**Sixth Cargese Workshop on Combinatorial Optimization** September 16, 2015

# Lecture 2: Matrix Chernoff bounds

*Nick Harvey* *University of British Columbia*

**Abstract**

The purpose of my second and third lectures is to discuss spectral sparsifiers, which are the second key ingredient in most of the fast Laplacian solvers. In this lecture we will discuss concentration bounds for sums of random matrices, which are an important technical tool underlying the simplest sparsifier construction.

## 1 Introduction

*I thought it was a rather trivial lemma, but many things are only trivial once you know them.*
*Herman Chernoff*

The use of randomization in algorithms has been increasingly prevalent since the mid 1970s. The Chernoff bound has been a hugely important tool in randomized algorithms and learning theory since the mid 1980s. It is a concentration inequality for random variables that are the sum of many independent, bounded random variables. A formal statement is:

**Theorem 1.** Let $X_1, \ldots, X_k$ be independent, random, real variables with $0 \le X_i \le R$. Let $\mu_{\min} \le \sum_i \mathrm{E}[X_i] \le \mu_{\max}$. Then, for all $\delta \ge 0$,

$$\begin{aligned}
\Pr\left[\sum_{i=1}^k X_i \ge (1+\delta)\mu_{\max}\right] &\le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mu_{\max}/R} \le \begin{cases} e^{-\delta^2\mu_{\max}/3R} & (\text{if } \delta \le 1) \\ e^{-\delta\mu_{\max}/3R} & (\text{if } \delta > 1) \end{cases} \\
\Pr\left[\sum_{i=1}^k X_i \le (1-\delta)\mu_{\min}\right] &\le \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu_{\min}/R} \le e^{-\delta^2\mu_{\min}/2R} \quad (\text{if } \delta \le 1).
\end{aligned}$$

In modern research it has been increasingly useful to study concentration for sums of random *matrices*. Many uses have appeared in compressed sensing, machine learning, and randomized numerical linear algebra. There was a long literature proving matrix concentration bounds, culminating in the following result of Tropp, which shows that the Chernoff bound has a perfect generalization to matrices. The formal statement is syntactically almost identical to the previous theorem.

**Theorem 2** (Tropp). Let $X_1, \ldots, X_k$ be independent, random, symmetric, real matrices of size $d \times d$ with $0 \preceq X_i \preceq R \cdot I$. Let $\mu_{\min} \cdot I \preceq \sum_i \mathrm{E}[X_i] \preceq \mu_{\max} \cdot I$. Then, for all $\delta \in [0,1]$,

$$\begin{aligned}
\Pr\left[\lambda_{\max}(\textstyle\sum_{i=1}^k X_i) \ge (1+\delta)\mu_{\max}\right] &\le d \cdot \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mu_{\max}/R} \le d \cdot \begin{cases} e^{-\delta^2\mu_{\max}/3R} & (\text{if } \delta \le 1) \\ e^{-\delta\mu_{\max}/3R} & (\text{if } \delta > 1) \end{cases} \\
\Pr\left[\lambda_{\min}(\textstyle\sum_{i=1}^k X_i) \le (1-\delta)\mu_{\min}\right] &\le d \cdot \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu_{\min}/R} \le d \cdot e^{-\delta^2\mu_{\min}/2R} \quad (\text{if } \delta \le 1).
\end{aligned}$$

Portions of these notes are based on scribe notes written by Zachary Drudi.

Here $\lambda_{\max}$ and $\lambda_{\min}$ respectively refer to the maximum and minimum eigenvalues of their argument. We are also using the partial ordering $\preceq$ on symmetric matrices defined by $A \preceq B$ if and only if $B - A$ is positive semi-definite (i.e. $\lambda_{\min}(B - A) \geq 0$). This is called the Löwner ordering.

We will use Theorem 2 in the next lecture to construct "graph sparsifiers" — subgraphs which approximation the graph in a very strong sense. Sparsifying the graph is a very useful subroutine for fast algorithms, such as fast Laplacian solvers.

Although the Chernoff bound is quite powerful (and certainly useful), its proof is quite straightforward, as Chernoff himself remarks in the quotation above. The matrix Chernoff bound is also quite comprehensible, especially since its proof has a very similar structure to the scalar bound. The main difficulty is understanding what scalar inequalities have generalizations to matrices. The field that studies such inequalities is called *matrix analysis*.

## 2 Matrix Analysis

In this section we briefly discuss some results that we will need.

**Definition 3** (Spectral mapping). Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We extend $f$ to a new function $f(A)$ on symmetric matrices by applying $f$ to the eigenvalues of $A$. That is, let $A = UDU^T$ be the spectral decomposition of $A$, where $U$ is orthogonal and $D$ is diagonal. Define $f(A) = Uf(D)U^T$, where $f(D)$ is the diagonal matrix with $f(D)_{i,i} = f(D_{i,i})$.

We will primarily be interested in the case $f = \exp$ or $f = \log$.

**Definition 4.** A function $f : \mathbb{R} \to \mathbb{R}$ is

- ***Operator monotone*** if $f(A) \succeq f(B)$ whenever $A \succeq B$.
- ***Operator concave*** if $f\big((1-x)A + xB\big) \succeq (1-x)f(A) + xf(B)$ for all $x \in [0,1]$ and all $A, B$.

In many ways, the Löwner ordering interacts nicely with the algebra of matrices and with spectral mapping. Many familiar scalar inequalities generalize to the Löwner ordering. However, many inequalities that one might hope to be true are actually false, so much care is needed!

### 2.1 The pitfalls of matrix analysis

**Hope 1.** If $f : \mathbb{R} \to \mathbb{R}$ is monotone (on a certain interval), is $f$ also operator monotone (for matrices whose eigenvalues lie in that interval)?

This is false. Consider function $f(x) = x^2$, which is monotone on the positive reals. Define $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. Then $A \preceq B$ since $B - A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, but $f(A) \not\preceq f(B)$ since $f(B) - f(A) = \left(\begin{smallmatrix} 3 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, which has a negative eigenvalue. The same $A$ and $B$ also show that exp is not operator monotone.

**Hope 2.** If $f : \mathbb{R} \to \mathbb{R}$ is concave (on a certain interval), is it also operator concave (for matrices whose eigenvalues lie in that interval)?

Also false. Consider $f(x) = -x^3$, which is concave on the positive reals. The same matrices $A$ and $B$ provide a counterexample with $x = 0.5$.

Before the reader becomes too dismayed, the take-away lesson should not be that matrix analysis

is useless. We will see next that there are many useful and powerful inequalities in matrix analysis. The lesson is that one much carefully determine which matrix inequalities are true before using them!

## 2.2 Some inequalities of matrix analysis

Proofs for the results of this section may be found in my 2013 lecture notes, with the exception of Theorem 9.

**Claim 5.** Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ satisfy $f(x) \leq g(x)\ \forall x \in [l, u]$. Suppose $A$ is symmetric and the eigenvalues of $A$ all lie in $[l, u]$. Then $f(A) \preceq g(A)$.

**Claim 6.** If $X$ and $Y$ are random matrices and $X \preceq Y$, then $\mathrm{E}[X] \preceq \mathrm{E}[Y]$.

**Claim 7** (Weyl's Monotonicity Theorem). Let $A$ and $B$ be symmetric, $n \times n$ matrices. Let $\lambda_i(A)$ denote the $i^{\text{th}}$ largest eigenvalue of $A$. If $A \preceq B$, then $\lambda_i(A) \leq \lambda_i(B)$ for all $i$.

**References:** Bhatia page 63.

**Corollary 8.** If $f$ is monotone, then so is the composition $\operatorname{tr} f$ defined on matrices.

**Theorem 9** (Löwner). log is operator concave.

**References:** Carlen Theorem 2.6, Horn and Johnson Exercise 6.6.18.

## 2.3 A commutative multiplication operation

One annoyance with matrix multiplication is that it is not commutative. We will define a new commutative multiplication operation that turns out to be useful.

**Definition 10.** If $A, B$ are positive definite, define $A \odot B = \exp(\log(A) + \log(B))$.

**References:** Warmuth-Kuzmin Section 4.

Note that if $A$ and $B$ commute then $A \odot B$ is the usual product $AB$.

**Theorem 11** (Lieb). Fix any symmetric $H$. The map $A \mapsto \operatorname{tr} \exp\big(\log(A) + H\big)$ is concave on the set of positive definite matrices.

**References:** Lieb 1973, Epstein 1973, Ohya-Petz 1993 Theorem 3.7.

**Corollary 12.** For any fixed $B$, the map $A \mapsto \operatorname{tr}(A \odot B)$ is concave.

Proof: $\operatorname{tr}(A \odot B) = \operatorname{tr} \exp(\log A + \log B)$. Apply Lieb's theorem with $H = \log B$. □

**Corollary 13.** Let $B$ be fixed, and $A$ a random matrix. Then $\mathrm{E}[\operatorname{tr}(A \odot B)] \leq \operatorname{tr}(\mathrm{E}[A] \odot B)$.

Proof: Apply Jensen's inequality. □

**Corollary 14.** Let $A_1, \ldots, A_k$ be independent random positive definite matrices. Then

$$\mathrm{E}[\operatorname{tr}(A_1 \odot \ldots \odot A_k)] \ \leq \ \operatorname{tr}(\mathrm{E}[A_1] \odot \ldots \odot \mathrm{E}[A_k]).$$

Proof: By symmetry Corollary 12 and Corollary 13 are also true if the roles of $A$ and $B$ are swapped. Thus we may inductively apply Corollary 13 to obtained the desired result. □

## 3 The Chernoff Bound

We now prove only the first inequality of Theorem 1.

**Claim 15.**

$$\Pr\left[\sum_{i=1}^{k} X_i \geq t\right] \;\leq\; \inf_{\theta>0} e^{-\theta t} \cdot \prod_{i=1}^{k} \mathrm{E}\left[e^{\theta X_i}\right].$$

Proof: Fix $\theta > 0$.

$$\begin{array}{rcll} \Pr\left[\sum_i X_i \geq t\right] & = & \Pr\left[\sum_i \theta X_i \geq \theta t\right] & \\ & = & \Pr\left[\exp(\sum_i \theta X_i) \geq \exp(\theta t)\right] & \text{(monotonicity of } e^x\text{)} \\ & \leq & e^{-\theta t} \cdot \mathrm{E}\left[\exp(\sum_i \theta X_i)\right] & \text{(Markov's inequality)} \end{array}$$

This expectation can be simplified:

$$\mathrm{E}\left[\exp(\textstyle\sum_i \theta X_i)\right] \;=\; \mathrm{E}\left[\prod_i e^{\theta X_i}\right] \;=\; \prod_i \mathrm{E}\left[e^{\theta X_i}\right] \qquad \text{(by independence).}$$

Combining these proves the claim. □

**Claim 16.** Let $X$ be a random variable with $0 \leq X \leq 1$. Then

$$\mathrm{E}\left[e^{\theta X}\right] \;\leq\; 1 + (e^{\theta} - 1) \cdot \mathrm{E}[X].$$

Proof: For $x \in [0,1]$ we have $e^{\theta x} \leq 1 + (e^{\theta} - 1) \cdot x$, by convexity of the left-hand side. Since $X \in [0,1]$,

$$\begin{array}{rcl} e^{\theta X} & \leq & 1 + (e^{\theta} - 1) \cdot X \\ \Longrightarrow \qquad \mathrm{E}\left[e^{\theta X}\right] & \leq & 1 + (e^{\theta} - 1) \cdot \mathrm{E}[X], \end{array}$$

since inequalities are preserved under taking expectation. □

**Proof** (of Chernoff Upper Bound). Without loss of generality $R = 1$.

$$\begin{array}{rcll} \prod_{i=1}^{k} \mathrm{E}\left[e^{\theta X_i}\right] & \leq & \prod_{i=1}^{k} \left(1 + (e^{\theta} - 1) \cdot \mathrm{E}[X_i]\right) & \text{(by Claim 16)} \\ & = & \exp\left(\sum_{i=1}^{k} \log\left(1 + (e^{\theta} - 1) \cdot \mathrm{E}[X_i]\right)\right) & \\ & \leq & \exp\left(\sum_{i=1}^{k} (e^{\theta} - 1) \cdot \mathrm{E}[X_i]\right) & \text{(using } \log(1+x) \leq x\text{)} \\ & \leq & \exp\left((e^{\theta} - 1)\mu_{\max}\right) & \end{array}$$

Applying Claim 15 with $t = (1+\delta)\mu_{\max}$ and $\theta = \ln(1+\delta)$

$$\begin{aligned} \Pr\left[\sum_i X_i \geq (1+\delta)\mu_{\max}\right] &\leq \exp\left(-\ln(1+\delta) \cdot (1+\delta)\mu_{\max}\right) \cdot \exp(\delta \cdot \mu_{\max}) \\ &= \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu_{\max}} \end{aligned}$$

■

# 4 Tropp's Matrix Chernoff Bound

We now prove only the first inequality of Theorem 2.

**Claim 17.**

$$\Pr\left[\lambda_{\max}\Big(\sum_{i=1}^{k} X_i\Big) \geq t\right] \;\leq\; \inf_{\theta>0} e^{-\theta t} \cdot \operatorname{tr}\left(\bigodot_{i=1}^{k} \mathrm{E}\left[e^{\theta X_i}\right]\right).$$

Proof: Fix $\theta > 0$.

$$\begin{aligned}
\Pr\left[\lambda_{\max}(\textstyle\sum_i X_i) \geq t\right] &= \Pr\left[\lambda_{\max}(\textstyle\sum_i \theta X_i) \geq \theta t\right] && \text{(homogeneity of max eigenvalue)}\\
&= \Pr\left[\exp\left(\lambda_{\max}(\textstyle\sum_i \theta X_i)\right) \geq \exp(\theta t)\right] && \text{(monotonocity of } e^x\text{)}\\
&\leq e^{-\theta t} \cdot \mathrm{E}\left[\exp\left(\lambda_{\max}(\textstyle\sum_i \theta X_i)\right)\right] && \text{(Markov's inequality)}
\end{aligned}$$

We can bound the maximum eigenvalue by a trace:

$$\begin{aligned}
\exp\left(\lambda_{\max}(\textstyle\sum_i \theta X_i)\right) &= \lambda_{\max}\left(\exp(\textstyle\sum_i \theta X_i)\right) && \text{(definition of matrix exponentiation)}\\
&\leq \operatorname{tr}\left(\exp(\textstyle\sum_i \theta X_i)\right) && \text{(max eigenvalue} \leq \text{sum of eigenvalues)}
\end{aligned}$$

Taking the expectation gives the bound:

$$\Pr\left[\lambda_{\max}(\textstyle\sum_i X_i) \geq t\right] \;\leq\; e^{-\theta t} \cdot \mathrm{E}\left[\operatorname{tr}\left(\exp(\textstyle\sum_i \theta X_i)\right)\right].$$

This expectation can be bounded:

$$\begin{aligned}
\mathrm{E}\left[\operatorname{tr}\left(\exp(\textstyle\sum_i \theta X_i)\right)\right] &= \mathrm{E}\left[\operatorname{tr}\left(\exp(\textstyle\sum_i \log A_i)\right)\right] && (\text{let } A_i = e^{\theta X_i})\\
&= \mathrm{E}\left[\operatorname{tr}(A_1 \odot \cdots \odot A_k)\right] && \text{(definition of } \odot\text{)}\\
&\leq \operatorname{tr}\left(\mathrm{E}\left[A_1\right] \odot \cdots \odot \mathrm{E}\left[A_k\right]\right) && \text{(by Corollary 14)}
\end{aligned}$$

Combining these inequalities proves the claim. □

**Claim 18.** Let $X$ be a random symmetric $d \times d$ matrix with $0 \preceq X \preceq I$. Then

$$\mathrm{E}\left[e^{\theta X}\right] \;\preceq\; I + (e^{\theta} - 1) \cdot \mathrm{E}\left[X\right].$$

Proof: For $x \in [0,1]$ we have $e^{\theta x} \leq 1 + (e^{\theta} - 1) \cdot x$, by convexity of the left-hand side. Since $X$ has all eigenvalues in $[0,1]$, Claim 5 gives

$$\begin{aligned}
e^{\theta X} &\preceq I + (e^{\theta} - 1) \cdot X\\
\Longrightarrow \qquad \mathrm{E}\left[e^{\theta X}\right] &\preceq I + (e^{\theta} - 1) \cdot \mathrm{E}\left[X\right],
\end{aligned}$$

since the Löwner ordering is preserved under taking expectation by Claim 6. □

**Proof** (of Matrix Chernoff Upper Bound). Without loss of generality $R = 1$. Our first observation is a bound for a sum of logs:

$$\begin{aligned}
\textstyle\sum_{i=1}^{k} \log \mathrm{E}\left[e^{\theta X_i}\right] &= k \cdot \textstyle\sum_{i=1}^{k} \frac{1}{k} \log \mathrm{E}\left[e^{\theta X_i}\right]\\
&\preceq k \cdot \log\left(\textstyle\sum_{i=1}^{k} \frac{1}{k} \mathrm{E}\left[e^{\theta X_i}\right]\right) \quad \text{(by Theorem 9)} \qquad (1)
\end{aligned}$$

Next:

$$
\begin{aligned}
&\operatorname{tr}\left(\mathrm{E}\left[\,e^{\theta X_1}\,\right] \odot \cdots \odot \mathrm{E}\left[\,e^{\theta X_k}\,\right]\right) \\
&\quad= \operatorname{tr}\exp\left(\textstyle\sum_{i=1}^{k} \log \mathrm{E}\left[\,e^{\theta X_i}\,\right]\right) && \text{(definition of } \odot\text{)} \\
&\quad\overset{(a)}{\leq} \operatorname{tr}\exp\left(k \cdot \log\left(\textstyle\sum_{i=1}^{k} \frac{1}{k} \mathrm{E}\left[\,e^{\theta X_i}\,\right]\right)\right) && \text{(by (1) and Corollary 8)} \\
&\quad\leq d \cdot \lambda_{\max}\left(\exp\left(k \cdot \log\left(\textstyle\sum_{i=1}^{k} \frac{1}{k} \mathrm{E}\left[\,e^{\theta X_i}\,\right]\right)\right)\right) && \text{(sum of eigenvalues} \leq d \text{ times maximum)} \\
&\quad= d \cdot \exp\left(k \cdot \log \lambda_{\max}\left(\textstyle\sum_{i=1}^{k} \frac{1}{k} \mathrm{E}\left[\,e^{\theta X_i}\,\right]\right)\right) && \text{(definition of spectral mapping)} \\
&\quad\leq d \cdot \exp\left(k \cdot \log \lambda_{\max}\left(I + \textstyle\sum_{i=1}^{k} \frac{1}{k}(e^{\theta}-1)\,\mathrm{E}\left[\,X_i\,\right]\right)\right) && \text{(by Claim 18 and Claim 7)} \\
&\quad= d \cdot \exp\left(k \cdot \log\left(1 + \tfrac{e^{\theta}-1}{k} \lambda_{\max}\left(\textstyle\sum_{i=1}^{k} \mathrm{E}\left[\,X_i\,\right]\right)\right)\right) \\
&\quad\leq d \cdot \exp\left((e^{\theta}-1) \cdot \lambda_{\max}\left(\textstyle\sum_{i=1}^{k} \mathrm{E}\left[\,X_i\,\right]\right)\right) && \text{(using } \log(1+x) \leq x\text{)} \\
&\quad\leq d \cdot \exp\left((e^{\theta}-1) \cdot \mu_{\max}\right)
\end{aligned}
$$

Apply Claim 17 with $t = (1+\delta)\mu_{\max}$ and $\theta = \ln(1+\delta)$:

$$
\begin{aligned}
\Pr\left[\,\lambda_{\max}\left(\textstyle\sum_i X_i\right) \geq (1+\delta)\mu_{\max}\,\right] &\leq \exp\left(-\ln(1+\delta) \cdot (1+\delta)\mu_{\max}\right) \cdot \left(d \cdot \exp(\delta \cdot \mu_{\max})\right) \\
&= d \cdot \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu_{\max}}
\end{aligned}
$$

■

**Remark.** In inequality (a), it is *not* true that exp is operator monotone, but it is true (by Corollary 8) that tr exp is monotone.

## References


[1] R. Bhatia. *Matrix Analysis*. Springer, 1997.

[2] E. Carlen. Trace inequalities and quantum entropy: An introductory course. *Contemporary Mathematics*, 529, 2009.

[3] H. Epstein. Remarks on Two Theorems of E. Lieb. *Communications in Mathematical Physics*, 31:317–325, 1973.

[4] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[5] E. H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Advances in Mathematics*, 11:267–288, 1973.

[6] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer-Verlag, 1993.

[7] J. A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 2011.

[8] M. K. Warmuth and D. Kuzmin. Bayesian generalized probability calculus for density matrices. *Machine Learning*, 78(1), 2010.