Concentration inequalities for sums of independent r.v.s

$\lesssim \iff \lesssim$

$X_1, X_2, \dots, X_n$ indep. random variables ✓

$X := X_1 + X_2 + \dots + X_n$ ✓

{ Assume that $\boxed{|X_i - \mathbb{E}X_i| \le L}$ (w/ prob. one) $\forall i = 1, \dots, n$

$$\boxed{\mathbb{P}[X \le \mathbb{E}X + t\sqrt{n}] \ge 1 - e^{-t^2/2L^2}}$$

$$\boxed{\mathbb{P}[X \ge \mathbb{E}X + t\sqrt{n}] \le e^{-t^2/2L^2}} \quad \forall t \ge 0$$

Can assume $\mathbb{E}X_i = 0$ and $|X_i| \le L$ $\forall i$

(1)

$$\mathbb{P}[X \ge t\sqrt{n}] = \mathbb{P}[e^{\beta X} \ge e^{\beta t\sqrt{n}}] \le \left[\frac{\mathbb{E}[e^{\beta X}]}{e^{\beta t\sqrt{n}}}\right] \quad (\beta > 0)$$

independence

$$\mathbb{E}[e^{\beta X}] = \mathbb{E}[e^{\beta(X_1 + \dots + X_n)}] = \left[\underbrace{\prod_{i=1}^{n} \mathbb{E}[e^{\beta X_i}]}\right] \quad (2)$$

$$\left[\quad \underbrace{1 + x} \le e^x \le \underbrace{1 + x + x^2} \quad \forall x \in [-1, 1] \right. \quad (3)$$

$$\underline{\mathbb{E}[e^{\beta X_i}]} \le \mathbb{E}[1 + \underline{\beta X_i} + \beta^2 X_i^2] \qquad \beta \le 1/L$$

$$= 1 + \mathbb{E}[\beta^2 X_i^2] \le 1 + \beta^2 L^2 \le \underline{e^{\beta^2 L^2}}$$

$\boxed{X \le nL}$

$$\mathbb{P}[X \ge t\sqrt{n}] \le e^{\boxed{-\beta t\sqrt{n}}} e^{n\beta^2 L^2} \qquad \forall \beta \le 1/L$$

$$t \le 2L\sqrt{n}$$

$$\beta := \frac{t}{2L^2\sqrt{n}} \quad \le \quad e^{-t^2/2L^2}$$

$A_1, A_2, \dots, A_n \in M_d(\mathbb{R})$ $\quad A_i = A_i^T \quad \forall i$

independent random symmetric matrices

$A := A_1 + A_2 + \dots + A_n$ $\quad \boxed{-L \cdot I \preceq A_i - \mathbb{E}A_i \preceq L \cdot I}$ $\forall i$

$$\boxed{\mathbb{P}[A \preceq t\sqrt{n} \cdot I] \geq 1 - e^{-t^2/2L^2}}$$

$\boxed{\mathbb{E}A_i = 0}$ $\quad -L \cdot I \preceq A_i \preceq L \cdot I$ $\quad \boxed{\|A_i\| \leq L \quad \forall i}$

$$\boxed{\mathbb{P}[\lambda_{max}(A) \geq t\sqrt{n}] \leq d e^{-t^2/2L^2}}$$

$$\|A\| = \max_{\substack{x \in \mathbb{R}^d \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2}$$

$$A_i = \begin{bmatrix} a_{11}^{(i)} & & \\ & \ddots & \\ & & a_{dd}^{(i)} \end{bmatrix} \qquad A = \begin{bmatrix} \boxed{a_{11}^{(1)} + \dots + a_{11}^{(n)}} & & \\ & \ddots & \\ & & \boxed{a_{dd}^{(1)} + \dots + a_{dd}^{(n)}} \end{bmatrix}$$

$\mathbb{E}A_i = 0 \quad \|A_i\| \leq L \quad \forall i$

$$\mathbb{P}[\lambda_{max}(A) \geq t\sqrt{n}] \leq \mathbb{P}[e^{\beta \lambda_{max}(A)} \geq e^{\beta t \sqrt{n}}]$$

$$\leq \frac{\mathbb{E}[e^{\beta \lambda_{max}(A)}]}{e^{\beta t\sqrt{n}}}$$

$$\left[\underline{\lambda_{max}(A_1 + \dots + A_n)} \stackrel{?}{=} \lambda_{max}(A_1) + \dots + \lambda_{max}(A_n) \right]$$

$$\mathbb{P}[\lambda_{max}(A) \geq t\sqrt{n}] \leq \mathbb{P}[\lambda_1(A) + \dots + \lambda_n(A) \geq t\sqrt{n}]$$

$$\mathbb{P}[\lambda_{max}(A) \geq t\sqrt{n}] = \mathbb{P}[\lambda_{max}(e^{\beta A}) \geq e^{\beta t\sqrt{n}}]$$

$f : \mathbb{R} \to \mathbb{R}$

$f : I \to \mathbb{R}$

$spec(A) \subseteq I$

$$A = U \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} U^T$$

$$f(A) := U \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} U^T$$

$$A = \sum_i \lambda_i \underline{\underline{u_i u_i^T}} \qquad f(A) := \sum_i f(\lambda_i) u_i u_i^T$$

$$A^n = U \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_d^n \end{pmatrix} U^T$$

$f(x) = x^n$

$f(A)$

$$e^x = \sum_{n \geq 0} \frac{x^n}{n!} \qquad e^A = U \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_d} \end{pmatrix} U^T$$

$$\boxed{e^A = \sum_{n \geq 0} \frac{A^n}{n!}}$$

$A^0, A^1, \ldots, \ldots$

$\lambda_{max}(e^{\beta(A_1 + \cdots + A_n)})$

$$\mathbb{P}[\lambda_{max}(A) \geq t\sqrt{n}]$$

$$= \mathbb{P}[\lambda_{max}(e^{\beta A}) \geq e^{\beta t\sqrt{n}}]$$

$$\leq \mathbb{P}[\mathrm{Tr}(e^{\beta A}) \geq e^{\beta t\sqrt{n}}]$$

$$\leq \frac{\mathbb{E}[\mathrm{Tr}(e^{\beta A})]}{e^{\beta t\sqrt{n}}}$$

$$\boxed{\mathrm{Tr}(A) = \sum_i A_{ii}}$$

$$\mathrm{Tr}(AB) = \sum_{i,j} A_{ij} B_{ji}$$

$$= \mathrm{Tr}(BA)$$

$$\mathrm{Tr}(PAP^{-1}) = \mathrm{Tr}(A)$$

$\forall$ invertible P

$$\mathbb{E}\left[\mathrm{Tr}\left(e^{\beta(A_1+A_2+\cdots+A_n)}\right)\right]$$

$\mathrm{Tr}(A) = \lambda_1+\cdots+\lambda_d$

$\mathrm{Tr}(UDU^T) = \mathrm{Tr}(D)$

$e^{A+B} \neq e^A e^B$ unless A, B commute

False: $e^{A+B} \preceq e^A e^B$ (doesn't always hold)

Golden-Thompson: For all $A, B \in \mathbb{M}_d(\mathbb{R})$ symmetric,

$$\mathrm{Tr}(e^{A+B}) \le \mathrm{Tr}(e^A e^B)$$

$$\overset{(GT)}{\le} \mathbb{E}\left[\mathrm{Tr}\left(e^{\beta(A_1+\cdots+A_{n-1})} e^{\beta A_n}\right)\right]$$

$$= \mathbb{E}_{A_1,\dots,A_{n-1}}\left[\mathrm{Tr}\left(\underbrace{e^{\beta(A_1+\cdots+A_{n-1})}}_{N}\ \underbrace{\mathbb{E}_{A_n} e^{\beta A_n}}_{M}\right)\right]$$

$|\lambda_1(N)|+\cdots+|\lambda_d(N)|$

$\mathrm{Tr}(e^{(0\cdots 0)}) = \mathrm{Tr}(I)$

M, N symmetric

$\mathrm{Tr}(MN) \le \|M\| \cdot \|N\|_1$

$\mathrm{Tr}(N)$ for $N \succeq 0$

$$\le \mathbb{E}_{A_1,\dots,A_{n-1}}\left[\mathrm{Tr}\left(e^{\beta(A_1+\cdots+A_{n-1})}\right)\right] \cdot \left\|\mathbb{E}_{A_n} e^{\beta A_n}\right\|$$

$$\le \cdots \le \left(\prod_{i=1}^{n} \left\|\mathbb{E}_{A_i} e^{\beta A_i}\right\|\right) \mathrm{Tr}(I)$$

$\cancel{\mathrm{Tr}(e^{\beta A_1} e^{\beta A_2} \cdots e^{\beta A_n})}$

$$= d \cdot \prod_{i=1}^{n} \left\|\mathbb{E}_{A_i} e^{\beta A_i}\right\|$$

$$\boxed{I + A \preceq e^A \preceq I + A + A^2} \qquad \|A\| \le 1$$

$$1 + \lambda \le e^\lambda \le 1 + \lambda + \lambda^2 \qquad \lambda \in [-1, 1]$$

$$\le d \cdot \prod_{i=1}^{n} \underbrace{\| \mathbb{E}(I + \beta A_i + \beta^2 A_i^2) \|}$$

$A \succeq cI$

$$\le \| I + \mathbb{E}[\beta^2 A_i^2] \| \qquad \|A_i\| \le L$$

$\langle u, Au \rangle \ge c\|u\|^2$

$$\le 1 + \beta^2 L^2 \qquad \|\beta A_i\| \le \beta L$$

$$\boxed{\mathbb{P}[\langle u, Au \rangle \ge t\sqrt{n}]} \le e^{\beta^2 L^2} \qquad \boxed{A_i} \qquad \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & \square \end{pmatrix}$$

$$\boxed{\frac{1}{n}\sum_{i,j} A_{ij}} \le \boxed{d \cdot e^{\beta^2 L^2 n}} \qquad \beta \le 1/L$$

$$\boxed{\mathbb{P}[\lambda_{max}(A) \ge t\sqrt{n}] \le d e^{-\beta t \sqrt{n}} e^{\beta^2 L^2 n}} \; = \underbrace{\qquad}$$

$$\boxed{\underline{GT}: \; Tr(e^{A+B}) \le Tr(e^A e^B)} \qquad \forall A, B \text{ symmetric}$$

$$e^x = \sum_{n \ge 0} \frac{x^n}{n!}, \text{ so } e^{A+B} = \sum_{n \ge 0} \frac{(A+B)^n}{n!}$$

e^{A+B}

$$(A+B)^n = \boxed{\sum_{i_1, i_2, \ldots, i_n \in \{0,1\}} A^{i_1} B^{1-i_1} A^{i_2} B^{1-i_2} \cdots A^{i_n} B^{1-i_n}}$$

$e^A e^B$ $\qquad$ $A^j B^{n-j}$

$\left(\sum \frac{A^i}{i!}\right)\left(\sum \frac{B^j}{j!}\right)$ $\qquad$ "$\mathrm{Tr}(ABA^2BAB^2) \le \mathrm{Tr}(A^4B^4)$"

Suppose A, B symmetric $\qquad$ $\boxed{\|A\|_2 = \sqrt{\mathrm{Tr}(A^TA)}}$

$\mathrm{Tr}(AB) \le \|A\| \cdot \|B\|_1$

$\to \boxed{\mathrm{Tr}(AB) \le \|A\|_2 \cdot \|B\|_2}$ $\qquad$ $\|A\|_2 = \sqrt{\sum |\lambda_i(A)|^2}$

$\boxed{\mathrm{Tr}((AB)^2)} = \boxed{\mathrm{Tr}(ABAB)} \le \|AB\|_2^2$

$= \mathrm{Tr}((AB)^T AB)$

$\lambda_1^3 \; \ldots \; \lambda_n^3$ $\qquad$ $= \mathrm{Tr}(B^TA^TAB)$

$A^3B \qquad ABA^2$ $\qquad$ $= \mathrm{Tr}(BAAB)$

$= \boxed{\mathrm{Tr}(A^2B^2)}$

$\mathrm{Tr}((AB)^4) = \mathrm{Tr}(AB\,AB\,AB\,AB) \le \mathrm{Tr}((A^2B^2)^2)$

$\boxed{\mathrm{Tr}(e^{A+B}) \le \mathrm{Tr}(e^Ae^B)}$ $\qquad$ $\le \mathrm{Tr}(A^4B^4)$

$$\boxed{\|A\|_p := \left(\mathrm{Tr}\left((A^TA)^{p/2}\right)\right)^{1/p}}$$