Goldon-Thompson via pinching inequality

Saket Choudhary

saketkc@gmail.com

October 20, 2018

Given A, B hermitian matrices, Golden-Thompson inequality [1, 2] states that:

$$\text{tr}\left[\exp\left((A+B)\right)\right] \leq \text{tr}\left[\exp\left((A)\right)\exp\left((B)\right)\right]$$

It is trivial if A, B commute i.e. $AB = BA$. A quick example would be $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $B = \mathbb{I}_2$, while it does not hold for $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ as $AB \neq BA$.

There have been multiple proofs [3], but they are not so straight forward to follow. Sutter *et al.* [4] presented a more intutive proof using spectral pinching. This document summarizes their approach and at places is slightly more elaborated than the original version that appeared in [4]. We will stick to real matrices here.

Spectral Pinching Method

Consider a square complex matrix A partitioned as a $r \times r$ block matrix: $A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{rr} \\ \vdots & \ddots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{pmatrix}$.

We can decompose this matrix intro two matrices comprising the diagonal and the off-diagonal elements respectively.

$$A = A_D + A_{\tilde{D}}$$

$$A_D = \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{rr} \end{pmatrix}$$

$$A_{\tilde{D}} = \begin{pmatrix} 0 & A_{12} & \cdots & A_{1r} \\ A_{21} & 0 & \cdots & A_{2r} \\ \vdots & \ddots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & 0 \end{pmatrix}$$

A_D is called a *pinching* of A. The simplest case is that of A_{ij} being 1 dimensional that we will be using here.

Any positive semi-definite matrix A can be written as $A = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$ where λ_i are n distinct eigen values of A. P_{λ_i} are orthogonal projectors such that $\sum_{i=1}^{n} P_{\lambda_i} = \mathbb{I}$ and hence $P_{\lambda_i}^2 = P_{\lambda_i}$

The *spectral pinching map* of A is then given by:

$$\mathcal{P}_A : X \mapsto \sum_{\lambda} P_\lambda X P_\lambda$$

The entire idea here is to use some form of convex combination resulting in an averaging operation. The pinching map in turn has the following properties:

(i) $\mathcal{P}_A[X]A = A\mathcal{P}_A[X]$

(ii) $\mathrm{tr}[\mathcal{P}_A[X]A] = \mathrm{tr}[XA]$

(iii) $\mathcal{P}_A[X] \geq \frac{1}{n} X$

Lemma 1. $\mathcal{P}_A[X]A = A\mathcal{P}_A[X]$

Proof:

$P_{\lambda_i}^2 = P_{\lambda_i}$ and $P_{\lambda_i} \perp P_{\lambda_j}$

$$\begin{aligned}
\mathcal{P}_A[X]A &= \sum_{i=1}^{n} P_{\lambda_i} X P_{\lambda_i} \sum_{i=1}^{n} \lambda_i P_{\lambda_i} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_j P_{\lambda_i} X P_{\lambda_i} P_{\lambda_j} \\
&= \sum_{i=1}^{n} \lambda_i P_{\lambda_i} X P_{\lambda_i} P_{\lambda_i} \\
&= \sum_{i=1}^{n} \lambda_i P_{\lambda_i} X P_{\lambda_i} \\
&= \sum_{i=1}^{n} \lambda_i P_{\lambda_i} P_{\lambda_i} X P_{\lambda_i} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_j P_{\lambda_j} P_{\lambda_i} X P_{\lambda_i} \\
&= \sum_{j=1}^{n} \lambda_j P_{\lambda_j} \sum_{i=1}^{n} P_{\lambda_i} X P_{\lambda_i} \\
&= A\mathcal{P}_A[X]
\end{aligned}$$

Lemma 2. $tr[\mathcal{P}_A[X]A] = tr[XA]$

Proof:

$$
\begin{aligned}
\text{tr}[\mathcal{P}_A[X]A] &= \text{tr}[\sum_{i=1}^{n} P_{\lambda_i} X P_{\lambda_i} \sum_{i=1}^{n} \lambda_i P_{\lambda_i}] \\
&= \text{tr}[\sum_{i=1}^{n} \lambda_i P_{\lambda_i} X P_{\lambda_i}] \\
\text{tr}[P_{\lambda_i} X P_{\lambda_i}] &= \sum_j (P_{\lambda_i} X)_{jj} \\
\implies \text{tr}[\mathcal{P}_A[X]A] &= \sum_{i=1}^{n} \lambda_i \sum_j (P_{\lambda_i} X)_{jj} \\
&= \text{tr}\ [AX]
\end{aligned}
$$

Lemma 3. $\mathcal{P}_A[X] \geq \frac{1}{n} X$

Proof:

Proving this part is probably the trickiest among the four lemmata here, but is the entire key behind deducing the final Golden-Thompson inequality. Consider a unitary matrix U_y defined as $U_y = \sum_{u=1}^{n} e^{i2\pi yu/n} P_{\lambda_u}$. It is easy to verify that $U_y U_y^* = \mathbb{I}$ as $\sum_{i=1}^{n} P_{\lambda_i} = \mathbb{I}$. Also $U_y \geq 0$ and $U_n = \mathbb{I}$

$$
\begin{aligned}
\sum_{y=1}^{n} U_y X U_y^* &= \sum_{y=1}^{n} \sum_{s=1}^{n} \sum_{t=1}^{n} e^{i2\pi ys/n} P_{\lambda_s} X P_{\lambda_t} e^{-i2\pi yt/n} \\
&= \sum_{s=1}^{n} \sum_{t=1}^{n} P_{\lambda_s} X P_{\lambda_t} \sum_{y=1}^{n} e^{i2\pi y(s-t)/n} \\
&= \sum_{s=1}^{n} \sum_{t=1}^{n} P_{\lambda_s} X P_{\lambda_t} (n \mathbf{1}_{\{s=t\}}) \\
&= n \sum_{s=1}^{n} P_{\lambda_s} X P_{\lambda_s} \\
\implies \mathcal{P}_A[X] &= \sum_{s=1} P_{\lambda_s} X P_{\lambda_s} = \frac{1}{n} \sum_{y=1}^{n} U_y X U_y^* \\
&\geq \frac{1}{n} X
\end{aligned}
$$

Now once (iii) is proved, the rest of the steps for proving the GT are straightforward. For a semi-positive definite $d \times d$ matrix A, we have the following lemma.

Lemma 4. $|spec(A^{\bigotimes m})| \leq O(poly(m))$

Proof:

The number of eigen values for $A^{\bigotimes m}$ is bounded by the number of the number of possible possible combinations of a sequence of d sy,bols (the maximum possible distinct eigen values of A) of length m which is given by $\binom{m+d-1}{d-1} \leq \frac{(m+d-1)^{d-1}}{(d-1)\,!} = O(\text{poly}(m))$

Golden Thompson Inequality

Given positive definite matrices A, B and using the facts that $\exp()$ and $\mathrm{tr}[\exp()]$ are operator monotone:

$$
\begin{aligned}
\log \mathrm{tr}[\exp(\log A + \log B)] &= \frac{1}{m} \log \mathrm{tr}[\exp\left(\log A^{\bigotimes m} + \log B^{\bigotimes m}\right)] \\
&\leq \frac{1}{m} \log \mathrm{tr}[\exp\left(\log \mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}]|\mathrm{spec}(A^{\bigotimes m})| + \log B^{\bigotimes m}\right)] && \text{Using Lemma 3} \\
&= \frac{1}{m} \log \mathrm{tr}[\exp\left(\log \mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}] + \log|\mathrm{spec}(A^{\bigotimes m})| + \log B^{\bigotimes m}\right)] \\
&= \frac{1}{m} \log \mathrm{tr}[\exp\left(\log \mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}] + \log B^{\bigotimes m}\right)] + \frac{1}{m}|\mathrm{spec}(A^{\bigotimes m})| \\
&\leq \frac{1}{m} \log \mathrm{tr}[\exp\left(\log \mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}] + \log B^{\bigotimes m}\right)] + \frac{O(\mathrm{poly}(m))}{m} && \text{Using Lemma 4} \\
&= \frac{1}{m} \log \mathrm{tr}[\exp\left(\log \mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}]B^{\bigotimes m}\right)] + \frac{O(\mathrm{poly}(m))}{m} && \text{Using Lemma 1} \\
&= \frac{1}{m} \log \mathrm{tr}[\mathcal{P}_{B^{\otimes m}}[A^{\bigotimes m}]B^{\bigotimes m}] + \frac{O(\mathrm{poly}(m))}{m} \\
&= \frac{1}{m} \log \mathrm{tr}[A^{\bigotimes m}B^{\bigotimes m}] + \frac{O(\mathrm{poly}(m))}{m} && \text{Using Lemma 2} \\
&= \log \mathrm{tr}[AB] + \frac{O(\mathrm{poly}(m))}{m}
\end{aligned}
$$

$$
\begin{aligned}
\implies \log \mathrm{tr}[\exp(\log A + \log B)] &\leq \log \mathrm{tr}(AB) \\
\implies \mathrm{tr}[\exp(\log A + \log B)] &\leq \mathrm{tr}(AB) \\
\implies \mathrm{tr}[\exp(A + B)] &\leq \mathrm{tr}(\exp(A)\exp(B))
\end{aligned}
$$

References

[1] S. Golden, "Lower bounds for the helmholtz function," *Physical Review*, vol. 137, no. 4B, p. B1127, 1965.

[2] C. J. Thompson, "Inequality with applications in statistical mechanics," *Journal of Mathematical Physics*, vol. 6, no. 11, pp. 1812–1813, 1965.

[3] P. J. Forrester and C. J. Thompson, "The golden-thompson inequality: Historical aspects and random matrix applications," *Journal of Mathematical Physics*, vol. 55, no. 2, p. 023503, 2014.

[4] D. Sutter, M. Berta, and M. Tomamichel, "Multivariate trace inequalities," *Communications in Mathematical Physics*, vol. 352, no. 1, pp. 37–58, 2017.