}^T X_k \\
&= \frac{n(n-1)(n-2)(n-3)}{6} \text{tr}(\Sigma + \mu\mu^T)^4 \\
&\leq \frac{n(n-1)(n-2)(n-3)}{6} 8(\text{tr}(\Sigma)^4 + (\mu^T \mu)^4) \\
&\leq O(n^4)(\lambda_{\max}^2(\Sigma) \text{tr}(\Sigma)^2 + (\mu^T \mu)^4) \\
&= o\left(n^4(\text{tr}\Sigma^2)^2\right).
\end{aligned} \tag{31}$$

It follows that (26) holds. (27) and (28) can be proved similarly by (23) and (22).

25

□

As Janssen's paper, let d denote any metric on the set of probability measures $\mathcal{M}_1(\mathbb{R})$ on \mathbb{R} such that convergence in $(\mathcal{M}_1(\mathbb{R}), d)$ is equivalent to weak convergence.

Theorem 1. *Suppose*

$$\frac{\lambda_{\max}^2(\Sigma)}{\text{tr}\Sigma^2} \rightarrow 0, \tag{32}$$

and

$$\mu^T \mu = o(\sqrt{\text{tr}\Sigma^2}), \tag{33}$$

then

$$d\left(\mathcal{L}\left(\frac{T_2(\epsilon_1 X_1, \dots, \epsilon_i X_i, \dots, \epsilon_n X_n)}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} \middle| X_1, \dots, X_n\right), N(0, 1)\right) \xrightarrow{P} 0. \tag{34}$$

Proof. Similar to the proof of Chen's theorem and by Lemma 2, we have

$$\frac{T_2(X_1, \dots, X_n) - \frac{n(n-1)}{2} \mu^T \mu}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} \xrightarrow{\mathcal{L}} N(0, 1). \tag{35}$$

For every subsequence of $\{n\}$, there is a further subsequence $\{n(k)\}$ along which (26), (27) and (28) holds almost surely. By Proposition 1, we have

$$\mathcal{L}\left(\frac{T_2(\epsilon_1 X_1, \dots, \epsilon_i X_i, \dots, \epsilon_n X_n)}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} \middle| X_1, \dots, X_n\right) \xrightarrow{\mathcal{L}} N(0, 1) \tag{36}$$

almost surely. Denote by ξ_α^* the $1 - \alpha$ quantile of the above conditional distribution, then along $\{n(m)\}$ we have $\xi_\alpha^* \rightarrow \Phi(1 - \alpha)$ almost surely, where $\Phi(\cdot)$

is the CDF of standard normal distribution. Then $\xi_\alpha^* \xrightarrow{P} \Phi(1 - \alpha)$. Now the asymptotic power function can be derived by Slutsky's theorem

$$\begin{aligned}
& \Pr \left(\frac{T_2(X_1, \dots, X_n)}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} > \xi_\alpha^* \right) \\
&= \Pr \left(\frac{T_2(X_1, \dots, X_n) - \frac{n(n-1)}{2} \mu^T \mu}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} > \xi_\alpha^* - \frac{\frac{n(n-1)}{2} \mu^T \mu}{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}} \right) \\
&= \Pr \left(\frac{T_2(X_1, \dots, X_n) - \frac{n(n-1)}{2} \mu^T \mu}{\sqrt{\frac{n(n-1)}{2} \text{tr} \Sigma^2}} - \frac{\sqrt{\sum_{1 \leq j < i \leq n} (X_i^T X_j)^2}}{\sqrt{\frac{n(n-1)}{2} \text{tr} \Sigma^2}} \xi_\alpha^* > -\frac{\sqrt{n(n-1)} \mu^T \mu}{\sqrt{2 \text{tr} \Sigma^2}} \right) \\
&= \Pr \left(N(0, 1) - \Phi(1 - \alpha) > -\frac{\sqrt{n(n-1)} \mu^T \mu}{\sqrt{2 \text{tr} \Sigma^2}} \right) + o(1) \\
&= \Phi(-\Phi(1 - \alpha) + \frac{\sqrt{n(n-1)} \mu^T \mu}{\sqrt{2 \text{tr} \Sigma^2}}) + o(1),
\end{aligned} \tag{37}$$

where the last two equality holds because an exersize of durrett.

30

□

References