SCALAR AND MATRIX TAIL BOUND

LONG CHEN

CONTENTS

1. Tail Bound of One Random Variable 1
2. Tail Bound of Sum of Random Variables 3
3. Trace and Trace Inequality 4
3.1. Trace 4
3.2. Function of matrices 5
3.3. Trace inequalities 8
3.4. Functions of random matrices 9
4. Tail Bound of One Random Matrix 9
5. Tail Bound of Sum of Random Matrices 10
References 13

1. TAIL BOUND OF ONE RANDOM VARIABLE

We collect several inequalities on the tail bound of random variables.

Taking expectation of the inequality

$$\chi(\{X \geq a\}) \leq X/a,$$

we obtain the Markov's inequality.

Markov's inequality. Let X be a *non-negative* random variable, i.e., $X \geq 0$. Then, for any value $a > 0$,

$$\Pr\{X \geq a\} \leq \frac{\mathbb{E}[X]}{a}.$$

Apply Markov's inequality to the non-negative RV: $(X - \mathbb{E}[X])^2$ to get the Chebyshev's inequality.

Chebyshev's inequality. If X is a random variable with finite mean and variance, then, for any value $a > 0$,

$$\Pr\{|X - \mathbb{E}[X]| \geq a\} \leq \frac{\mathrm{Var}(X)}{a^2}.$$

If we know more moment of X, we can obtain more effective bounds. For example, if $\mathbb{E}[X^r]$ is finite for a positive integer r, then for any $a > 0$

$$\Pr\{X \geq a\} = \Pr\{X^r \geq a^r\} \leq \frac{\mathbb{E}[X^r]}{a^r}, \tag{1}$$

a bound that falls off as $1/a^r$. The larger the r, the greater the rate is, given a bound on $\mathbb{E}[X^r]$ is available. If we write the probability $\bar{F}(a) := \Pr\{X > a\} = 1 - \Pr\{X \leq a\} =$

Date: December 14, 2015.

$1 - F(a)$, then the bound (1) tells how fast the function $\bar{F}$ decays. The moments $\mathbb{E}[X^r]$ is finite implies the PDF f decays faster than $1/x^{r+1}$ and $\bar{F}$ decays like $1/x^r$.

We can improve the bound to be strictly less than one.

Chebyshev-Cantell inequality. If $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$, then, for any $a > 0$,

$$\Pr\{X \geq \mu + a\} \leq \frac{\sigma^2}{\sigma^2 + a^2},$$
$$\Pr\{X \leq \mu - a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Proof. Let $b > 0$ and note that $X \geq a$ is equivalent to $X + b \geq a + b$. Hence

$$\Pr\{X \geq a\} = \Pr\{X + b \geq a + b\} \leq \Pr\{(X + b)^2 \geq (a + b)^2\}.$$

Upon applying Markov's inequality, the preceding yields that

$$\Pr\{X \geq a\} \leq \frac{\mathbb{E}[(X + b)^2]}{(a + b)^2} = \frac{\sigma^2 + b^2}{(a + b)^2}.$$

Letting $b = \sigma^2/a$, which minimizes the upper bound, gives the desired result. $\square$

When the moment generating function $M_X(t) = \mathbb{E}[e^{tX}]$ is available (all moments are finite), we have the Chernoff bound which usually implies exponential decay of the tail.

Chernoff bounds.

$$\Pr\{X \geq a\} \leq \inf_{t>0} e^{-ta} M_X(t), \tag{2}$$
$$\Pr\{X \leq a\} \leq \inf_{t<0} e^{-ta} M_X(t). \tag{3}$$

A proof of the first inequality is as follows: for all $t > 0$

$$\Pr\{X \geq a\} = \Pr\left\{e^{tX} \geq e^{ta}\right\} \leq e^{-ta} M_X(t).$$

Taking the inf over all $t > 0$, we get the Chernoff bounds. Note that the moment generating function $M_X(t)$ might be exist only for a bounded interval $t \in I$. Then in the inf of (2), the $t \in I^+$.

Example 1.1 (Chernoff bounds for the standard normal distribution)**.** Let $X \sim N(0, 1)$ be the standard normal distribution. Then $M(t) = e^{t^2/2}$. So the Chernoff bound is given by

$$\Pr\left\{X \geq a\right\} \leq e^{-ta} e^{t^2/2} \quad \text{for all } t > 0.$$

The minimum is achieved at $t = a$ which gives the exponential decay tail bound

$$\Pr\left\{X \geq a\right\} \leq e^{-a^2/2} \quad \text{for all } a > 0. \tag{4}$$

In general, if $M_X(t) \leq e^{Ct^2/2}$, which is called sub-Gaussian, for some constant C and for all $t > 0$, then X has a sub-Gaussian upper tail $\Pr\left\{X \geq a\right\} \leq e^{-a^2/(2C)}$ with the same constant C. If the M-bound only holds for $t \in (0, t_0)$, then the T-bound holds for $a \in (0, Ct_0)$.

Example 1.2. *Let X be the random variable taking values ± 1 with probability $1/2$. Then*

$$\mathbb{E}\left[e^{tX}\right] = \frac{1}{2}(e^t + e^{-t}) \leq e^{t^2/2}.$$

In general for a bounded random variable, we have the Hoeffding's inequality.

Proposition 1.3 (Hoeffding's inequality)**.** *Let X be a random variable with $\mathbb{E}[X] = 0$ and $a \le X \le b$. Then for $t > 0$,*

$$\mathbb{E}[e^{tX}] \le e^{t^2(b-a)^2/8}.$$

Proof. We use the convexity of the exponential function to get the inequality

$$e^{tx} \le w_1(x)e^{ta} + w_2(x)e^{tb}$$

with $w_1(x) = (x-a)/(b-a)$ and $w_2 = (b-x)/(b-a)$. Apply the expectation operator and notice that $\mathbb{E}[X] = 0$ to get

$$\mathbb{E}[e^{tX}] \le w_2(0)e^{tb} - w_1(0)e^{ta} = (1 - p + pe^{t(b-a)})e^{-pt(b-a)} = e^{\Phi(u)},$$

where $p = -a/(b-a), u = t(b-a)$, and $\Phi(u) = -pu + \log(1 - p + pe^u)$. Now it is a calculus problem to show $\Phi(u) \le t^2(b-a)^2/8$, e.g. by Taylor series. $\square$

2. Tail Bound of Sum of Random Variables

Consider n-independent random variables. We want to pass properties of individual random variable to the sum. A sequence of random variables $X_1, X_2, \dots, X_n$ have a *uniform sub-Gaussian tail* if all of them have sub-Gaussian tails with the same constant.

The target is a tail bound for the sum of X_i. But when passing the properties of each random variable to the summation, working on the M-bound is much easier. So we first apply "T-bound to M-bound" procedure to each X_i and then "M-bound to T-bound"; see the diagram below

$$\begin{array}{ccc} \text{T-bound} & \longrightarrow & \text{M-bound} \\ \Sigma\Big\downarrow & & \Sigma\Big\downarrow \\ \text{T-bound} & \longrightarrow & \text{M-bound} \end{array}$$

Lemma 2.1. *Let $X_1, \dots, X_n$ be independent random variables satisfying $\mathbb{E}[X_i] = 0$, $\mathrm{Var}(X_i) = 1$, and having a uniform sub-Gaussian M-bound. Let $\alpha_1, \dots, \alpha_n$ be real coefficients satisfying $\sum_{i=1}^n \alpha_i^2 = 1$. Then the sum $Y = \sum_{i=1}^n \alpha_i X_i$ has $\mathbb{E}[Y] = 0, \mathrm{Var}(Y) = 1$, and a sub-Guassian tail.*

Proof. By the linearity of expectation, we get $\mathbb{E}[Y] = 0$. For independent random variables, the variance is additive and thus $\mathrm{Var}(Y) = \sum_{i=1}^n \alpha_i^2 \mathrm{Var}(X_i) = \sum_{i=1}^n \alpha_i^2 = 1$.

With assumption, $M_{X_i}(t) \le e^{Ct^2/2}$ for all $t > 0$ and all $i = 1, \dots, n$, we have

$$M_Y(t) = \mathbb{E}\left[e^{tY}\right] = \mathbb{E}\left[\prod_{i=1}^n e^{t\alpha_i X_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{t\alpha_i X_i}\right] \le e^{\frac{1}{2}Ct^2 \sum_{i=1}^n \alpha_i^2} = e^{Ct^2/2}.$$

Then the M-bound of Y implies the desired tail bound. $\square$

Example 2.2 (The 2-stability of Gaussian distribution)**.** Let $X_1, \dots, X_n$ be i.i.d $N(0,1)$. Let $\alpha_1, \dots, \alpha_n$ be real coefficients satisfying $\sum_{i=1}^n \alpha_i^2 = 1$. Then $Y = \sum_{i=1}^n \alpha_i X_i$ is still the standard normal distribution, i.e., $Y \sim N(0,1)$ since the moment generation function uniquely determines the random variable.

Example 2.3 (Hoeffding's inequality)**.** *Let $X_1, \dots, X_n$ be independent random variables with zero mean $\mathbb{E}[X_i] = 0$ and the bound $a_i \le X_i \le b_i$ for $i = 1, \dots, n$. Let $Y = \sum_{i=1}^n X_i$. Then by the Hoeffding's inequality for one random variables, we have, for any $\epsilon > 0$,*

$$\Pr\{Y \ge \epsilon\} \le \exp(-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2).$$

If the variance is finite, we could improve to the Bernstein's inequality. We refer to [3] for a proof.

Theorem 2.4 (Bernstein's inequality). *Let $X_1, \dots, X_n$ be independent random variables with zero mean $\mathbb{E}[X_i] = 0$ and uniform bound $|X_i| \leq M$ for $i = 1, \dots, n$. Let $\sigma^2 = \frac{1}{n}\sum_{i=1}^n \operatorname{Var}(X_i)$. Then for any $\epsilon > 0$,*

$$\Pr\left\{\frac{1}{n}\sum_{i=1}^n X_i \geq \epsilon\right\} \leq \exp\left(-\frac{n\epsilon^2}{2\sigma^2 + 2M\epsilon/3}\right).$$

When $X_i \sim N(0, \sigma_i)$, the power in the bound is $-n\epsilon^2/(2\sigma^2)$. For general independent random variables, by the central limit theorem, the bound holds with power $-n\epsilon^2/(2\sigma^2)$, if n is large enough. The Bernstein's inequality is qualitatively right up to a term $2c\epsilon/3$ which is small if $\epsilon \ll 1$. On the other hand, if $\epsilon > \sigma^2$, then the bound behaves like $e^{-n\epsilon}$ which is an improvement of $e^{-n\epsilon^2}$. It is an improvement since now to get the probability less than $\delta \in (0,1)$, we can chose smaller $n = C\epsilon^{-1}\ln(1/\delta)$ instead of $n = C\epsilon^{-2}\ln(1/\delta)$.

We also point out this improvement is only for bounded random variables which rules out the most popular Gaussian.

The independence of X_i can relaxed. Consider a martingale.

xxx definition of martingale xxx

Theorem 2.5 (Azuma-Hoeffding's inequality). *Suppose $S_n, n = 0, 1, 2\dots$ is a martingale and $|S_n - S_{n-1}| < c_n$ almost surely. Then for all positive integers N and all $t > 0$*

$$\Pr\{X_N - X_0 \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^N c_k^2}\right).$$

3. Trace and Trace Inequality

Let A be a $n \times n$ matrix. The trace of A is:

$$\operatorname{tr}(A) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i(A).$$

We follow Carlen [2] to provide some background on the trace inequalities involving functions of matrices.

3.1. Trace.

If A is an $m \times n$ matrix and B is an $n \times m$ matrix, then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

which can be easily verified by direct computation.

The trace $\operatorname{tr}(A)$ is the sum of all eigenvalues of A and thus is invariant with respect to a change of basis:

$$\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(APP^{-1}) = \operatorname{tr}(A).$$

In particular,

Lemma 3.1. *For any orthonormal basis $\{u_1, u_2, \dots u_n\}$*

$$\operatorname{tr}(A) = \sum_{i=1}^n (u_j, Au_j). \tag{5}$$

Proof. Let $U = (u_1, u_2, \dots u_n)$ be the orthonormal matrix. Then $\operatorname{tr}(A) = \operatorname{tr}(U^*AU) = \sum_{i=1}^n (u_j, Au_j)$. □

The trace of a projection matrix is the dimension of the target space. That is if $P = A(A^TA)^{-1}A^T$, then

$$\mathrm{tr}(P) = \mathrm{rank}(A).$$

The trace of a product can be rewritten as the sum of entry-wise products of elements:

$$\mathrm{tr}(X^TY) = X : Y = \sum_{ij} X_{ij}Y_{ij} = \sum_{ij}(X.*Y)_{ij} = \mathrm{vec}(X)^T\mathrm{vec}(Y).$$

This means that the trace of a product of matrices functions similarly to a dot product of vectors. For this reason, generalizations of vector operations to matrices often involve a trace of matrix products. The norm induced by the above inner product is called the Frobenius norm.

Unlike the determinant, the trace of the product is not the product of traces. It is true, however, when apply to the tensor product:

$$\mathrm{tr}(X \otimes Y) = \mathrm{tr}(X)\,\mathrm{tr}(Y).$$

3.2. **Function of matrices.** Motivation: generalize properties of functions of single variable to functions of matrices. The matrix should be restricted to Hermitian space and most generalization works for trace of matrix functions.

Denoted by

- M_n the space of $n \times n$ matrices;
- H_n the space of Hermitian $n \times n$ matrices;
- H_n^+ the set of positive semi-definite Hermitian matrices;
- H_n^{++} the set of positive definite Hermitian matrices;
- S_n the set of density matrices i.e. $\mathrm{tr}(\rho) = 1$ for $\rho \in M_n$.

We can introduce a partial ordering of H_n: for $A, M \in H_n$,

$$A \leq M \iff x^TAx \leq x^TMx, \forall x \in \mathbb{R}^n.$$

Definition of function of matrices. Consider a scalar function $f : \mathbb{R} \to \mathbb{R}$. For a diagonal matrix $D = \mathrm{diag}(d_1, \cdots, d_n)$, $f(D) := \mathrm{diag}(f(d_1), \cdots, f(d_n))$. For $A \in H_n$, let $A = Q\Lambda Q^*$ be the eigen-decomposition of A. We define

$$f(A) := Qf(\Lambda)Q^*.$$

If we write

$$Q\Lambda Q^* = \sum_{i=1}^{n} \lambda_i Q_{(:,i)}Q^*_{(:,i)} := \sum_{i=1}^{n} \lambda_i P_i$$

with the rank-1 projection $P_i = Q_{(:,i)}Q^*_{(:,i)}$, an equivalent definition of matrix function is

$$f(A) = \sum_{i=1}^{n} f(\lambda_i)P_i. \tag{6}$$

Taking trace and using $\mathrm{tr}(P_i) = 1$, we obtain

$$\mathrm{tr}\, f(A) = \sum_{i=1}^{n} f(\lambda_i). \tag{7}$$

Two important matrix functions are exponential and logarithm functions of matrices. We have the power series expansion:

$$\exp(A) = I + \sum_{k=1}^{\infty} \frac{A^k}{k!}.$$

Inequalities of scalar functions can be generalized to matrix version by the following transfer rule:

Lemma 3.2. *If $f(x) \leq g(x)$ for $x \in I$, then for all $A \in H_n$ and $\sigma(A) \subseteq I$, $f(A) \leq g(A)$.*

As a consequence, some inequalities of exponential functions can be generalized to exponential of matrices.

Exercise 3.3. Prove Lemma 3.2 and consequently, for $A \in H_n$, prove that

(1) $e^A \geq 0$.
(2) $I + A \leq e^A$.
(3) $\cosh(A) \leq \exp(A^2/2)$.

□

The von Neuman *entropy* of $\rho \in S_n, \mathcal{S}(\rho)$ is defined by

$$\mathcal{S}(\rho) = -\operatorname{tr}(\rho \log \rho).$$

Exercise 3.4. *Prove that, for $\rho \in S_n$,*

$$0 \leq \mathcal{S}(\rho) \leq \log n.$$

And given conditions when the equality holds

A related function is

$$A \to \log \operatorname{tr}(e^A).$$

It is shown that the function $\log \operatorname{tr}(e^A)$ is the Legendre transforms of $\mathcal{S}(A)$ if we extend the definition of $\mathcal{S}(A) = -\infty$ for $A \notin S_n$.

We want to extend some properties of scalar functions to matrix functions. In the following we restrict matrices in H_n with spectrum restricted to the domain of f. The properties we concern are:

Operator monotone. A function f is operator monotone if the following holds

$$A \geq B \Longrightarrow f(A) \geq f(B).$$

Obviously if f is operator monotone, then $f(x)$ is monotone. But not all monotone functions are operator monotone due to the non-commutative algebra structure of M_n.

Operator convex. A function $f : I \to \mathbb{R}$ is operator convex if for all $A, B \in H_n$ and $\sigma(A), \sigma(B) \subset I$ and for $\theta \in (0,1)$, the following holds

$$f(\theta A + (1-\theta)B) \leq \theta f(A) + (1-\theta) f(B).$$

A function f is operator concave if $-f$ is operator convex.

Example 3.5 (A^2 is not operator monotone)**.** The function $f(x) = x^2$ is monotone but for $A, B \in H_n^+$,

$$(A+B)^2 = A^2 + (AB + BA) + B^2.$$

Due to the non-coummutative property of matrix product, the term $AB + BA$ could have negative eigenvalue and thus $(A + tB)^2 \geq A^2$ could fail for sufficiently small t. An example is $A = [1 \quad 1;\ 1 \quad 1], B = [1 \quad 0;\ 0 \quad 0]$.

Example 3.6 (A^2 is operator convex)**.** With the parallelogram law

$$\left(\frac{A+B}{2}\right)^2 + \left(\frac{A-B}{2}\right)^2 = \frac{1}{2}A^2 + \frac{1}{2}B^2,$$

we get the convexity for $\theta = 1/2$, which is known as *midpoint convexity*. From that, we can prove the convexity for all $\theta \in (0,1)$.

Example 3.7 ($A^{1/2}$ is operator monotone)**.** We can prove that if $A, B \in H_n^+$ and $A^2 \leq B^2$, then $A \leq B$.

Example 3.8 ($-A^{-1}$ is operator monotone and concave)**.** The monotonicity can be proved by the following identity: for $A, B \in H_n^+$

$$A^{-1} - (A+B)^{-1} = A^{-1/2}\left[I - (I+C)^{-1}\right]A^{-1/2},$$

with $C = A^{-1/2}BA^{-1/2} \in H_n^+$.

The midpoint convexity can be proved by the identity:

$$\frac{1}{2}A^{-1} + \frac{1}{2}B^{-1} - \left(\frac{A+B}{2}\right)^{-1} = A^{-1/2}\left[\frac{1}{2}I + \frac{1}{2}C^{-1} - \left(\frac{I+C}{2}\right)^{-1}\right]A^{-1/2},$$

and then the convexity of the function $f(x) = 1/x$.

Theorem 3.9 (Löwner-Heinz)**.** *On the function $f(x) = x^p$, we have*

- *For $-1 \leq p \leq 0$, the function $f(A) = -A^p$ is operator concave and operator monotone.*
- *For $0 \leq p \leq 1$, the function $f(A) = A^p$ is operator concave and operator monotone.*
- *For $1 < p \leq 2$, the function $f(A) = A^p$ is operator convex but not operator monotone when p is near 2.*

In short for the power $p \in [-1,1]$, the monotonicity and convexity/concavity is preserved for function x^p. When $p \in [1,2]$, the convexity is still preserved. When $p > 2$, even the convexity could be missing.

An elementary proof given by Carlen [2] is based on the integral form of A^p. For example, for $p \in (-1,0)$,

$$A^p = \frac{\pi}{\sin((p+1)\pi)} \int_0^\infty t^p (tI + A)^{-1}\, \mathrm{d}t. \tag{8}$$

The integral is a weighted sum of monotone and convex function $A \to (tI+A)^{-1}$ and thus preserves the monotonicity and convexity. The case for $1 < p < 2$ is different since it is related to the sum of two operator convex functions but the difference of two operator monotone functions.

Exercise 3.10. (1) Prove (8) using a contour integral for scalar, i.e., $A = a$.
(2) Write out integral form for A^p when $p \in (0,1)$ and $p \in (1,2)$.

Corollary 3.11. *(1) $\log(A)$ is operator concave and operator monotone*
(2) $A\log(A)$ is operator convex but not monotone.

Proof. We have the limit

$$\log(A) = \lim_{p\to 0} \frac{A^p - I}{p}$$

and A^p is nice for p near 0. Thus the properties pass to the limit.

The other limit is

$$A \log(A) = \lim_{p \to 1} \frac{A^p - A}{p - 1}$$

□

We emphasize that the function $\exp(A)$ and A^p for $p > 2$ are neither operator monotone nor operator convex.

3.3. **Trace inequalities.** Monotone and convex operators are very rare. Taking trace of matrix functions, however, we can preserve the convexity and monotonicity property.

Theorem 3.12. *The trace of a matrix function will preserve the convexity and monotonicity property. Namely*

- *if f is monotone increasing, so is $\operatorname{tr} f(A)$ on H_n.*
- *if f is convex (or concave), so is $\operatorname{tr} f(A)$ on H_n.*

Proof. Let $A, B \in H_n$ and $C = A - B > 0$. We first consider the differentiable monotone function f with $f' > 0$. Then

$$\begin{aligned}\operatorname{tr}(f(A)) - \operatorname{tr}(f(B)) &= \int_0^1 \frac{d}{dt} \operatorname{tr}(f(B + t(A - B)))\, \mathrm{d}t \\ &= \int_0^1 \operatorname{tr}(C^{1/2} f'(B + tC) C^{1/2})\, \mathrm{d}t \geq 0.\end{aligned}$$

Using the continuity argument, the differentiability can be relaxed to monotone only.

To prove the convexity, we need the following Peierls inequality: for any orthonormal basis $\{u_1, \ldots, u_n\}$ and for a convex function f

$$\sum_{i=1}^n f((u_i, Au_i)) \leq \operatorname{tr}(f(A)). \tag{9}$$

And the equality holds when each u_i is an eigenvector of A. Note that we can write the right hand side $\operatorname{tr}(f(A)) = \sum_{i=1}^n f(\lambda_i) = \sum_{i=1}^n f((v_i, Av_i))$ with (λ_i, v_i) being the eigen-pair of A. This verifies the equality.

To prove inequality (9), we write $A = \sum_{k=1}^n \lambda_k P_k$ and thus

$$(u_i, Au_i) = \sum_{k=1}^n \lambda_k (u_i, P_k u_i) = \sum_{k=1}^n \lambda_k \|P_k u_i\|^2.$$

As the nonnegative weight $w_k = \|P_k u_i\|^2$ satisfies $\sum_k w_k = \|u_i\|^2 = 1$, we use the convexity of f and the definition of $f(A)$ to conclude that

$$f\left(\sum_{k=1}^n w_k \lambda_k\right) \leq \sum_{k=1}^n \|P_k u_i\|^2 f(\lambda_k) = \sum_{k=1}^n (u_i, f(\lambda_k) P_k u_i) = (u_i, f(A) u_i).$$

Sum over index i and use the identity for $\operatorname{tr}(A)$, c.f. Lemma 3.1, we finish the proof. □

Example 3.13 (Golden-Thompson inequality)**.** *For any $A, B \in H_n$,*

$$\operatorname{tr}(e^{A+B}) \leq \operatorname{tr}(e^A e^B).$$

Theorem 3.14 (Lieb)**.** *For a fixed Hermitian matrix $L \in H_n$, the function $f(A) = \operatorname{tr}(\exp(L + \log A))$ is concave on H_n^{++}.*

3.4. **Functions of random matrices.** Let X be a random matrix in H_n. We define the moment generating function (mgf) and the cumulant-generating function (cgf):

$$M_X(t) := \mathbb{E}\left[e^{tX}\right], \quad \text{and} \quad \Xi(t) := \log \mathbb{E}\left[e^{tX}\right],$$

for t in an open interval containing zero. They admit a formal power series expansions:

$$M_X(t) = I + \sum_{k=1}^{\infty} \frac{t^k}{k!}\mathbb{E}[X^k],$$
$$\Xi_X(t) = \sum_{k=1}^{\infty} \frac{t^k}{k!}\Psi_k.$$

The coefficients $\mathbb{E}[X^k]$ are called *matrix moment* and Ψ_k as a *matrix cumulant*. In particular, $\Psi_1 = \mathbb{E}[X]$ is the mean and $\Psi_2 = \mathbb{E}[X^2] - \mathbb{E}^2[X]$ is the variance.

The expectation operator $\mathbb{E}$ is exchangeable to a linear operator especially

$$\mathbb{E}[\operatorname{tr}(X)] = \operatorname{tr}\mathbb{E}[X].$$

For an operator convex function f, as the expectation is a weighted average, we have the Jensen's inequality

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

In particular, as $f(x) = x^2$ is operator convex,

$$\mathbb{E}^2[X] \leq \mathbb{E}[X^2].$$

And the expectation preserves the partial ordering in H_n, i.e.

$$\mathbb{E}[A] \leq \mathbb{E}[B] \quad \text{if } A \leq B.$$

4. Tail Bound of One Random Matrix

We provide matrix version of various tail bounds.

Lemma 4.1 (Matrix Markov inequality)**.** *Let X be a random matrix in H_n^+. Then for all deterministic matrix $A \in H_n^{++}$, we have*

$$\Pr\{X \not\leq A\} \leq \operatorname{tr}(A^{-1}\mathbb{E}[X]).$$

Proof. We consider the matrix $A^{-1}X$ which is Hermitian w.r.t the $(\cdot,\cdot)_A$ inner product. $X \not\leq A$ is equivalent to $A^{-1}X \not\leq_A I$ which implies that $\|A^{-1}X\| = \lambda_{\max}(A^{-1}X) > 1$. Then we obtain the inequality

$$\chi_{\{X \not\leq A\}} \leq \lambda_{\max}(A^{-1}X) \leq \operatorname{tr}(A^{-1}X).$$

Taking the expectation and using the fact $\mathbb{E}$ is linear, we obtain the desired inequality. $\square$

Exercise 4.2. *Use the fact $f(x) = x^{1/2}$ is operator monotone to derive a matrix version Chebyschev inequality.*

We now present a matrix Chernoff bound established by Oliverira [4].

Theorem 4.3 (Matrix Chernoff bound)**.** *Let X be a random matrix in H_n. For all $a \in \mathbb{R}$,*

$$\Pr\{\lambda_{\max}(X) \geq a\} \leq \inf_{t>0} e^{-ta}\mathbb{E}\left[\operatorname{tr} e^{tX}\right].$$

Proof. By the scalar Chernoff bound, we obtain

$$\Pr\{\lambda_{\max}(X) \geq a\} \leq \inf_{t>0} e^{-ta}\mathbb{E}\left[e^{t\lambda_{\max}(X)}\right].$$

Use

$$e^{t\lambda_{\max}(X)} = \lambda_{\max}(e^{tX}) \leq \operatorname{tr}(e^{tX})$$

to get the desired inequality. $\square$

5. Tail Bound of Sum of Random Matrices

We shall follow Tropp [5] to present tail bounds of sums of random matrices.

In the scalar case, we use the fact

$$\mathbb{E}\left[e^{t\sum_i X_i}\right] = \prod_i \mathbb{E}\left[e^{tX_i}\right] \tag{10}$$

to get the additivity of independent sub-Gaussians. For matrices, first of all,

$$e^{A+B} \neq e^A e^B$$

again due to the non-commutative algebra structure. Taking trace, we could get a desired inequality (using Golden-Thompson inequality) for *two* independent random matrices

$$\mathbb{E}\operatorname{tr}\left[e^{X_1+X_2}\right] \leq \mathbb{E}\operatorname{tr}\left(e^{X_1}e^{X_2}\right) = \operatorname{tr}\left(\mathbb{E}\left[e^{X_1}\right]\mathbb{E}\left[e^{X_1}\right]\right). \tag{11}$$

Unfortunately, the above inequality cannot be generalized to three or more matrices.

The route Ahlswede and Winter [1] take is to recrusively apply inequality (11) and the inequality $\operatorname{tr}(AB) \leq \operatorname{tr}(A)\lambda_{\max}(B)$ for $A, B \in H_n^+$ and end up with

$$\mathbb{E}\operatorname{tr}\left[e^{\sum_k X_k}\right] \leq d\exp\left(\sum_k \lambda_{\max}(\log\mathbb{E}[e^{X_k}])\right) \tag{12}$$

We shall follow Tropp to present a shaper result with upper bound involving a smaller quantity $\lambda_{\max}(\sum_k \log\mathbb{E}[e^{X_k}])$. To do so, we use a random matrix version of Lieb's concave result, c.f. Theorem 3.14.

Lemma 5.1. *Let $L \in H_n$ be a fixed Hermitian matrix and let X be a random matrix in H_n. Then*

$$\mathbb{E}[\operatorname{tr}\exp(L+X)] \leq \operatorname{tr}\exp(L+\log\mathbb{E}[e^X]).$$

Proof. Let $A = e^X$. By Lieb's concave theorem, $A \to \operatorname{tr}\exp(L+\log A)$ is concave. Then apply Jensen's inequality to the expectation to get the desired result. $\square$

Lemma 5.2 (Subadditivity of Matrix cgf's)**.** *Let $\{X_k\}$ be a sequence of independent random matrices in H_d. Then*

$$\mathbb{E}\left[\operatorname{tr}\exp\left(\sum_k X_k\right)\right] \leq \operatorname{tr}\exp\left(\sum_k \log\mathbb{E}\left[e^{X_k}\right]\right).$$

Proof. Tropp's Lemma 3.4. $\square$

Theorem 5.3 (Master Tail Bound)**.** *Let $\{X_k\}$ be a sequence of independent random matrices in H_d. For all $a \in \mathbb{R}$,*

$$\Pr\{\lambda_{\max}(X) \geq a\} \leq \inf_{t>0} e^{-ta}\operatorname{tr}\exp\left(\sum_{k=1}^n \log\mathbb{E}\left[e^{tX_k}\right]\right).$$

Proof. It is a combination of matrix Chernoff bound, c.f. Theorem 4.3 and the sub-additivity of matrix cgf. ☐

Then we combine bounds of mgf/cgf to a Chernoff bound.

Corollary 5.4. *Let $\{X_k\}$ be a sequence of independent random matrices in H_d. Assume*

$$\mathbb{E}\left[e^{tX_k}\right] \leq e^{g(t)A_k}$$

with a function $g : \mathbb{R}^+ \to \mathbb{R}^+$ and deterministic matrices $A_k \in H_d$ for $k = 1, 2, \ldots, n$. Then, for all $a \in \mathbb{R}$,

$$\Pr\left\{\lambda_{\max}\left(\sum_k X_k\right) \geq a\right\} \leq d \inf_{t>0} e^{-ta+g(t)\rho_A}, \tag{13}$$

with parameter $\rho_A = \lambda_{\max}\left(\sum_k A_k\right)$.

Proof. Use the inequality $\mathrm{tr}(M) \leq d\lambda_{\max}(M)$ for $M \in H_d$. ☐

Lemma 5.5. *Let X be a random matrix in H_d. Assume*

$$0 \leq \lambda_{\min}(X) \leq \lambda_{\max}(X) \leq 1.$$

Then for all $t \in \mathbb{R}$

$$\mathbb{E}\left[e^{tX}\right] \leq I + (e^t - 1)\mathbb{E}[X] \leq \exp\left((e^t - 1)\mathbb{E}[X]\right). \tag{14}$$

Theorem 5.6 (Matrix Chernoff I)**.** *Let $\{X_k\}$ be a sequence of independent random matrices in H_d. Assume*

$$0 \leq \lambda_{\min}(X_k) \leq \lambda_{\max}(X_k) \leq R.$$

Let

$$\mu_{\min} = \lambda_{\min}\left(\sum_k \mathbb{E}[X_k]\right) \quad \text{and} \quad \mu_{\max} = \lambda_{\max}\left(\sum_k \mathbb{E}[X_k]\right).$$

Then

$$\Pr\left\{\lambda_{\min}\left(\sum_k X_k\right) \leq (1-\delta)\mu_{\min}\right\} \leq d\left[\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right]^{\mu_{\min}/R} \quad \text{for } \delta \in [0,1], \tag{15}$$

and

$$\Pr\left\{\lambda_{\max}\left(\sum_k X_k\right) \geq (1+\delta)\mu_{\max}\right\} \leq d\left[\frac{e^{-\delta}}{(1+\delta)^{1+\delta}}\right]^{\mu_{\max}/R} \quad \text{for } \delta \geq 0. \tag{16}$$

To obtain Bernstein type inequality, we need to refine the bound of mgf.

Lemma 5.7. *Let X be a random matrix in H_n and*

$$\mathbb{E}[X] = 0 \quad \text{and} \quad \lambda_{\max}(X) \leq 1.$$

Then for all $t > 0$

$$\mathbb{E}\left[e^{tX}\right] \leq \exp\left((e^t - t - 1)\mathbb{E}[X^2]\right).$$

Proof. We can prove the scalar inequality

$$e^{tx} \leq 1 + tx + (e^t - t - 1)x^2, \quad \forall x \leq 1, \tag{17}$$

by power series of exponential functions, and then transfer to the matrix inequality by Lemma 3.2

$$e^{tX} \leq I + tX + (e^t - t - 1)X^2.$$

Taking expectation and using the inequality $I + A \leq e^A$ to get the desired inequality. ☐

Theorem 5.8 (Matrix Bennett and Bernstein)**.** *Let $\{X_k\}$ be a sequence of independent random matrices in H_d. Assume*

$$\mathbb{E}[X_k] = 0 \quad \text{and} \quad \lambda_{\max}(X_k) \leq R.$$

Let

$$\sigma^2 = \frac{1}{n}\left\|\sum_{k=1}^{n} \mathrm{Var}(X_k)\right\| = \frac{1}{n}\left\|\sum_{k=1}^{n} \mathbb{E}(X_k^2)\right\|.$$

Then for all $\epsilon \geq 0$

(18)

$$\Pr\left\{\lambda_{\max}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \geq \epsilon\right\} \leq d\exp\left(-\frac{n\sigma^2}{R^2}h(\frac{R\epsilon}{\sigma^2})\right) \leq d\exp\left(-\frac{n\epsilon^2}{2\sigma^2+2R\epsilon/3}\right),$$

where the function $h(u) = (1+u)\log(1+u) - u$ for $u \geq 0$.

Proof. The first one is Bennett and the second is Bernstein. From Bennett to Bernstein is from the bound

$$h(u) \geq \frac{u^2}{2+2u/3}.$$

□

Remark 5.9. *If replace the boundedness of $\lambda_{\max}(X_k)$ to the boundedness of*

$$\|X_k\| \leq R,$$

by applying the Bernstein estimate to X_k and $-X_k$, we could obtain the tail bound for the spectral norm

$$\Pr\left\{\left\|\frac{1}{n}\sum_{k=1}^{n} X_k\right\| \geq \epsilon\right\} \leq 2d\exp\left(-\frac{n\epsilon^2}{2\sigma^2+2R\epsilon/3}\right). \tag{19}$$

We present a concentration result for sum of rank-1 matrices which is quite useful in the randomized numerical linear algebra.

Corollary 5.10. *Let $y_1, y_2, \ldots, y_n$ be i.i.d. random column vectors in $\mathbb{C}^d$ with*

$$\|y_i\| \leq M \quad \text{and} \quad \|\,\mathbb{E}\left[y_1 y_1^*\right]\| \leq 1.$$

Then for all $0 \leq \epsilon \leq 1$

$$\Pr\left\{\left\|\frac{1}{n}\sum_{k=1}^{n} y_k y_k^* - \mathbb{E}\left[y_1 y_1^*\right]\right\| \geq \epsilon\right\} \leq 2d\exp\left(-\frac{3n\epsilon^2}{8(M^2+1)}\right).$$

In [4], the bound is: for all $\epsilon \geq 0$

$$\Pr\left\{\left\|\frac{1}{n}\sum_{k=1}^{n} y_k y_k^* - \mathbb{E}\left[y_1 y_1^*\right]\right\| \geq \epsilon\right\} \leq (2\min(d,n))^2\exp\left(-\frac{n}{16M^2}\min(\epsilon^2, 4\epsilon-4)\right),$$

which leads to meaningful results even the ambient dimension d is arbitrarily large.

To control the spectral norm of rectangular matrices, we can use the dilation to form a larger matrix, i.e., for $B_{d_1\times d_2}$, we let

$$\mathscr{F}(B) := \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}.$$

Then $\mathscr{F}(B) \in H_{d_1+d_2}$ and

$$\mathscr{F}(B)^2 := \begin{pmatrix} BB^* & 0 \\ 0 & B^*B \end{pmatrix}.$$

Consequently

$$\lambda_{\max}(\mathscr{F}(B)) = \|\mathscr{F}(B))\| = \|B\|.$$

Corollary 5.11 (Rectangular Matrix Bernstein)**.** *Let $\{Z_k\}$ be a sequence of independent random matrices of size $d_1 \times d_2$. Assume*

$$\mathbb{E}[Z_k] = 0 \quad \textit{and} \quad \lambda_{\max}(Z_k) \le R.$$

Let

$$\sigma^2 = \max\left\{\left\|\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(Z_k Z_k^*)\right\|, \left\|\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(Z_k^* Z_k)\right\|\right\}.$$

Then for all $\epsilon \ge 0$

$$\Pr\left\{\left\|\frac{1}{n}\sum_{k=1}^{n} Z_k\right\| \ge \epsilon\right\} \le d\exp\left(-\frac{n\epsilon^2}{2\sigma^2 + 2R\epsilon/3}\right), \tag{20}$$

where the function $h(u) = (1+u)\log(1+u) - u$ for $u \ge 0$ and $d = d_1 + d_2$.

References

[1] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48:569–579, 2002.

[2] E. Carlen. Trace inequalities and quantum entropy: an introductory course. *Contemporary Mathematics*, 0000:73–140, 2010.

[3] G. Lugosi. Concentration-of-measure inequalities. *Lecture Notes*, XXXIII(2):81–87, 2014.

[4] R. I. Oliveira. Sums of random hermitian matrices and an inequality by Rudelson. *Electronic Communications in Probability*, 15:203–212, 2010.

[5] J. A. Tropp. User-Friendly Tail Bounds for Sums of Random Matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.