Operations with Concentration Inequalities

Cosme Louart

Assistant Professor at CUHK Shenzhen

SCHOOL OF DATA SCIENCE

Operations with Concentration Inequalities

Cosme Louart

Assistant Professor at CUHK Shenzhen

SCHOOL OF DATA SCIENCE

Given two random variables: $X, Y \in \mathbb{R}$

$$\forall t > 0 : \quad \mathbb{P}\left(X \geq t\right) \leq \alpha(t)$$

$$\times$$

$$\forall t > 0 : \quad \mathbb{P}\left(Y \geq t\right) \leq \beta(t)$$

$$=$$

$$\forall t > 0 :$$

$$\mathbb{P}\left(X \cdot Y \geq t\right) \leq 2\, \alpha \boxtimes \beta(t)$$

Operations with Concentration Inequalities

Cosme Louart

Assistant Professor at CUHK Shenzhen

SCHOOL OF DATA SCIENCE

Given two random variables: $X, Y \in \mathbb{R}$

$$\forall t > 0 : \quad \mathbb{P}(X \geq t) \leq \alpha(t)$$

$$\times$$

$$\forall t > 0 : \quad \mathbb{P}(Y \geq t) \leq \beta(t)$$

$$=$$

"Parallel product"

$$\forall t > 0 : \quad \mathbb{P}(X \cdot Y \geq t) \leq 2\, \alpha \boxtimes \beta(t)$$

Operations with Concentration Inequalities

Cosme Louart

Assistant Professor at CUHK Shenzhen

SCHOOL OF DATA SCIENCE

- Given a random vector : $X \in \mathbb{R}^n$:
  $\forall t > 0 : \quad \mathbb{P}\left(|f(X) - \mathbb{E}[f(X)]| \geq t\right) \leq \alpha(t)$
  $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz

Operations with Concentration Inequalities

Cosme Louart

Assistant Professor at CUHK Shenzhen

SCHOOL OF DATA SCIENCE

Content

I - **Motivation:** Hanson Wright for Random Matrix Theory

II - Parallel Sum and Product.

$\alpha \boxplus \beta$? $\alpha \boxtimes \beta$? $\mathbb{P}(X + Y \geq t) \leq$?

III - Concentration in High Dimension

$\mathbb{P}(|f(X) - \mathbb{E}[f(X)]| \geq t) \leq \alpha(t)$, $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, Talagrand result
Concentration of Φ where $\|\Phi(Z) = \Phi(Z')\| \leq V\|Z - Z'\|$.

IV - Application to Hanson-Wright inequality

Large tail concentration, Random matrix hypothesis ?

I - **Motivation:** Hanson Wright for Random Matrix Theory

Given $x_1, \dots, x_n \sim \mathcal{N}(0, \Sigma)$, i.i.d. random vectors, note $X \equiv (x_1, \dots, x_n) \in \mathbb{R}^{n \times p}$.

Goal: Eigen value distribution of $\frac{1}{n} X X^T$: $\mu \equiv \frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i}$ **??**

Eigen values of $\frac{1}{p} X X^T$

$$\left(\text{Sp} \left(\frac{1}{p} X X^T \right) = \{ \lambda_1, \dots, \lambda_p \} \right)$$

- Correspondance $\mu \longleftrightarrow m : z \mapsto \int_{\mathbb{R}} \frac{1}{z - \lambda} d\mu(\lambda)$

"Steiltjes Transform" (similar to Cauchy Transform)

- Link with the *"Resolvent"*: $m(z) = \frac{1}{p} \text{Tr} Q(z)$, where $Q(z) \equiv \left(z I_p - \frac{1}{n} X X^T \right)^{-1}$.

Strategy: Find deterministic $\tilde{Q} \in \mathcal{M}_p$ such that $Q \approx \tilde{Q}$

I - **Motivation:** Hanson Wright for Random Matrix Theory

Goal: Approach $\mathbb{E}[Q] = \mathbb{E}\left[\left(zI_p - \frac{1}{n}XX^T\right)^{-1}\right]$

- Of course $\mathbb{E}[Q]$ far from $(zI_p - \Sigma)^{-1}$

$\Sigma \equiv \mathbb{E}\left[\frac{1}{n}XX^T\right] = \mathbb{E}[x_i x_i^T], \ \forall i \in [n]$

Solution: Look for $\tilde{Q} \equiv \left(zI_p - \frac{\Sigma}{1+\delta}\right)^{-1}$

δ to be determined

Given $A \in \mathcal{M}_p$, deterministic:

Dependence between Q and x_i

$$\mathrm{Tr}\left(A(\mathbb{E}[Q] - \tilde{Q})\right) = \mathbb{E}\left[\mathrm{Tr}\left(AQ\left(\frac{\Sigma}{1+\delta} - \frac{1}{n}XX^T\right)\tilde{Q}\right)\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\mathrm{Tr}\left(\frac{AQ\Sigma\tilde{Q}}{1+\delta} - AQx_ix_i^T\tilde{Q}\right)\right]$$

I - **Motivation:** Hanson Wright for Random Matrix Theory

$$\mathrm{Tr}\left(A(\mathbb{E}[Q] - \tilde{Q}_\delta)\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathrm{Tr}\left(\left(\frac{1}{1+\delta} - \frac{1}{1+\frac{1}{n}x_i^T Q_{-i} x_i}\right) A Q_{-i} x_i x_i^T \tilde{Q}_\delta\right)\right] + O\left(\frac{1}{\sqrt{n}}\right)$$

Use the *Schur Formula:* $Qx_i = \frac{Q_{-i}x_i}{1+\frac{1}{n}x_i^T Q_{-i} x_i}$, with $Q_{-i} \equiv \left(zI_p - \frac{1}{n}XX^T - x_i x_i^T\right)^{-1}$.

Independent with x_i

independent with p, n.

1. Chose $\delta_1 \equiv \frac{1}{n}\mathbb{E}[x_i^T Q_{-i} x_i] \approx \frac{1}{n}\mathrm{Tr}(\Sigma\mathbb{E}[Q]) \approx \frac{1}{n}\mathrm{Tr}(\Sigma\tilde{Q}_{\delta_1})$

   Hanson-Wright Inequality: $\mathbb{P}\left(\left|\frac{1}{n}x_i^T Q_{-i} x_i - \delta_1\right| \geq t\right) \leq Ce^{-ct^2}$

2. Chose δ_2 solution to $\delta = \frac{1}{n}\mathrm{Tr}(\Sigma\tilde{Q}_\delta)$ ⟶ $\mathrm{Tr}\left(A(\mathbb{E}[Q] - \tilde{Q}_{\delta_2})\right) = O\left(\frac{1}{\sqrt{n}}\right)$

Theorem: (Hanson Wright) Given $A \in \mathcal{M}_n$ deterministic, $Z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ such that:

- $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz:
  $$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq C' e^{-c' t^2}$$
- $\|\mathbb{E}[Z]\| \leq K$

$C, c, C', c', K > 0$, independent with n

$$\mathbb{P}\left(\left|Z^T A Z - \mathbb{E}[Z^T A Z]\right| \geq t\right) \leq C e^{-\frac{ct^2}{\|A\|_F^2}} + C e^{-\frac{ct}{\|A\|}}$$

$\Phi(Z)$ satisfying: $|\Phi(Z) - \Phi(Z')| \leq \underbrace{(\|AZ\| + \|AZ'\|)}_{V\text{:variations of }\Phi} \|Z - Z'\|$

Adamczak, Radosław (2014) *A note on the Hanson-Wright inequality for random vectors with dependencies.* Electronic Communications in Probability. 20. 10.1214/ECP.v20-3829.

II - Parallel Sum and Product.

Definition: $\alpha \boxplus \beta = (\alpha^{-1} + \beta^{-1})^{-1}$

Proposition: Given $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}_+$, two random variables $X, Y \in \mathbb{R}$ such that $\forall t \in \mathbb{R}$:

$$\mathbb{P}(X \geq t) \leq \alpha(t) \quad \text{and} \quad \mathbb{P}(Y \geq t) \leq \beta(t)$$

Then $\mathbb{P}(X + Y \geq t) \leq 2\alpha \boxplus \beta(t)$

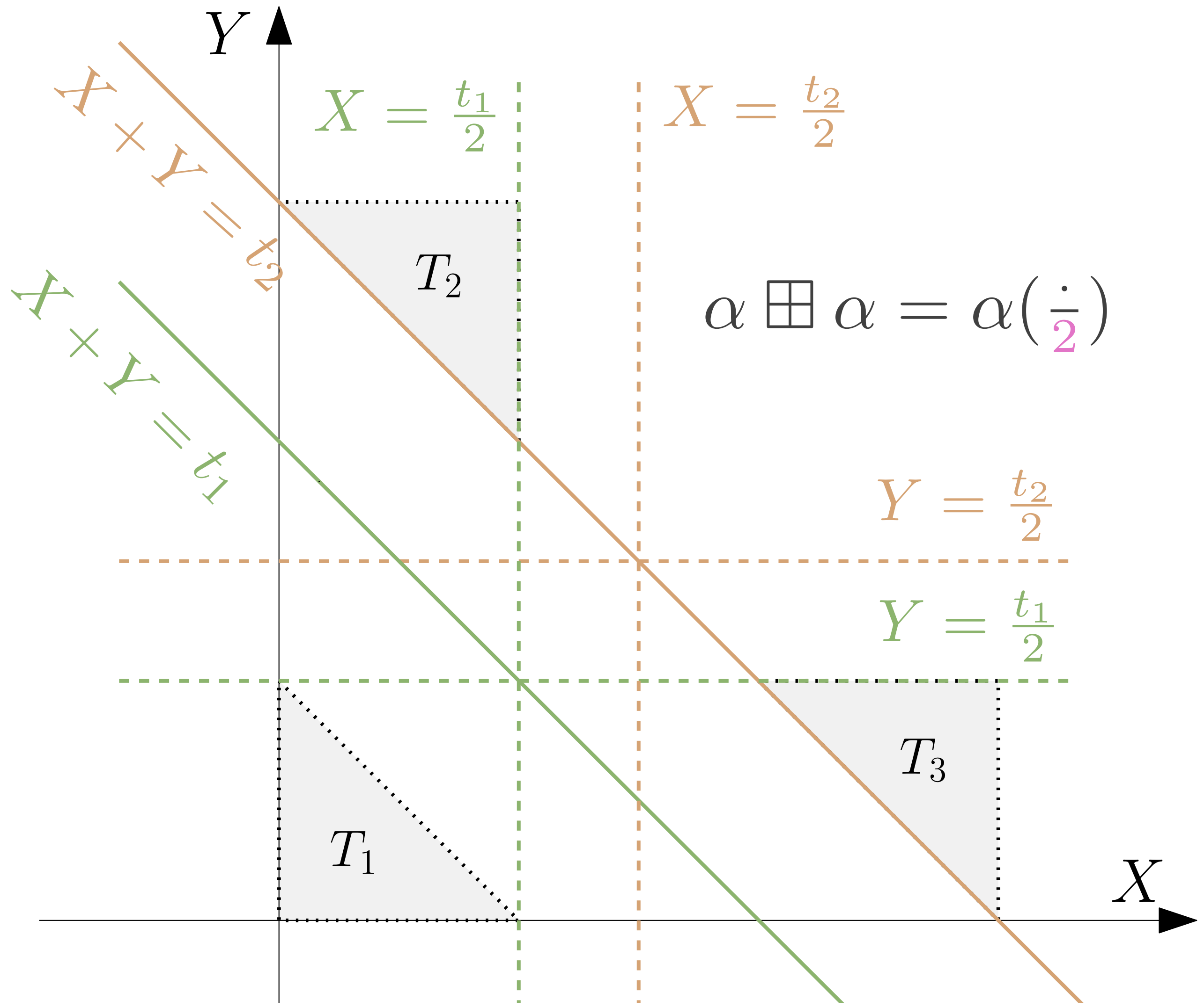
Proof: Denoting $\gamma \equiv \alpha \boxplus \beta$, for any $t \in \mathbb{R}$:

In particular: $\alpha^{-1}(\gamma(t)) + \beta^{-1}(\gamma(t)) = t$

$$\begin{aligned}\mathbb{P}(X + Y \geq t) &\leq \mathbb{P}\left(X + Y \geq \alpha^{-1}(\gamma(t)) + \beta^{-1}(\gamma(t))\right)\\ &\leq \mathbb{P}\left(X \geq \alpha^{-1}(\gamma(t))\right) + \mathbb{P}\left(Y \geq \beta^{-1}(\gamma(t))\right)\\ &\leq 2\gamma(t)\end{aligned}$$

$\forall t \in [t_1, t_2]$:

$$\mathbb{P}(X + Y \geq t) = \tfrac{2}{3} = \mathbb{P}(X \geq \tfrac{t}{2}) + \mathbb{P}(Y \geq \tfrac{t}{2})$$



Uniform distribution of (X, Y) on T_1, T_2, T_3

II - Parallel Sum and **Product**.

Definition: $\alpha \boxtimes \beta \equiv (\alpha^{-1} \cdot \beta^{-1})^{-1}$ $(\alpha, \beta > 0)$

$\alpha, \beta : (-\infty, 0) \to \{+\infty\}$

Proposition: Given $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}_+$, $X, Y > 0$ s.t.:

$\forall t > 0 : \quad \mathbb{P}(X \geq t) \leq \alpha(t) \quad$ and $\quad \mathbb{P}(Y \geq t) \leq \beta(t)$

Then $\mathbb{P}(X \cdot Y \geq t) \leq 2\alpha \boxtimes \beta(t)$

Proof: Denoting $\gamma \equiv \alpha \boxtimes \beta = (\alpha^{-1} \cdot \beta^{-1})^{-1}$, $\forall t > 0$:

$$\mathbb{P}(X \cdot Y \geq t) \leq \mathbb{P}\left(X \cdot Y \geq \alpha^{-1}(\gamma(t)) \cdot \beta^{-1}(\gamma(t))\right)$$

$$\leq \mathbb{P}\left(X \geq \alpha^{-1}(\gamma(t))\right) + \mathbb{P}\left(Y \geq \beta^{-1}(\gamma(t)\right.$$

$$\leq 2\gamma(t)$$

II - Parallel Sum and **Product**.

Introduce:

$$\mathrm{inc}_u : \quad \mathbb{R} \longrightarrow \bar{\mathbb{R}}$$
$$t \longmapsto \begin{cases} -\infty & \text{if } t \leq u \\ +\infty & \text{if } t > u, \end{cases}$$

Lemma: Given α decreasing:

$$\mathbb{P}(|V - u| \geq t) \leq \alpha(t)$$
$$\implies \mathbb{P}(|V| \geq t) \leq \alpha \circ \min\left(\mathrm{inc}_{2u}, \frac{\mathrm{Id}}{2}\right)(t)$$

Proof: $t \geq 2u \implies \frac{t}{2} \leq t - u.$

Lemma: $\alpha \circ (f \boxplus g) = (\alpha \circ f) \boxplus (\alpha \circ g)$

- $\min(f, g)^{-1} = \max(f^{-1}, g^{-1})$
- $\mathrm{inc}_u^{-1} : t \mapsto u$

Now, consider X, V:

$$\mathbb{P}(X \geq t) \leq \alpha \qquad \mathbb{P}(|V - u| \geq t) \leq \alpha(t/\lambda),$$
$$\to \mathbb{P}(XV \geq t) \leq \alpha \circ \mathrm{Id} \boxtimes \min(\mathrm{inc}_{2u}, \mathrm{Id}/2\lambda)(t)$$

Lemma: $\mathrm{Id} \boxtimes \min\left(\mathrm{inc}_u, \frac{\mathrm{Id}}{\lambda}\right) = \min\left(\frac{\mathrm{Id}}{u}, \sqrt{\frac{\mathrm{Id}}{\lambda}}\right)$

Proof: $\mathrm{Id}^{-1} \cdot \min\left(\mathrm{inc}_u, \frac{\mathrm{Id}}{\lambda}\right)^{-1} = \mathrm{Id} \cdot \max(u, \lambda \,\mathrm{Id})$

$$= \max\left(u\,\mathrm{Id}, \lambda\,\mathrm{Id}^2\right)$$

If $\alpha : t \mapsto e^{-t^2}$:

$$\leq e^{-\frac{t^2}{u^2}} + e^{-\frac{t}{\lambda}}$$

Retrieve Hanson Wright right-hand term with:

- $u = \|A\|_F$
- $\lambda = \|A\|$

III - Concentration in High Dimension

Theorem: Given $Z \sim \mathcal{N}(\mu, I_n)$, $\forall f : \mathbb{R}^n \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq 2e^{-\frac{t^2}{2}} \quad Z, Z' \ i.i.d.$$

Given $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ λ-Lipschitz and $f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz:

$$\mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z'))| \geq t\right)$$

$$= \mathbb{P}\left(\left|\frac{1}{\lambda} f(\Phi(Z)) - \frac{1}{\lambda} f(\Phi(Z'))\right| \geq \frac{t}{\lambda}\right) \leq 2e^{-\frac{t^2}{2\lambda^2}}.$$

$$\|\Phi(Z) - \Phi(Z')\| \leq V \|Z - Z'\| \quad a.s.$$

Random

Theorem: **(Talagrand)**
Given $Z = (Z_1, \dots Z_n) \in [0,1]^n$ s.t. $Z_1, \dots, Z_n$ independent
$\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz and convex:

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq 2e^{-\frac{t^2}{4}}.$$

Michel Talagrand (1995) *Concentration of measure and isoperimetric inequalities in product spaces.* Publications mathématiques de l'IHÉS, 104:905–909.

III - Concentration in High Dimension

Theorem: Consider $Z \in \mathbb{R}^n$, random, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq \alpha(t) \qquad (Z, Z' \text{ i.i.d.})$$

- Consider $V \in \mathbb{R}_+$ random s.t.:

$$\forall t > 0 : \quad \mathbb{P}\left(V \geq t\right) \leq \beta(t)$$

- Consider $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s.t.:

$$\|\Phi(Z) - \Phi(Z')\| \leq V \|Z - Z'\| \quad a.s.$$

Then: $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\forall t > 0 : \quad \mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z')| \geq t\right) \leq 2\, \alpha \boxtimes \beta(t)$$

Proof: Denote $\gamma \equiv \alpha \boxtimes \beta = (\alpha^{-1} \cdot \beta^{-1})^{-1}$ In particular, $\forall t > 0 : \ \alpha^{-1}(\gamma(t)) \cdot \beta^{-1}(\gamma(t)) = t$

$$\begin{aligned}
\mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z')| \geq t\right) &\leq \mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z')| \geq t, V \leq \beta^{-1}(\gamma(t))\right) + \mathbb{P}\left(V \geq \beta^{-1}(\gamma(t))\right) \\
&\leq \alpha\left(\frac{t}{\beta^{-1}(\gamma(t))}\right) + \beta(\beta^{-1}(\gamma(t))) \\
&\leq 2\gamma(t)
\end{aligned}$$

III - Concentration in High Dimension

Theorem: Consider $Z \in \mathbb{R}^n$, random, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq \alpha(t) \qquad (Z, Z' \text{ i.i.d.})$$

- Consider $V \in \mathbb{R}_+$ random s.t.:

$$\forall t > 0 : \quad \mathbb{P}\left(V \geq t\right) \leq \beta(t)$$

- Consider $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s.t.:

$$\|\Phi(Z) - \Phi(Z')\| \leq V \|Z - Z'\| \quad a.s.$$

Then: $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\forall t > 0 : \quad \mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z')| \geq t\right) \leq 2\, \alpha \boxtimes \beta(t)$$

Theorem: Consider $X = (X_1, \dots, X_n) \in \mathbb{R}^n$,
Such that $\forall i \in [n]$:
- $X_i = \phi(Z_i)$
- $Z_i \sim \mathcal{N}(0, 1)$

Then: $\mathbb{P}(|f(X) - f(X')|| \geq t)$

$$\leq \exp\left(- \min\left(\frac{t}{\phi'(2\sqrt{\log(2n)})}, \frac{\left(\mathrm{Id} \cdot \phi'\right)^{-1}(t)}{2} \right)^2 \right)$$

" Large-tailed concentration inequality "

IV - Application to Hanson-Wright inequality

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq \alpha(t) \qquad (Z, Z' \text{ i.i.d.})$$

- Consider $V \in \mathbb{R}_+$ random s.t.:

$$\forall t > 0 : \quad \mathbb{P}\left(V \geq t\right) \leq \beta(t)$$

Then: $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\forall t > 0 : \quad \mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z')| \geq t\right) \leq 2\, \alpha \boxtimes \beta(t)$$

- Consider $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s.t.:

$$\|\Phi(Z) - \Phi(Z')\| \leq V\|Z - Z'\| \quad a.s.$$

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq 2e^{-t^2/2}$$

$$\Longrightarrow \mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq Ce^{-ct^2}$$

$$\Longrightarrow \mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq C'e^{-c't}$$

For $C, C', c', c > 0$ numerical constant.

Yes, **IF** $\alpha, \beta : t \mapsto 2e^{-t^2/2}$

Other choices for α, β ??

Question: Possible to replace $\begin{cases} f(Z') \\ f(\Phi(Z')) \end{cases}$ with $\begin{cases} \mathbb{E}[f(Z)] \\ \mathbb{E}[f(\Phi(Z))] \end{cases}$ **??**

IV - Application to Hanson-Wright inequality

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^n \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq \alpha(t)$$

- Consider $V \in \mathbb{R}_+$ random s.t.:

$$\forall t > 0: \quad \mathbb{P}\left(|V - \mathbb{E}[V]| \geq t\right) \leq \alpha\left(\frac{t}{\lambda}\right)$$

- Consider $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s.t.:

$$\|\Phi(Z) - \Phi(Z')\| \leq V\|Z - Z'\| \quad a.s.$$

- Assume α independent with n and:

$$\sigma_\alpha \equiv \sqrt{\int_{\mathbb{R}_+} t\alpha(t)dt} \leq \infty \quad (\mathbb{E}[|f(Z) - \mathbb{E}[f(Z)]|^2] \leq \sigma_\alpha^2)$$

Then: $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\forall t > 0: \quad \mathbb{P}\left(|f(\Phi(Z)) - \mathbb{E}[f(\Phi(Z))]| \geq t\right) \leq C\,\alpha\left(\frac{t}{c|\mathbb{E}[V]|}\right) + C\,\alpha\left(\sqrt{\frac{t}{c\lambda}}\right).$$

Recall:

$$\alpha \boxtimes \alpha \circ \min\left(\mathrm{inc}_{\mathbb{E}[V]}, \frac{\mathrm{Id}}{\lambda}\right) = \alpha \circ \min\left(\frac{\mathrm{Id}}{\mathbb{E}[V]}, \sqrt{\frac{\mathrm{Id}}{\lambda}}\right)$$

IV - Application to Hanson-Wright inequality

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq \alpha(t)$$

- Consider $V \in \mathbb{R}_+$ random s.t.:

$$\forall t > 0 : \quad \mathbb{P}\left(|\|AZ\| - \mathbb{E}[\|AZ\|]| \geq t\right) \leq \alpha\left(\frac{t}{\|A\|}\right)$$

- Consider $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s.t.:

$$\left\|Z^T AZ - Z'^T AZ'\right\| \leq \underbrace{\left(\|AZ\| + \|AZ'\|\right)}_{V} \|Z - Z'\| \quad a.s.$$

- Assume α independent with n and:

$$\sigma_\alpha \equiv \sqrt{\int_{\mathbb{R}_+} t\alpha(t)dt} \leq \infty$$

Then: $\forall A \in \mathcal{M}_n$, $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\forall t > 0 : \quad \mathbb{P}\left(\left|Z^T AZ - \mathbb{E}[Z^T AZ]\right| \geq t\right) \leq C\, \alpha\left(\frac{t}{c\mathbb{E}[\|AZ\|]}\right) + C\, \alpha\left(\sqrt{\frac{t}{c\|A\|}}\right).$$

IV - Application to Hanson-Wright inequality

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq \alpha(t) \qquad (*)$$

with $\alpha : \mathbb{R}^+ \to \mathbb{R}_+$ independent with n.

- $\sigma_\alpha \equiv \sqrt{\int_{\mathbb{R}_+} t\alpha(t)dt} \leq \infty$

- Assume $\|\mathbb{E}[Z]\| \leq \sigma_\alpha$.

Then: $\forall A \in \mathcal{M}_n$, $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\mathbb{P}\left(\left|Z^T AZ - \mathbb{E}[Z^T AZ]\right| \geq t\right) \leq C\ \alpha\left(\frac{ct}{\sigma_\alpha \|A\|_F}\right) + C\alpha\left(\sqrt{\frac{ct}{\|A\|}}\right).$$

$$\begin{aligned}\mathbb{E}[\|AZ\|] &\leq \sqrt{\mathbb{E}[\|AZ\|^2]} \\ &= \sqrt{\mathbb{E}[\mathrm{Tr}(A^T AZZ^T)]} \\ &= \|A\|_F \sqrt{\|\mathbb{E}[ZZ^T]\|}\end{aligned}$$

Lemma: Given $Z \in \mathbb{R}^n$ satisfying $(*)$:

$$\|\mathbb{E}[ZZ^T]\| \leq \|\mathbb{E}[Z]\|^2 + C\sigma_\alpha^2$$

for some numerical constant $C > 0$

IV - Application to Hanson-Wright inequality

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - \mathbb{E}[f(Z)]| \geq t\right) \leq \alpha(t) \qquad (*)$$

with $\alpha : \mathbb{R}^+ \to \mathbb{R}_+$ independent with n.

- $\sigma_\alpha \equiv \sqrt{\int_{\mathbb{R}_+} t\alpha(t)dt} \leq \infty$

- Assume $\|\mathbb{E}[Z]\| \leq \sigma_\alpha$.

Then: $\forall A \in \mathcal{M}_n$, $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz, $\forall t > 0$:

$$\mathbb{P}\left(\left|Z^T AZ - \mathbb{E}[Z^T AZ]\right| \geq t\right) \leq C\ \alpha\left(\frac{ct}{\sigma_\alpha \|A\|_F}\right) + C\alpha\left(\sqrt{\frac{ct}{\|A\|}}\right).$$

Comparison *Adamczak's result:* $\alpha : t \mapsto e^{-\frac{t^2}{2\sigma_\alpha{}^2}}$

V - Multi-level concentration and conjugate of parallel sum

What happened ?

Compute: $$\left(\frac{\mathrm{Id}}{\nu_1}\right)^{\frac{1}{1}} \boxtimes \min\left(\mathrm{inc}_{\theta_0}, \left(\frac{\mathrm{Id}}{\theta_1}\right)^{\frac{1}{1}}\right)$$

Convention: $$\left(\frac{t}{\theta_0}\right)^{\frac{1}{0}} = \mathrm{inc}_{\theta_0}$$

$$\left(\frac{\mathrm{Id}}{\nu_1}\right)^{\frac{1}{1}} \boxtimes \min\left(\left(\frac{\mathrm{Id}}{\theta_0}\right)^{\frac{1}{0}}, \left(\frac{\mathrm{Id}}{\theta_1}\right)^{\frac{1}{1}}\right) = \min\left(\left(\frac{\mathrm{Id}}{\nu_1\theta_0}\right)^{\frac{1}{1+0}}, \left(\frac{\mathrm{Id}}{\nu_1\theta_1}\right)^{\frac{1}{1+1}}\right).$$

V - Multi-level concentration and conjugate of parallel sum

Theorem:

$$\forall t \in \mathbb{R} : \qquad \alpha(t) = \min_{a \in A} \left(\frac{t}{\check{\alpha}_a} \right)^{\frac{1}{a}} \qquad \text{and} \qquad \beta(t) = \min_{b \in B} \left(\frac{t}{\check{\beta}_b} \right)^{\frac{1}{b}},$$

where $(\check{\alpha}_a)_{a \in A} \in \mathbb{R}_+^A$ and $(\check{\beta}_b)_{b \in B} \in \mathbb{R}_+^B$, for $A, B \subset \mathbb{R}_+$

Then : $$\alpha \boxtimes \beta = \min_{(a,b) \in A \times B} \left(\frac{t}{\check{\alpha}_a \check{\beta}_b} \right)^{\frac{1}{a+b}}.$$

V - Multi-level concentration and conjugate of parallel sum

Theorem: If $\forall t \in \mathbb{R}, \forall k \in [n]$:

$$\alpha^{(k)}(t) = \min_{a \in A^{(k)}} \left(\frac{t}{\sigma_a^{(k)}} \right)^{\frac{1}{a}}$$

where $(\sigma_a^{(k)})_{a \in A^{(k)}} \in \mathbb{R}_+^{A^{(k)}}$ and $A^{(k)} \subset \mathbb{R}_+$

Then : $$\alpha^{(1)} \boxtimes \cdots \boxtimes \alpha^{(n)} = \min_{(a_1, \ldots, a_n) \in A^{(1)} \times \cdots \times A^{(n)}} \left(\frac{t}{\sigma_a^{(1)} \cdots \sigma_b^{(n)}} \right)^{\frac{1}{a_1 + \cdots + a_n}}.$$

Useful when $\|\Phi(Z) - \Phi(Z')\| \leq V_2 \cdots V_n \|Z - Z'\|)$

$\alpha^{(2)}$ $\alpha^{(n)}$ $\alpha^{(1)}$

Proof:

Hypothesis: $\alpha^{(k)}(t) = \min_{a \in A^{(k)}} \left(\frac{t}{\sigma_a^{(k)}} \right)^{\frac{1}{a}}$

$$\left(\alpha^{(1)} \boxtimes \cdots \boxtimes \alpha^{(n)}\right)^{-1} = \left(\inf_{a_1 \in A^{(1)}} \left(\frac{\mathrm{Id}}{\sigma_{a_1}^{(1)}} \right)^{\frac{1}{a_1}} \right)^{-1} \cdots \left(\inf_{a_n \in A^{(n)}} \left(\frac{\mathrm{Id}}{\sigma_{a_n}^{(n)}} \right)^{\frac{1}{a_n}} \right)^{-1}$$

$$= \sup_{a_1 \in A^{(1)}} \sigma_{a_1}^{(1)} \mathrm{Id}^{a_1} \cdots \sup_{a_n \in A^{(n)}} \sigma_{a_n}^{(n)} \mathrm{Id}^{a_n}$$

$$= \sup_{a_1 \in A^{(1)}, \dots, a_n \in A^{(n)}} \sigma_{a_1}^{(1)} \cdots \sigma_{a_n}^{(n)} \mathrm{Id}^{a_1 + \cdots + a_n}$$

$$= \left(\inf_{a_1 \in A^{(1)}, \dots, a_n \in A^{(n)}} \left(\frac{\mathrm{Id}}{\sigma_{a_1}^{(1)} \cdots \sigma_{a_n}^{(n)}} \right)^{\frac{1}{a_1 + \cdots + a_n}} \right)^{-1}$$

V - Multi-level concentration and conjugate of parallel sum

Theorem:

- Consider $Z \in \mathbb{R}^n$, s.t. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(Z) - f(Z')| \geq t\right) \leq \alpha(t) \qquad (Z, Z' \text{ i.i.d.})$$

- $\Phi : \mathbb{R}^p \to \mathbb{R}^p$: $\qquad \|\Phi(Z) - \Phi(Z')\| \leq \Lambda^{(1)} \cdots \Lambda^{(n)} \cdot \|Z - Z'\|, a.s.$

- $\forall k \in [n]$, $\exists A^{(k)} \subset \mathbb{R}_+$, $(\sigma^{(k)}_{a \in A^{(k)}}) \in \mathbb{R}_+^{A^{(k)}}$:

$$\mathbb{P}\left(\left|\Lambda^{(k)} - \sigma_0^{(k)}\right| \geq t\right) \leq \alpha \circ \inf_{a \in A^{(k)} \setminus \{0\}} \left(\frac{t}{\sigma_a^{(k)}}\right)^{\frac{1}{a}}$$

Then: $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

$$\mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z'))| \geq t\right) \leq (n+1)\alpha \circ \inf_{a_k \in A^{(k)}, k \in [n]} \left(\frac{t}{\sigma_{a_1}^{(1)} \cdots \sigma_{a_n}^{(n)}}\right)^{\frac{1}{1 + a_1 + \cdots + a_n}}.$$

Remark:

If $A^{(1)} = \{0\}$: $\Lambda^{(1)} = \sigma^{(1)}$

- Appears in speed
- Do not affect powers

V - Multi-level concentration and conjugate of parallel sum

$$\mathbb{P}\left(|f(\Phi(Z)) - f(\Phi(Z'))| \geq t\right) \leq \alpha \circ \inf_{a_k \in A^{(k)}, k \in [n]} \left(\frac{t}{\sigma^{(1)}_{a_1} \cdots \sigma^{(n)}_{a_n}}\right)^{\frac{1}{1+a_1+\cdots+a_n}}.$$

"Multi-level concentration inequalities"

Our assumption:

$$\|\Phi(Z) - \Phi(Z')\| \leq V_2 \cdots V_n \|Z - Z'\|)$$

Rigorous proofs in:

arXiv > math > arXiv:2402.08206

Search...

Help | Ad

Mathematics > Probability

[Submitted on 13 Feb 2024 (v1), last revised 12 Jun 2024 (this version, v5)]

Operation with Concentration Inequalities and Conjugate of Parallel Sum

Cosme Louart

THANK YOU!