Theorem: We fix δ and we give a lower bound for the probability that a random matrix is δ- approximately scaled by the scaling factors (x, y) of $E[\tilde{A}]$. $P(\delta, M, N)$ = Probability that $\tilde{A}$ is δ-app. sc. by x, y

* Let's restrict ourselves to the case of doubly stochastic scaling of square matrices

Hoeffding (Boundedness):

$$\mathbb{P}(\mathcal{A}) = P(\delta, M) \geq 1 - 4M \exp\left(-\frac{\delta^2 M^{\cancel{2}}}{\cancel{M} \cdot C_p^2}\right)$$

independent of M

depends only on the bounds $[a, b]$

Let $\mathcal{A} = \{\tilde{A} - \delta\text{-approx. scaled}\}$

For any $\tilde{A} \in \mathcal{A}$ $\xrightarrow{\text{deterministic procedure}}$ $\tilde{x}, \tilde{y}$: Exact scaling factors that don't deviate too much from x, y pointwise.

Bound: $\frac{|\tilde{x}_i - x_i|}{x_i} \leq C_e \cdot \delta$

$\boxed{\text{Independent of N}}$

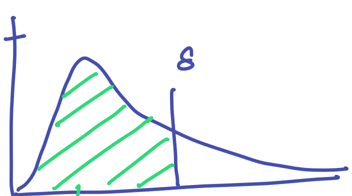
depends only on ele. bounds $[a, b]$

Question 1: Is $C_e \cdot \delta$ the best bound you can get ?

Let $\mathbb{B} = \left\{ \frac{|\tilde{x}_i - x_i|}{x_i} \leq C_e \cdot \delta \text{ and } \frac{|\tilde{y}_j - y_j|}{y_j} \leq C_e\, \delta \;\; \forall\, i, j \text{ for some } x, y \right\}$

$\mathbb{B} \supseteq \mathcal{A} \Rightarrow \boxed{\mathbb{P}(\mathbb{B}) > \mathbb{P}(\mathcal{A})}$

To estimate the Bound $\rightsquigarrow$ Fix a δ $\rightsquigarrow$ Estimator for $\mathbb{P}(\mathcal{A})$ from this estimate $\longrightarrow$ p (say)



scaling level

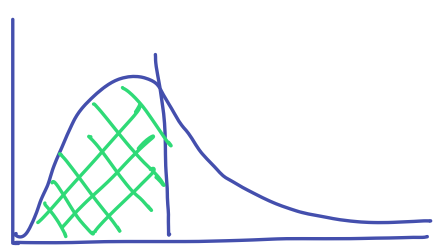
Or compute from this cutoff : normalised — Loose upper Bound ??

└ To study dependence on δ

▨ $= P(\delta, M) = \mathbb{P}(\mathcal{A})$

$\mathcal{A} \subseteq \mathbb{B}$

Area ⊠ $= \mathbb{P}(\mathcal{A})$



$\varepsilon(\tilde{x}, \tilde{y})$

normalise

Discrepancy relate to

checks how effective going from approximate scaling to actual scaling is.

This is different from what we want though

Estimating exact upper bound

How much bigger is $\mathbb{P}(B)$ than $\mathbb{P}(A)$

* This is what we want in practice ?

Question 2: Dependence on N: upper bound.

Empirically: change N

Fix δ ⇝ Change N and compute the maximum.

⇝ <u>weak upper bound</u>. $B(N)$

But probability also changes: Print that out.

For fixed probability: we can find δ_N using quantiles

Find Bounds: $B(\delta_N, N)$

---

§ Question 3: Dependence on a

Keep N fixed ⇝ plot the quantiles and see how approximate scaling changes

- Bound for 80% of them
- Look at approximately scaled and then estimate the bound from that → would give a very weak upper bound
  - Dealing with A instead of B → even though δ is fixed prob is increasing
  - $\|x\| = \|y\|$ ⇝ used to guess the bound

---

§ THEORETICAL DEPENDENCE OF δ ON N FOR FIXED PROBABILITY

Empirically: For a fixed probability of p

we would need $4N\exp(-K\delta^2 N) = O(1)$

$\delta^2 N = O(\log N)$ ⇝ $\delta = O\left(\sqrt{\frac{\log N}{N}}\right)$ → which also matches rate of convergence

Empirically : For both chi-squared and uniform : seems to resemble $O\left(\sqrt{\frac{\log N}{N}}\right)$

we get for scaling factors WHP ⇝ connection *

which is better $O\left(\sqrt{\frac{\log N}{N}}\right)$

* In practice : ↳ $\boxed{\text{How do you compare ? Curve fitting ?}}$

---

§ ASYMPTOTICS FROM LANDA'S PAPER: Doubly stochastic case

$$\varepsilon(\tilde{x}, \tilde{y}) = \max\left\{ \max_{i \in [N]} \frac{|x_i - \tilde{x}_i|}{x_i}, \max_{i \in [N]} \frac{|y_i - \tilde{y}_i|}{y_i} \right\}$$

Claim: Boundedness, + ty : assumptions carry forward.

$\exists$ scaling factors $\{\tilde{x}^{(N)}, \tilde{y}^{(N)}\}_{N \in \mathbb{N}}$ s.t

$$\varepsilon(\tilde{x}^{(N)}, \tilde{y}^{(N)}) = O_{whp}\left(\sqrt{\frac{\log N}{N}}\right)$$

↳ doubly stochastic case.

Question: How does dependence on N feature in this proof?

↳ Not through the approximate scaling bound's dependence on N

It is through choosing δ_N increasingly small enough to shrink the bound but while also ensuring the probability tends to 1.

Hence, $\sqrt{\frac{\log N}{N}}$ matches * ⇝ can this be improved or is it tight ?

---

§ PRACTICE ?

- To get scaling factors close to the ones we have.
- $\boxed{\text{Deterministic}}$ if we know if they are approximately scaled at the a suitable δ-level