A Simple Generalization of a Result for Random Matrices with Independent Sub-Gaussian Rows

Namrata Vaswani and Seyedehsara Nayer

Abstract—In this short note, we give a very simple but useful generalization of a result of Vershynin (Theorem 5.39 of [1]) for a random matrix with independent sub-Gaussian rows. We also explain with an example where our generalization is useful.

I. INTRODUCTION

In this note, we obtain a generalization of a result of Vershynin, Theorem 5.39 of [1]. This result bounds the minimum and maximum singular values of an $N \times n$ matrix $W$ with mutually independent, sub-Gaussian, and isotropic rows. We use $\| \cdot \|$ to denote the $l_2$ norm of a vector or the induced $l_2$ norm of a matrix, and we use $\text{tr}$ to denote matrix or vector transpose. Let $W = [w_1, w_2, \ldots, w_N]'$. Thus, $w_j$ is its $j$-th row. As explained in [1, Section 5.2], “isotropic” means that $E[x|p]^{1/p} \leq K_g \sqrt{p}$ for all integers $p \geq 1$. The smallest such $K_g$ is referred to as the sub-Gaussian norm of $x$, denoted $\|x\|_{\psi_2}$. Thus, $\|x\|_{\psi_2} = \sup_{p \geq 1} p^{-1/2} E|x|^{1/p}$. A sub-Gaussian random vector $x$, is one for which, for all unit norm vectors $z$, $x'z$ is sub-Gaussian. Also, its sub-Gaussian norm, $\|x\|_{\psi_2} = \sup_{\|z\|=1} \|x'z\|_{\psi_2}$.

Let $K$ denote the maximum of the sub-Gaussian norms of the rows of $W$. Theorem 5.39 of [1] shows that, for any $t > 0$, with probability at least $1 - \exp(-c_K t^2)$, the minimum singular value of $W$ is more than $\sqrt{N} - (C_K \sqrt{n} + t)$ and the maximum is less than $\sqrt{N} + (C_K \sqrt{n} + t)$. Here $C_K$ and $c_K$ are numerical constants that depend only on $K$. These bounds are obtained by bounding the deviation of $\frac{1}{N} W' W$ from its expected value, which is equal to $I$.

Our generalization of this result does two extra things. First, it bounds $\|W'W - E[W'W]\|$, even when the different rows of $W$ do not have the same second moment matrix. Second, it states a separate result that bounds $\|Wz\|^2$ for one specific vector $z$. This bound clearly holds with much higher probability than the bound on $\|W'W - E[W'W]\|$. The proof approach for getting our result is the same as that used to get Theorem 5.39 of [1]. Thus, our generalization would be obvious to a reader who understands the proof of that result. However, it is a useful addition to the literature for readers who would like to just use results from [1] in their work, without having to understand all their proof techniques.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$

for a numerical constant $c$.

Here, and throughout the paper, the letter $c$ is reused to denote different numerical constants.

Proof: The proof is given in the Supplementary Material.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$

for a numerical constant $c$.

Here, and throughout the paper, the letter $c$ is reused to denote different numerical constants.

Proof: The proof is given in the Supplementary Material.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$

for a numerical constant $c$.

Here, and throughout the paper, the letter $c$ is reused to denote different numerical constants.

Proof: The proof is given in the Supplementary Material.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$

for a numerical constant $c$.

Here, and throughout the paper, the letter $c$ is reused to denote different numerical constants.

Proof: The proof is given in the Supplementary Material.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$

for a numerical constant $c$.

Here, and throughout the paper, the letter $c$ is reused to denote different numerical constants.

Proof: The proof is given in the Supplementary Material.

II. OUR RESULT

Theorem 2.1. Suppose that $w_j$, $j = 1, 2, \ldots, N$, are $n$-length independent, sub-Gaussian random vectors with sub-Gaussian norms bounded by $K$. Let

$$D := \frac{1}{N} \sum_{j=1}^{N} w_j w_j' - \frac{1}{N} \sum_{j=1}^{N} E[w_j w_j']. $$

For an $\varepsilon > 0$, 1) for a given vector $z$, with probability (w.p.) at least $1 - 2 \exp(-c \min(\varepsilon, \varepsilon^2) N)$,

$$|z'Dz| \leq 4 \varepsilon K^2 \|z\|^2;$$

2) w.p. at least $1 - 2 \exp(n \log 9 - c \min(\varepsilon, \varepsilon^2) N)$,

$$\|D\| \leq 4 \varepsilon K^2.$$
with \( i = 1, 2, \ldots, m \) and \( k = 1, 2, \ldots, q \); and \( n \)-length deterministic vectors \( x_k, k = 1, 2, \ldots, q \). Assume that \( q \leq n^2 \). Consider bounding

\[
 b := \left| \frac{1}{m} \sum_{i=1}^{m} (a_{i,k}' x_k)^2 - \|x_k\|^2 \right|
\]

By applying item 1 of Theorem 2.1 with \( N = m \) and \( w_j = a_{j,k} \), \( b_k \leq \varepsilon \|x_k\|^2 \) w.p. at least \( 1 - \exp(-\varepsilon c \varepsilon^2 m) \). Such a bound holds for all \( k = 1, 2, \ldots, q \) w.p. at least \( 1 - q \exp(-\varepsilon c \varepsilon^2 m) \). Thus, to ensure that this bound holds w.p. at least \( 1 - 1/poly(n) \), we need \( m \geq \frac{\varepsilon (\log n + \log q)}{\varepsilon^2 c^2} \) since \( q \leq n^2 \). Here \( poly(n) \) means polynomial in \( n \).

On the other hand, to apply [1] Theorem 5.39, we first need to upper bound the \( b_k \)'s as

\[
 b_k = |x_k' \left( \frac{1}{m} \sum_{i=1}^{m} a_{i,k} a_{i,k}' - I \right) x_k |
\]

\[
 \leq \|x_k\|^2 \left| \frac{1}{m} \sum_{i=1}^{m} a_{i,k} a_{i,k}' - I \right|
\]

With this, we can get the same bound as above on the \( b_k \)'s by applying [1] Theorem 5.39 with \( t = \sqrt{m^2 K^2 \varepsilon - C_K \sqrt{n}} \) (or, equivalently, by applying item 2 of Theorem 2.1 above). But the bound would hold with probability lower bounded by \( 1 - \exp(n \log 9 - c \varepsilon^2 m) \). For a given \( m \), this is a much smaller probability. Said another way, one would need \( m \geq \frac{9}{\varepsilon^2} \) for the probability to be high enough (at least \( 1 - 1/poly(n) \)). This is a much larger lower bound on \( m \) than the earlier one.

To see an application of item 2 of Theorem 2.1 consider bounding

\[
 
 b := \left| \frac{1}{mq} \sum_{k=1}^{q} \sum_{i=1}^{m} a_{i,k} a_{i,k}' f_k^2 - \frac{1}{q} \sum_{k=1}^{q} f_k^2 \right|
\]

where \( f_k \)'s are scalars. By conditioning on the \( f_k \)'s, we can apply item 2 of Theorem 2.1 on all the \( N = mq \) vectors \( (a_{i,k}, f_k) \) to conclude that \( b \leq \varepsilon_2 \max_k f_k^2 \) w.p. at least \( 1 - 2 \exp(n \log 9 - c \varepsilon^2 mq) \). Thus, the bound holds w.p. at least \( 1 - 1/poly(n) \) if \( m \geq \frac{cn}{\varepsilon^2} \).

Observe that the \( a_{i,k}'s \) are isotropic independent sub-Gaussian vectors but \( a_{i,k} f_k \)'s are not. In fact, \( \mathbb{E}[a_{i,k} a_{i,k}' f_k^2] = f_k^2 \) and hence the vectors \( a_{i,k} f_k \) also do not have the same second moment matrix for all \( k, i \). As a result, we cannot apply Theorem 3.59 or Remark 5.40 of [1] to bound \( \bar{b} \) if we want to average over all the \( mq \) vectors. To apply one of these, we first need to upper bound \( \bar{b} \) as

\[
 \bar{b} \leq \frac{1}{q} \sum_{k=1}^{q} \left| \frac{1}{m} \sum_{i=1}^{m} a_{i,k} a_{i,k}' - I \right| f_k^2
\]

Now using [1] Theorem 5.39, we get \( \bar{b} \leq \varepsilon_2 \frac{1}{q} \sum_{k=1}^{q} f_k^2 \leq \varepsilon_2 \max_k f_k^2 \) w.p. at least \( 1 - 2 \exp(n \log 9 - c \varepsilon^2 m) \). Observe that \( mq \) is replaced by \( m \) in the probability now. Thus, to get the probability to be high enough (at least \( 1 - \frac{1}{poly(n)} \)) we will need \( m \geq \frac{cn}{\varepsilon^2} \) which is, once again, a much larger lower bound than what we got by applying item 2 of Theorem 2.1

To understand the context, in [2], \( m \) is the sample complexity required for the initialization step of low-rank phase retrieval to get an estimate of the low-rank matrix \( X := [x_1, x_2, \ldots, x_q] \) that is within a relative error \( c\varepsilon \) of the true \( X \) with probability at least \( 1 - 1/poly(n) \). If we directly use the result from [1], we will need \( m \geq cn/\varepsilon^2 \), where as if we use Theorem 2.1 we can get a lower bound that is smaller than \( cn \) (when \( q \) is large enough).

### III. Conclusions

We proved a simple generalization of a result of Vershynin [1] for random matrices with independent, sub-Gaussian rows. We should mention that the first claim of Theorem 2.1 can be further generalized for two different vectors \( z_1 \) and \( z_2 \) as follows: with the same probability, \( |z_1' D z_2| \leq 4\varepsilon^2 K^2 (\|z_1\|^2 + \|z_2\|^2) \). This follows because, for two sub-Gaussian scalars, \( x, y, xy \) is sub-exponential with sub-exponential norm bounded by \( c(\|x\|_2^2 + \|y\|_2^2) \) [4].

### REFERENCES

[1] R. Vershynin, “Introduction to the non-asymptotic analysis of random matrices,” *Compressed sensing*, pp. 210–268, 2012.

[2] N. Vaswani, S. Nayer, and Y. Eldar, “Low rank phase retrieval,” *Under Revision for IEEE Trans. Sig. Proc.*, 2016, also at arXiv:1608.04141.

[3] S. Nayer, N. Vaswani, and Y. Eldar, “Low rank phase retrieval,” in *submitted to ICASSP*, 2017.

[4] Y. Q. Vu, I. Cho, J. Lei, and K. Rohe, “Fantope projection and selection,” in *Adv. Neural Info. Proc. Sys. (NIPS)*, 2013.

### APPENDIX

**Proof of our result**

**A. Preliminaries**

As explained in [1], nets are a convenient means to discretize compact metric spaces. The following definition is [1] Definition 5.1 for the unit sphere. For an \( \varepsilon > 0 \), a subset \( N_\varepsilon \) of the unit sphere in \( \mathbb{R}^n \) is called an \( \varepsilon \)-net if, for every vector \( x \) in the unit sphere, there exists a vector \( y \in N_\varepsilon \) such that \( \|y - x\| \leq \varepsilon \).

The covering number of the unit sphere in \( \mathbb{R}^n \), is the minimal cardinality of an \( \varepsilon \)-net on it. In other words, it is the size of the smallest \( \varepsilon \)-net, \( N_\varepsilon \), on it.

**Fact 1.1.**

1. By Lemma 5.2 of [1], the covering number of the unit sphere in \( \mathbb{R}^n \) is upper bounded by \( (1 + \frac{2}{\varepsilon})^n \).

2. By Lemma 5.4 of [1], for a symmetric matrix, \( D \), \( \|D\| \leq \max_{x: \|x\|_1 = 1} \|x'Dx\| \leq \frac{1}{\sqrt{2}} \max_{x \in N_1/4} \|x'Dx\|. \)

Thus, if \( \varepsilon = 1/4 \), then \( \|D\| \leq 2 \max_{x \in N_{1/4}} \|x'Dx\| \) and the cardinality of the smallest such net is at most \( 9^n \).

A r.v. \( x \) is sub-exponential if the following holds: there exists a constant \( K_\varepsilon \) such that \( \mathbb{E}[|x|^p]^{1/p} \leq K_\varepsilon p \) for all integers \( p \geq 1 \); the smallest such \( K_\varepsilon \) is referred to as the sub-exponential norm of \( x \), denoted \( \|x\|_2 \) [1] Section 5.2.

The following facts will be used in our proof.

**Fact 1.2.**

1. If \( x \) is a sub-Gaussian random vector with sub-Gaussian norm \( K \), then for any vector \( z \), (i) \( x'z \) is sub-Gaussian with sub-Gaussian norm bounded by \( K \|z\| \); (ii) \( (x'z)^2 \)
is sub-exponential with sub-exponential norm bounded by $2K^2\|z\|^2$; and (iii) $(x'z)^2 - E[(x'z)^2]$ is centered (zero-mean), sub-exponential with sub-exponential norm bounded by $4K^2\|z\|^2$. This follows from the definition of a sub-Gaussian random vector; Lemma 5.14 and Remark 5.18 of [1].

2) By [1, Corollary 5.17], if $x_i, i = 1, 2, \ldots N$, are a set of independent, centered, sub-exponential r.v.'s with sub-exponential norm bounded by $K\|z\|_2$, then, for any $\varepsilon > 0$,

$$
\Pr \left( \left| \sum_{i=1}^{N} x_i \right| > \varepsilon K \varepsilon N \right) \leq 2 \exp(-c \min(\varepsilon, \varepsilon^2) N).
$$

3) If $x \sim \mathcal{N}(0, \tilde{\Lambda})$ with $\tilde{\Lambda}$ diagonal, then $x$ is sub-Gaussian with

$$
\|x\|_{\psi_2} \leq c \sqrt{\lambda_{\max}}.
$$

B. Proof of Theorem 2.1

The proof strategy is similar to that of Theorem 5.39 of [1]. By Fact 1.2, item 1, for each $j$, the r.v.s $w_j'z$ are sub-Gaussian with sub-Gaussian norm bounded by $K\|z\|$; $(w_j'z)^2$ are sub-exponential with sub-exponential norm bounded by $2K^2\|z\|^2$; and $(w_j'z)^2 - E[(w_j'z)^2] = z'(w_jw_j')z - z'(E[w_jw_j'])z$ are centered sub-exponential with sub-exponential norm bounded by $4K^2\|z\|^2$. Also, for different $j$'s, these are clearly mutually independent. Thus, by applying Fact 1.2 (Corollary 5.17 of [1]) with $K = 4K^2\|z\|^2$ we get the first part.

To prove the second part, let $\mathcal{N}_{1/4}$ denote a 1/4-th net on the unit sphere in $\mathbb{R}^n$. Let $D := \frac{1}{N} \sum_{j=1}^{N} (w_jw_j' - E[w_jw_j'])$. Then by Fact 1.1 (Lemma 5.4 of [1])

$$
\|D\| \leq 2 \max_{z \in \mathcal{N}_{1/4}} |z'Dz| \quad (1)
$$

Since $\mathcal{N}_{1/4}$ is a finite set of vectors, all we need to do now is to bound $|z'Dz|$ for a given vector $z$ followed by applying the union bound to bound its maximum over all $z \in \mathcal{N}_{1/4}$. The former has already been done in the first part. By Fact 1.1 (Lemma 5.2 of [1]), the cardinality of $\mathcal{N}_{1/4}$ is at most $9^n$.

Thus, using the first part, $\Pr \left( \max_{z \in \mathcal{N}_{1/4}} |z'Dz| \geq \frac{4K^2}{2} \right) \leq 9^n \cdot 2 \exp(-c \min(\varepsilon, \varepsilon^2) N) = 2 \exp(n \log 9 - c \varepsilon^2 N)$. By (1), we get the result.