

# 随机矩阵概率不等式类的证明

郑珂, 宋儒瑛

太原师范学院数学系, 山西 晋中  
Email: 1071980702@qq.com

收稿日期: 2021年6月19日; 录用日期: 2021年7月11日; 发布日期: 2021年7月22日

## 摘要

在证明随机矩阵满足限制等距性质以及估计限制等距常数时, Khintchine不等式、Hoeffding型不等式以及Bernstein型不等式都发挥着重要的作用。在实际应用当中, 这三种不等式根据随机变量类别的不同又有很多变化。本文重点补充证明了当随机变量为矩阵时的Khintchine不等式、Hoeffding型不等式以及Bernstein型不等式, 这些不等式在随机测量矩阵稀疏恢复证明过程中起到关键作用。

## 关键词

限制等距性质, 亚高斯随机变量, 标准高斯随机变量, Khintchine不等式, Hoeffding型不等式, Bernstein型不等式

# The Proof of Probability Inequality of Random Matrix

Ke Zheng, Ruying Song

Department of Mathematics, Taiyuan Normal University, Jinzhong Shanxi  
Email: 1071980702@qq.com

Received: Jun. 19<sup>th</sup>, 2021; accepted: Jul. 11<sup>th</sup>, 2021; published: Jul. 22<sup>nd</sup>, 2021

## Abstract

Khintchine's inequality, Hoeffding's inequality and Bernstein's inequality play an important role in proving that random matrix satisfies the property of restricted isometry and estimating the constant of restricted isometry. In practical application, these three inequalities have many changes according to the different types of random variables. In this paper, we mainly prove the Khintchine inequality, Hoeffding type inequality and Bernstein type inequality when the random variable is a matrix. These inequalities play a key role in the proof of sparse restoration of random

measurement matrix.

## Keywords

Restricted Isometric Property, Sub-Gaussian Random Variable, Standard Gaussian Random Variable, Khintchine Inequality, Hoeffding Type Inequality, Bernstein Type Inequality

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

压缩感知的模型如下  $y = Ax$ , 当测量矩阵是随机测量矩阵的时候, 满足 RIP 性质的概率需要足够大, 在证明过程中需要借助 Khintchine 不等式(见文献[1])、Hoeffding 型不等式以及 Bernstein 型不等式(见文献[2]), 这三种不等式有很多种, 后两种分类是根据随机变量决定, 随机变量可能是一个单独的变量, 可能是向量, 也可能是矩阵。

在文献[3]中介绍了 Khintchine 不等式, 文献[1]介绍了非对易 Khintchine 不等式, 读者可以参考。Khintchine 不等式提供了 Rademacher 和以及高斯和的界定, 之后又延伸到矩阵 Rademacher 和以及矩阵高斯和的 Khintchine 不等式, 本文对这两种不等式进行了证明。

文献[4]介绍了标量 Bernstein 不等式, 文献[5]介绍了矩阵 Bernstein 不等式, 在文献[3]中 8.5 节重点介绍了 Noncommutative Bernstein 不等式, 文献[2] [6] [7]中也有相关介绍, 其中随机变量是独立零均值自伴随机矩阵, 而次指数版本 Noncommutative Bernstein 不等式在文献[3]练习题当中。本文在前人对这两种不等式总结以及文献[3]相关证明的基础上对又 Noncommutative Bernstein 不等式的次指数版本进行了证明。

在证明 Noncommutative Bernstein 不等式的同时需要 Hoeffding 型不等式, 文献[3]重点介绍了 Rademacher 和的 Hoeffding 型不等式, 在练习题中引入了矩阵高斯和的 Hoeffding 型不等式, 本文将对这个不等式进行证明, 同时将证明一些其他定理用以辅助证明 Hoeffding 型不等式。

## 2. 矩阵高斯和的 Khintchine 不等式

在证明之前首先需要以下引理来辅助证明 Khintchine 不等式。

**引理 2.1:** 设  $g$  是一标准高斯变量,  $B \in \mathbb{C}^{d \times d}$  是一自伴矩阵。那么满足以下等式

$$E \exp(g\theta B) = \exp(\theta^2 B^2 / 2). \quad (1)$$

证明: 对于标准高斯随机变量  $g$ , 期望和方差满足以下条件

$$Eg = 0, Eg^2 = 1 \quad (2)$$

见文献[3], [8]。所以当标准高斯随机变量是奇数次幂时,

$$Eg^k = E(g)E(g^{k-1}) = 0 \times E(g^{k-1}) = 0. \quad (3)$$

我们将定理中等式左边按照泰勒展开式展开, 即

$$E \exp(g\theta B) = \sum_{k=0}^{\infty} \frac{E(g\theta B)^k}{k!} = \sum_{k=0}^{\infty} \frac{E(g)^k E(\theta B)^k}{k!}$$

我们将(3)代入上式, 可得

$$\begin{aligned} E \exp(g\theta B) &= \sum_{k=0}^{\infty} \frac{E(g\theta B)^k}{k!} = \sum_{k=0}^{\infty} \frac{E(g)^k E(\theta B)^k}{k!} \\ &= 1 + \sum_{k=1}^{\infty} \frac{E(g)^{2k} E(\theta B)^{2k}}{(2k)!!} + \sum_{k=1}^{\infty} \frac{E(g)^{2k-1} E(\theta B)^{2k-1}}{(2k-1)!!} \\ &= 1 + \sum_{k=1}^{\infty} \frac{E(g)^{2k} E(\theta B)^{2k}}{(2k)!!} = 1 + \sum_{k=1}^{\infty} \frac{1 \times (\theta B)^{2k}}{(2k)!!} \\ &= \sum_{k=0}^{\infty} \frac{(\theta B)^{2k}}{(2k)!!} = \sum_{k=0}^{\infty} \frac{(\theta B)^{2k}}{2^k k!} = \exp(\theta^2 B^2/2) \end{aligned}$$

等式结果得以证明。

以上引理证明时可以借鉴文献[9] [10]。

**引理 2.2:** 设  $\varepsilon$  是一 Rademacher 随机变量,  $B \in \mathbb{C}^{d \times d}$  是一自伴矩阵。那么满足以下等式

$$E \exp(\varepsilon \theta B) \leq \exp(\theta^2 B^2/2). \tag{4}$$

证明: 对于 Rademacher 随机变量  $\varepsilon$ , 期望和方差满足以下条件

$$E\varepsilon = 0, E\varepsilon^2 = 1 \tag{5}$$

我们将定理中等式左边按照泰勒展开式展开, 即

$$E \exp(\varepsilon \theta B) = \frac{1}{2} (\exp(\theta B) + \exp(-\theta B)) = \sum_{k=0}^{\infty} \frac{(\theta B)^{2k}}{(2k)!} \leq \sum_{k=0}^{\infty} \frac{(\theta B)^{2k}}{2^k k!} = \exp(\theta^2 B^2/2) \tag{6}$$

等式结果得以证明。

该引理摘自文献[3]。

**定理 2.1:** 在以上定理成立的条件下, 满足以下不等式

$$E \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \leq \sqrt{2 \ln(2d)} \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}^{1/2} \tag{7}$$

成立。

并且, 对于 Rademacher 序列  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_M)$ ,

$$E \left\| \sum_{j=1}^M \varepsilon_j B_j \right\|_{2 \rightarrow 2} \leq \sqrt{2 \ln(2d)} \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}^{1/2} \tag{8}$$

成立。

证明: 首先证明不等式(7), 先从(7)中左边开始着手

$$\begin{aligned} E \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} &= E \left( \sqrt{\lambda_{\max} \left( \sum_{j=1}^M g_j B_j \right)} \right) \leq E \sqrt{\sum_{j=1}^d \lambda_j \left( \sum_{j=1}^M g_j B_j \right)} \\ &\leq E \sqrt{\text{tr} \left( \sum_{j=1}^M g_j B_j \right)} = E \sqrt{\sum_{j=1}^M \text{tr} (g_j B_j)} \leq \sqrt{\sum_{j=1}^M \text{tr} (\ln E \exp(g_j B_j))} \end{aligned}$$

由于  $\text{tr}(A+B) = \text{tr}A + \text{tr}B$  以及引理 2.1 可知, 当  $\theta=1$  时可得  $E \exp(g_j B_j) = \exp(B_j^2/2)$ , 上式继续推导可得

$$\begin{aligned}
&= \sqrt{\sum_{j=1}^M \operatorname{tr} \left( \ln \exp \frac{(B_j)^2}{2} \right)} = \sqrt{\sum_{j=1}^M \operatorname{tr} \left( \frac{(B_j)^2}{2} \right)} = \sqrt{\operatorname{tr} \sum_{j=1}^M \left( \frac{(B_j)^2}{2} \right)} \\
&= \frac{\sqrt{2}}{2} \sqrt{\sum_{i=1}^d \lambda_i \left( \sum_{j=1}^M (B_j)^2 \right)} \leq \frac{\sqrt{2d}}{2} \sqrt{\lambda_{\max} \left( \sum_{j=1}^M (B_j)^2 \right)} = \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2}
\end{aligned}$$

最终结果为  $E \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \leq \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2}$ , 当  $\frac{\sqrt{2d}}{2} \leq \sqrt{2 \ln(2d)}$  成立时, 即  $d \leq 4 \ln(2d)$ ,

$$E \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \leq \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2} \leq \sqrt{2 \ln(2d)} \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}^{1/2} \text{ 成立。}$$

类似证明(8)也成立, 先从(8)中左边开始着手

$$\begin{aligned}
E \left\| \sum_{j=1}^M \varepsilon_j B_j \right\|_{2 \rightarrow 2} &= E \left( \sqrt{\lambda_{\max} \left( \sum_{j=1}^M \varepsilon_j B_j \right)} \right) \leq E \sqrt{\sum_{j=1}^d \lambda_j \left( \sum_{j=1}^M \varepsilon_j B_j \right)} \\
&\leq E \sqrt{\operatorname{tr} \left( \sum_{j=1}^M \varepsilon_j B_j \right)} = E \sqrt{\sum_{j=1}^M \operatorname{tr} (\varepsilon_j B_j)} \leq \sqrt{\sum_{j=1}^M \operatorname{tr} (\ln E \exp (\varepsilon_j B_j))}
\end{aligned}$$

由于  $\operatorname{tr}(A+B) = \operatorname{tr}A + \operatorname{tr}B$  以及引理 2.2 可知, 当  $\theta = 1$  时可得  $E \exp(\varepsilon_j B_j) \leq \exp(B_j^2/2)$ , 上式继续推导可得

$$\begin{aligned}
&= \sqrt{\sum_{j=1}^M \operatorname{tr} \left( \ln \exp \frac{(B_j)^2}{2} \right)} = \sqrt{\sum_{j=1}^M \operatorname{tr} \left( \frac{(B_j)^2}{2} \right)} = \sqrt{\operatorname{tr} \sum_{j=1}^M \left( \frac{(B_j)^2}{2} \right)} \\
&= \frac{\sqrt{2}}{2} \sqrt{\sum_{i=1}^d \lambda_i \left( \sum_{j=1}^M (B_j)^2 \right)} \leq \frac{\sqrt{2d}}{2} \sqrt{\lambda_{\max} \left( \sum_{j=1}^M (B_j)^2 \right)} = \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2}
\end{aligned}$$

最终结果为  $E \left\| \sum_{j=1}^M \varepsilon_j B_j \right\|_{2 \rightarrow 2} \leq \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2}$ , 当  $\frac{\sqrt{2d}}{2} \leq \sqrt{2 \ln(2d)}$  成立时, 即  $d \leq 4 \ln(2d)$ ,

$$E \left\| \sum_{j=1}^M \varepsilon_j B_j \right\|_{2 \rightarrow 2} \leq \frac{\sqrt{2d}}{2} \left\| \sum_{j=1}^M (B_j)^2 \right\|_{2 \rightarrow 2}^{1/2} \leq \sqrt{2 \ln(2d)} \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}^{1/2} \text{ 成立。}$$

### 3. 矩阵高斯和的 Hoeffding 型不等式

在证明之前首先证明一个简单的定理, 如下:

**定理 3.1:** 对于随机变量  $X$ , 满足以下不等式

$$EX \leq \inf_{\theta > 0} \theta^{-1} \ln E[\exp(\theta X)].$$

证明:  $EX = \theta^{-1} E \ln[\exp(\theta X)] = \theta^{-1} \ln E[\exp(\theta X)]$

所以  $\theta^{-1} \ln E[\exp(\theta X)]$  取下确界可得  $EX \leq \inf_{\theta > 0} \theta^{-1} \ln E[\exp(\theta X)]$ 。

接下来开始证明矩阵高和 Hoeffding 型不等式, 文献[3]中介绍了矩阵 Rademacher 和的 Hoeffding 型不等式, 读者可以对比阅读。

**定理 3.2:** 设  $g = (g_1, \dots, g_M)$  是一个由独立标准高斯随机变量构成的向量, 并且  $B_1, \dots, B_M \in \mathbb{C}^{d \times d}$  是自伴矩阵。引入  $\sigma^2 = \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}$ , 满足不等式

$$E \exp \left( \theta \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \right) \leq 2d \exp(\theta^2 \sigma^2 / 2) \tag{9}$$

其中  $\theta > 0$ 。并且对于  $t > 0$ ，概率满足

$$P \left( \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \geq t \right) \leq 2d \exp \left( \frac{-t^2}{2\sigma^2} \right). \tag{10}$$

证明：对于不等式(9)，我们从(9)的左边开始证明

$$\begin{aligned} E \exp \left( \theta \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \right) &= E \exp \left( \theta \lambda_{\max} \left( \sum_{j=1}^M g_j B_j \right) \right) = E \lambda_{\max} \exp \left( \theta \sum_{j=1}^M g_j B_j \right) \\ &\leq E \sum_{j=1}^d \lambda_j \exp \left( \theta \sum_{j=1}^M g_j B_j \right) = E \operatorname{tr} e^{\theta \sum_{j=1}^M g_j B_j} \leq \operatorname{tr} \exp \left( \sum_{j=1}^M \ln E \exp(\theta g_j B_j) \right) \end{aligned}$$

其中第二行的不等号应用了文献[3]中命题 8.16 的证明思路，等号应用了文献[3]中命题 8.16 以及[11]中的命题 3.1 证明时的思路，最后一个等号应用了文献[3]中命题 8.18。接着上式继续证明

$$\begin{aligned} &\leq \operatorname{tr} \exp \left( \sum_{j=1}^M \ln \exp(\theta^2 B_j^2 / 2) \right) = \operatorname{tr} \exp \left( \frac{\theta^2 \sum_{j=1}^M B_j^2}{2} \right) = \exp \left( \frac{\theta^2 \sum_{j=1}^d \lambda_j \left( \sum_{j=1}^M B_j^2 \right)}{2} \right) \\ &\leq \exp \left( \frac{\theta^2 d \lambda_{\max} \left( \sum_{j=1}^M B_j^2 \right)}{2} \right) = \exp \left( \frac{\theta^2 d \left\| \sum_{j=1}^M B_j^2 \right\|_{2 \rightarrow 2}}{2} \right) = \exp \left( \frac{\theta^2 d \sigma^2}{2} \right) \end{aligned}$$

在  $e^{d/2} \leq 2d$  ( $d \leq 2 \ln 2d$ ) 成立的条件下，不等式  $\exp(d\theta^2 \sigma^2 / 2) \leq 2d \exp(\theta^2 \sigma^2 / 2)$  成立，即

$$E \exp \left( \theta \left\| \sum_{j=1}^M g_j B_j \right\|_{2 \rightarrow 2} \right) \leq \exp(d\theta^2 \sigma^2 / 2) \leq 2d \exp(\theta^2 \sigma^2 / 2).$$

不等式(9)得以证明。

对于概率不等式(10)，证明过程可以借鉴文献[3]中命题 8.20 的证明过程如下，根据引理 2.1 可知  $E \exp(g\theta B) = \exp(\theta^2 B^2 / 2)$ ，这里  $B^2$  是半正定的。因此，文献[3]中公式(8.35)里  $g(\theta) = \theta^2 / 2$  并且  $A_\ell = B_\ell^2$ 。[3]中命题 8.19 的参数  $\rho$  被给定

$$\rho = \left\| \sum_{\ell=1}^M B_\ell^2 \right\|_{2 \rightarrow 2} = \sigma^2$$

其中  $\sum_{\ell=1}^M B_\ell^2$  是半正定的。由此

$$\begin{aligned} P \left( \left\| \sum_{\ell=1}^M g_\ell B_\ell \right\|_{2 \rightarrow 2} \geq t \right) &\leq P \left( \lambda_{\max} \left( \sum_{\ell=1}^M g_\ell B_\ell \right) \geq t \right) + P \left( \lambda_{\max} \left( -\sum_{\ell=1}^M g_\ell B_\ell \right) \geq t \right) \\ &\leq 2d \inf_{\theta > 0} e^{-\theta t + \theta^2 \sigma^2 / 2} = 2d e^{-t^2 / (2\sigma^2)}. \end{aligned}$$

这里  $\theta$  的最优选择是  $\theta = t / \sigma^2$ 。定理 3.2 的(10)得以证明。

接下来我们介绍矩阵版本的非对易伯恩斯坦不等式的次指数版本。

#### 4. Noncommutative Bernstein 不等式的次指数版本

首先给出证明时所需引理来辅助证明, 如下:

**引理 4.1:** 设  $Y \in \mathbb{C}^{d \times d}$  是一个自伴随机矩阵。那么, 对于  $t \in \mathbb{R}$ ,

$$P(\lambda_{\max}(Y) \geq t) \leq \inf_{\theta > 0} \exp\{e^{-\theta t} \text{Etr} \exp(\theta Y)\}. \quad (11)$$

**引理 4.2:** 设  $X_1, \dots, X_M \in \mathbb{C}^{d \times d}$  是独立自伴随机矩阵。那么, 对于  $\theta \in \mathbb{R}$ ,

$$\text{Etr} \exp\left(\theta \sum_{\ell=1}^M X_{\ell}\right) \leq \text{tr} \exp\left(\sum_{\ell=1}^M \ln \text{E} \exp(\theta X_{\ell})\right). \quad (12)$$

**引理 4.3:** 设  $X_1, \dots, X_M \in \mathbb{C}^{d \times d}$  是独立自伴随机矩阵。假设它们存在函数  $g: (0, \infty) \rightarrow [0, \infty)$  以及固定的自伴矩阵  $A_1, \dots, A_M$  使得

$$E[\exp(\theta X_{\ell})] \prec \exp(g(\theta) A_k), \quad \text{对于 } \forall \theta > 0 \text{ 以及 } \forall k \in [M]. \quad (13)$$

那么,  $\rho := \lambda_{\max}\left(\sum_{\ell=1}^M A_{\ell}\right)$ 。概率满足以下不等式

$$P\left(\lambda_{\max}\left(\sum_{\ell=1}^M X_{\ell}\right) \geq t\right) \leq d \inf_{\theta > 0} e^{-\theta t + g(\theta) \rho}, \quad t \in \mathbb{R}.$$

以上三个引理均摘自文献[11]。接下来开始证明 Noncommutative Bernstein 不等式的次指数版本, 其中 Noncommutative Bernstein 不等式在文献[5]中已经给出。

**定理 4.1 (Noncommutative Bernstein 不等式的次指数版本):** 设  $X_1, \dots, X_M \in \mathbb{C}^{d \times d}$  是独立零均值自伴随机矩阵。假设

$$E[X_{\ell}^n] \prec n! R^{n-2} \sigma_{\ell}^2 B_{\ell}^2 / 2, \quad \ell \in [M], \quad (14)$$

对于某些自伴矩阵  $B_{\ell}$ , 设

$$\sigma^2 := \left\| \sum_{\ell=1}^M B_{\ell}^2 \right\|_{2 \rightarrow 2}.$$

则满足以下概率不等式

$$P\left(\lambda_{\max}\left(\sum_{\ell=1}^M X_{\ell}\right) \geq t\right) \leq d \exp\left(-\frac{t^2/2}{\sigma^2 + Rt}\right). \quad (15)$$

其中  $t > 0$ 。

证明: 通过克拉默定理的启示, 我们估计  $X_{\ell}$  的矩生成函数。将指数函数展开为泰勒级数, 利用 Fubini 定理交换期望和求和得到

$$E[\exp(\theta X_{\ell})] = 1 + \theta E[X_{\ell}] + \sum_{n=2}^{\infty} \frac{\theta^n E[X_{\ell}^n]}{n!} = 1 + \frac{\theta^2 \sigma_{\ell}^2 B_{\ell}^2}{2} \sum_{n=2}^{\infty} \frac{\theta^{n-2} E[X_{\ell}^n]}{n! \sigma_{\ell}^2 B_{\ell}^2 / 2},$$

由于  $E[X_{\ell}] = 0$ 。定义  $F_{\ell}(\theta) = \sum_{n=2}^{\infty} \frac{\theta^{n-2} E[X_{\ell}^n]}{n! \sigma_{\ell}^2 B_{\ell}^2 / 2}$ , 我们可得

$$E[\exp(\theta X_{\ell})] = 1 + \theta^2 \sigma_{\ell}^2 B_{\ell}^2 F_{\ell}(\theta) / 2 \leq \exp(\theta^2 \sigma_{\ell}^2 B_{\ell}^2 F_{\ell}(\theta) / 2).$$

引入  $F(\theta) = \max_{\ell \in [M]} F_\ell(\theta)$  并且应用  $\sigma^2 := \left\| \sum_{\ell=1}^M B_\ell^2 \right\|_{2 \rightarrow 2}$ , 我们从克拉默定理中(文献[3]中定理 7.18)中可得

$$\begin{aligned} P\left(\lambda_{\max}\left(\sum_{\ell=1}^M X_\ell\right) \geq t\right) &\leq \inf_{\theta>0} \left\{ \exp(-\theta t) t r \exp\left(\sum_{\ell=1}^M \ln E \exp(\theta X_\ell)\right) \right\} \\ &\leq \inf_{\theta>0} \left\{ \exp(-\theta t) t r \exp\left(\sum_{\ell=1}^M \ln \exp(\theta^2 \sigma_\ell^2 B_\ell^2 F_\ell(\theta)/2)\right) \right\} \\ &\leq \exp(-\theta t) t r \exp\left(\sum_{\ell=1}^M (\theta^2 \sigma_\ell^2 B_\ell^2 F_\ell(\theta)/2)\right) \end{aligned}$$

由于  $E[X_\ell^n] \leq E[|X_\ell|^n]$ , 假设文献[3]中(7.38)成立便推导出

$$F_\ell(\theta) \leq \sum_{n=2}^{\infty} \frac{\theta^{n-2} E[|X_\ell|^n]}{n! \sigma_\ell^2 B_\ell^2 / 2} \leq \sum_{n=2}^{\infty} \frac{\theta^{n-2} n! R^{n-2} \sigma_\ell^2 B_\ell^2 / 2}{n! \sigma_\ell^2 B_\ell^2 / 2} = \sum_{n=2}^{\infty} (\theta R)^{n-2} = \frac{1}{1-R\theta}$$

这里  $R\theta < 1$ 。所以  $F(\theta) \leq (1-R\theta)^{-1}$  代入到以上概率不等式可得

$$\begin{aligned} P\left(\lambda_{\max}\left(\sum_{\ell=1}^M X_\ell\right) \geq t\right) &\leq \exp(-\theta t) t r \exp\left(\sum_{\ell=1}^M (\theta^2 \sigma_\ell^2 B_\ell^2 F_\ell(\theta)/2)\right) \\ &\leq \exp(-\theta t) t r \exp\left(\sum_{\ell=1}^M \left(\frac{\theta^2 \sigma_\ell^2 B_\ell^2}{2(1-R\theta)}\right)\right) \\ &= \exp(-\theta t) \sum_{i=1}^d \lambda_i \left( \exp\left(\sum_{\ell=1}^M \left(\frac{\theta^2 \sigma_\ell^2 B_\ell^2}{2(1-R\theta)}\right)\right) \right) \\ &\leq d \exp(-\theta t) \lambda_{\max} \left( \exp\left(\sum_{\ell=1}^M \left(\frac{\theta^2 \sigma_\ell^2 B_\ell^2}{2(1-R\theta)}\right)\right) \right) \\ &\leq d \exp(-\theta t) \exp\left(\frac{\theta^2}{2(1-R\theta)} \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right) \\ &\leq d \exp\left(-\theta t + \frac{\theta^2}{2(1-R\theta)} \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right) \end{aligned}$$

现在我们选择  $\theta = t / \left(Rt + \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right)$ , 很明显满足  $R\theta < 1$ 。这便推导出

$$\begin{aligned} P\left(\lambda_{\max}\left(\sum_{\ell=1}^M X_\ell\right) \geq t\right) &\leq d \exp\left(-\theta t + \frac{\theta^2}{2(1-R\theta)} \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right) \\ &\leq d \exp\left(-\frac{t^2}{Rt + \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)} + \frac{\frac{t^2}{\left(Rt + \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right)^2}}{2 \left(1 - \frac{Rt}{Rt + \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)}\right)} \lambda_{\max}\left(\sum_{\ell=1}^M (\sigma_\ell^2 B_\ell^2)\right)\right) \end{aligned}$$

$$\begin{aligned}
&= d \exp \left( \frac{-t^2/2}{Rt + \lambda_{\max} \left( \sum_{\ell=1}^M (\sigma_{\ell}^2 B_{\ell}^2) \right)} \right) = d \exp \left( \frac{-t^2/2}{Rt + \left\| \sum_{\ell=1}^M \sigma_{\ell}^2 B_{\ell}^2 \right\|_{2 \rightarrow 2}} \right) \leq d \exp \left( \frac{-t^2/2}{Rt + \left\| \sum_{\ell=1}^M B_{\ell}^2 \right\|_{2 \rightarrow 2}} \right) \\
&= d \exp \left( \frac{-t^2/2}{Rt + \lambda_{\max} \left( \sum_{\ell=1}^M (\sigma_{\ell}^2 B_{\ell}^2) \right)} \right) = d \exp \left( \frac{-t^2/2}{Rt + \left\| \sum_{\ell=1}^M \sigma_{\ell}^2 B_{\ell}^2 \right\|_{2 \rightarrow 2}} \right) \\
&\leq d \exp \left( \frac{-t^2/2}{Rt + \left\| \sum_{\ell=1}^M B_{\ell}^2 \right\|_{2 \rightarrow 2}} \right) = d \exp \left( \frac{-t^2/2}{Rt + \sigma^2} \right)
\end{aligned}$$

其中  $\sigma_{\ell}^2 \in (0, 1]$ , 这时满足不等式

$$P \left( \lambda_{\max} \left( \sum_{\ell=1}^M X_{\ell} \right) \geq t \right) \leq d \exp \left( \frac{-t^2/2}{Rt + \sigma^2} \right).$$

结论得证。特别的, 当随机变量是标量时的 Bernstein 不等式如下:

**定理 4.2:** 设  $X_1, \dots, X_M$  是独立零均值随机变量, 对于任意的整数  $n \geq 2$ , 满足以下不等式

$$E |X_{\ell}|^n \leq n! R^{n-2} \sigma_{\ell}^2 / 2, \quad \forall \ell \in [M], \quad (16)$$

对于某些常数  $R > 0$  以及  $\sigma_{\ell} > 0$ ,  $\ell \in [M]$ 。那么, 对于  $\forall t > 0$ , 满足概率不等式

$$P \left( \left| \sum_{\ell=1}^M X_{\ell} \right| \geq t \right) \leq 2 \exp \left( -\frac{t^2/2}{\sigma^2 + Rt} \right). \quad (17)$$

其中  $\sigma^2 := \sum_{\ell=1}^M \sigma_{\ell}^2$ 。

以上内容摘自文献[3]。

## 5. 结束语

本文在文献[3]中研究的基础上, 对其中一些未解决的问题进行了证明, 补充了相关内容, 使得这部分不等式内容更加充分, 在今后的研究证明中起到一定作用。希望这些补充会对读者有一些启发。

## 参考文献

- [1] Vershynin, R. (2012) Introduction to the Non-Asymptotic Analysis of Random Matrices. In: Eldar, Y. and Kutyniok, G., Eds., *Compressed Sensing: Theory and Applications*, Cambridge University Press, Cambridge, xii+544.
- [2] Qiu, R. and Wicks, M. (2014) *Cognitive Networked Sensing and Big Data*. Springer, New York. <https://doi.org/10.1007/978-1-4614-4544-9>
- [3] Foucart, S. and Rauhut, H. (2013) *A Mathematical Introduction to Compressive Sensing*. Springer, New York. <https://doi.org/10.1007/978-0-8176-4948-7>
- [4] van der Vaart, A.W. and Wellner, J.A. (1996) *Weak Convergence and Empirical Processes*. Springer, New York. [https://doi.org/10.1007/978-1-4757-2545-2\\_3](https://doi.org/10.1007/978-1-4757-2545-2_3)
- [5] Gross, D. (2011) Recovering Low-Rank Matrices from Few Coefficients in Any Basis. *IEEE Transactions on Information Theory*, **57**, 1548-1566. <https://doi.org/10.1109/TIT.2011.2104999>
- [6] Recht, B. (2011) A Simpler Approach to Matrix Completion. *The Journal of Machine Learning Research*, **12**, 3413-



- 3430.
- [7] Oliveira, R. (2009) Concentration of the Adjacency Matrix and of the Laplacian in Random Graphs with Independent Edges. arXiv:0911.0600.
  - [8] Baraniuk, R., Davenport, M.A., Duarte, M.F. and Hegde, C. (2011) An Introduction to Compressive Sensing. Rice University, Houston.
  - [9] Rauhut, H. (2010) Compressive Sensing and Structured Random Matrices. In: Fornasier, M., Ed., *Theoretical Foundations and Numerical Methods for Sparse Recovery, Radon Series on Computational and Applied Mathematics*, de Gruyter, Berlin, Vol. 9, 1-92.
  - [10] Oliveira, R. (2010) Sums of Random Hermitian Matrices and an Inequality by Rudelson. *Electronic Communications in Probability*, **15**, 203-212. <https://doi.org/10.1214/ECP.v15-1544>
  - [11] Tropp, J.A. (2012) User-Friendly Tail Bounds for Sums of Random Matrices. *Foundations of Computational Mathematics*, **12**, 389-434. <https://doi.org/10.1007/s10208-011-9099-z>