Hanson Wright à queue large

Cosme Louart

Chinese University of Hong Kong Shenzhen

SCHOOL OF DATA SCIENCE

Cornell University

We gratefully acknowledge

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

Following the concentration of the measure theory formalism, we consider the transformation $\Phi(Z)$ of a random variable Z having a general concentration function α. If the transformation Φ is λ-Lipschitz with $\lambda > 0$ deterministic, the concentration function of $\Phi(Z)$ is immediately deduced to be equal to $\alpha(\cdot/\lambda)$. If the variations of Φ are bounded by a random variable Λ having a concentration function (around 0) $\beta : \mathbb{R}_+ \to \mathbb{R}$, this paper sets that $\Phi(Z)$ has a concentration function analogous to the so-called parallel produuct of α and β: $(\alpha^{-1} \cdot \beta^{-1})^{-1}$. We apply this result to (i) express the concentration of large-tailed random vectors, (ii) generalize Hanson Wright inequality, and (iii) provide useful insights on the so-called "multilevel concentration" that appears when Λ is the product of n random variables. This last result is obtained when we formulate the conjugate functions of the parallel sum of n real mappings.

Subjects: **Probability (math.PR)**; Functional Analysis (math.FA)

MSC classes: 60-08, 60B20, 62J07

Cite as: arXiv:2402.08206 **[math.PR]**

(or arXiv:2402.08206v5 **[math.PR]** for this version)

https://doi.org/10.48550/arXiv.2402.08206

Hanson Wright à queue large

Cosme Louart

Chinese University of Hong Kong Shenzhen

SCHOOL OF DATA SCIENCE

Soient 2 variables aléatoires: $X, Y \in \mathbb{R}$

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

$$\times$$

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

$$=$$

$$\forall t > 0 :$$

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

Hanson Wright à queue large

Cosme Louart
Chinese University of Hong Kong Shenzhen
SCHOOL OF DATA SCIENCE

Soient 2 variables aléatoires: $X, Y \in \mathbb{R}$

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

$$\times$$

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

$$=$$

"Produit parallèle"

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

Hanson Wright à queue large

Cosme Louart

Chinese University of Hong Kong Shenzhen

SCHOOL OF DATA SCIENCE

- Soit un vecteur aléatoire: $X \in \mathbb{R}^n$:

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

  $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz

Hanson Wright
à queue large

Cosme Louart

Chinese University of Hong Kong Shenzhen

SCHOOL OF DATA SCIENCE

Plan

I - **Motivation:** Hanson Wright pour la théorie des mat. al.

Méthode de la transformée de Stieltjes, resolvante, leave-one-out

II - Somme et produit parallèle

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

III - Concentration en Grande dimension

$\mathbb{P}(|f(X) - \mathbb{E}[f(X)]| \geq t) \leq \alpha(t)$, $\forall f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz

Concentration de Φ quand $\|\Phi(Z) = \Phi(Z')\| \leq V\|Z - Z'\|$.

IV - Application à l'inégalité de Hanson-Wright

Généralisation aux concentration à queue large

I - **Motivation:** Hanson Wright pour la théorie des mat. al.

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

But: Distribution des valeurs propres de $\frac{1}{n} X X^T$: $\mu \equiv \frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i}$ **??**

valeurs propres de $\frac{1}{p} X X^T$

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

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

"Transformée de Steiltjes" (similaire à la transformée de Cauchy)

- Lien avec la *"Resolvante"*: $m(z) = \frac{1}{p} \text{Tr} Q(z)$, où $Q(z) \equiv \left(z I_p - \frac{1}{n} X X^T \right)^{-1}$.

Strategie: trouver $\tilde{Q} \in \mathcal{M}_p$ deterministe t.q. $Q \approx \tilde{Q}$

I - **Motivation:** Hanson Wright pour la théorie des mat. al.

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

- Bien sûr $\mathbb{E}[Q]$ loin de $(zI_p - \Sigma)^{-1}$

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

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

δ à déterminer

Dependance entre
Q et x_i

Soit $A \in \mathcal{M}_p$, deterministe:

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

I - **Motivation:** Hanson Wright pour la théorie des mat. al.

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

Avec *Formule de Schur:* $Qx_i = \dfrac{Q_{-i}x_i}{1+\frac{1}{n}x_i^TQ_{-i}x_i}$, où $Q_{-i} \equiv \left(zI_p - \frac{1}{n}XX^T - x_ix_i^T\right)^{-1}$.

Independant avec x_i

independant avec p, n.

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

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

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

I - **Motivation:** Hanson Wright pour la théorie des mat. al.

Theorem: **(Hanson Wright)** Soit $A \in \mathcal{M}_n$ deterministe, $Z = (z_1, \dots, z_n) \in \mathbb{R}^n$ t.q.:

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

$C, c, C', c', K > 0$, independants avec n

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

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

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 - Somme et Produit Parallèle.

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

Proposition: Soient $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}_+$, deux var. al. $X, Y \in \mathbb{R}$ t.q. $\forall t \in \mathbb{R}$:

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

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

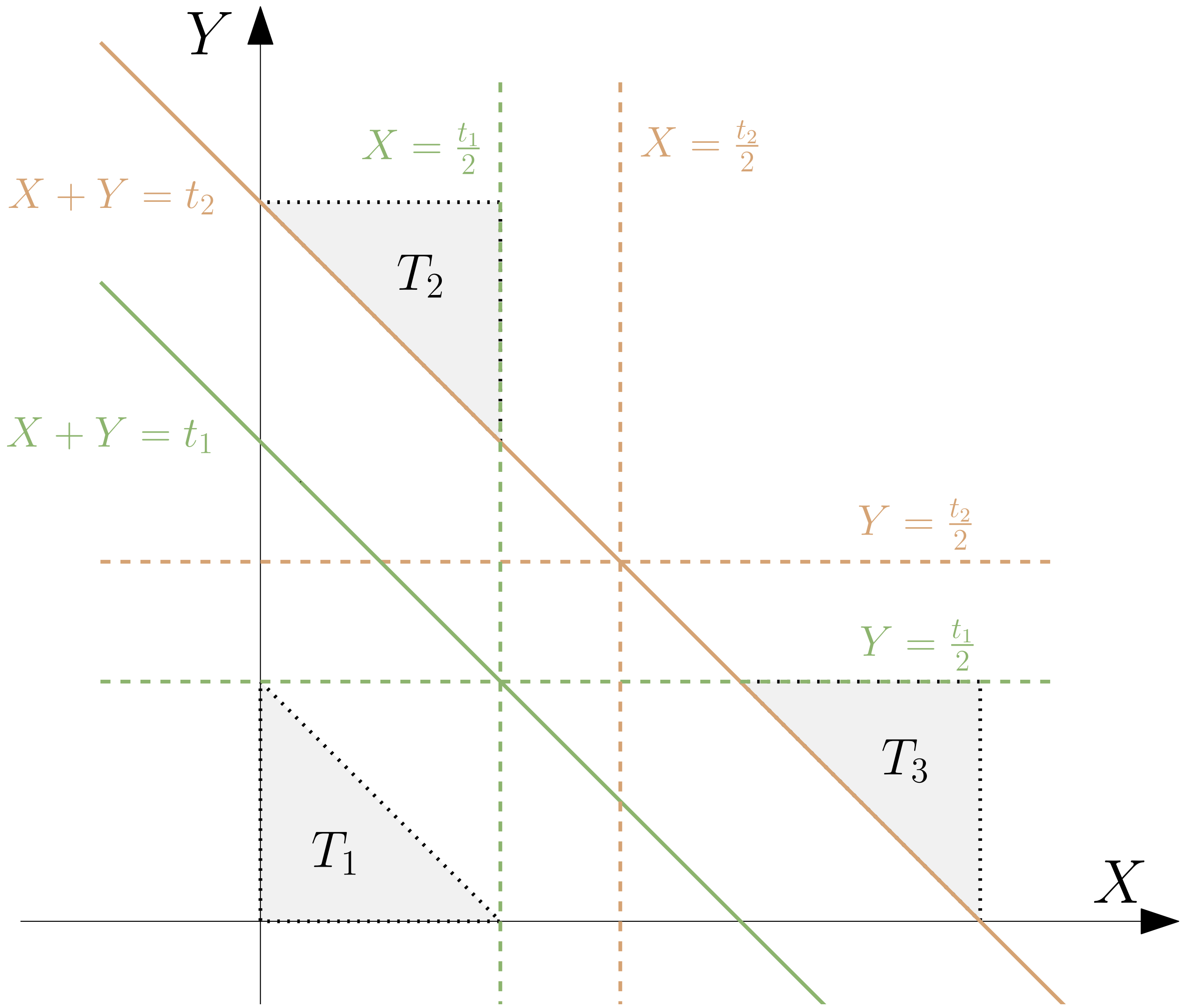
Preuve: Notons $\gamma \equiv \alpha \boxplus \beta$, pour tout $t \in \mathbb{R}$:

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

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

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

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



Distribution uniforme of (X, Y) on T_1, T_2, T_3

II - Somme et Produit Parallèle.

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

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

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

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

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

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

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

II - Somme et Produit Parallèle.

Introduisons:

$$\begin{aligned} \mathrm{inc}_u : \quad \mathbb{R} &\longrightarrow \bar{\mathbb{R}} \\ t &\longmapsto \begin{cases} -\infty & \text{si} \quad t < u \\ +\infty & \text{si} \quad t \geq u, \end{cases} \end{aligned}$$

Lemme: Soit α décroissante:

$$\mathbb{P}(|V - u| \geq t) \leq \alpha(t)$$

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

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

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

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

Maintenant, étant donnés X, V:

$$\mathbb{P}(X \geq t) \leq \alpha \qquad \mathbb{P}(|V - u| \geq t) \leq \alpha(t/\lambda),$$

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

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

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

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

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

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

On retrouve le terme de droite de Hanson Wright avec:

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

III - Concentration en grande Dimension

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

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

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

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

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

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

Aléatoire

Theorème: **(Talagrand)**
Soit $Z = (Z_1, \ldots Z_n) \in [0,1]^n$ s.t. $Z_1, \ldots, Z_n$ independants
$\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz et convexe:

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

"Diamètre de la Distribution" : $O(\sqrt{n})$

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

III - Concentration en grande Dimension

Theorème: Soit $Z \in \mathbb{R}^n$, aléatoire, t.q. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 1-Lipschitz:

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

- Soit $V \in \mathbb{R}_+$ aléatoires t.q.:

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

- Soit $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ t.q.:

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

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

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

"Preuve:" Notons $\gamma \equiv \alpha \boxtimes \beta = (\alpha^{-1} \cdot \beta^{-1})^{-1}$. En particulier, $\forall t > 0 : \ \alpha^{-1}(\gamma(t)) \cdot \beta^{-1}(\gamma(t)) = t$

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

III - Concentration en grande Dimension

Theorème:

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

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

- Soit $V \in \mathbb{R}_+$ aléatoire t.q.:

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

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

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

- Soit $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ t.q.:

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

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

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

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

Pour $C, C', c', c > 0$ constantes numériques.

Question: Peux-t-on remplacer $\begin{cases} f(Z') \\ f(\Phi(Z')) \end{cases}$ avec $\begin{cases} \mathbb{E}[f(Z)] \\ \mathbb{E}[f(\Phi(Z))] \end{cases}$??

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

D'autres choix pou α, β ??

III - Concentration en grande Dimension

Theorème:

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

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

- Soit $V \in \mathbb{R}_+$ aléatoire t.q.:

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

- Soit $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ t.q.:

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

- Suppose α independant avec n et:

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

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

$$\forall t > 0 : \quad \mathbb{P}\left(|f(\Phi(Z)) - \mathbb{E}[f(\Phi(Z))]| \geq t\right) \leq 2\,\alpha\left(\frac{t}{|\mathbb{E}[V]|}\right) + 2\,\alpha\left(\sqrt{\frac{t}{\lambda}}\right).$$

Rappel:

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

IV - Application à l'inégalité de Hanson-Wright

Theorème:

- Soit $Z \in \mathbb{R}^n$, t.q. $\forall f : \mathbb{R}^p \to \mathbb{R}$, 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.:

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

- Supposons α independant avec n et:

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

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

$$\forall t > 0 : \quad \mathbb{P}\left(\left|Z^T AZ - \mathbb{E}[Z^T AZ]\right| \geq t\right) \leq 2\,\alpha\left(\frac{t}{\mathbb{E}[\|AZ\|]}\right) + 2\,\alpha\left(\sqrt{\frac{t}{\|A\|}}\right).$$

IV - Application à l'inégalité de Hanson-Wright

Theorème:

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

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

avec $\alpha : \mathbb{R}^+ \to \mathbb{R}_+$ indépendant avec n.

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

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

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

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

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

Lemme: Soit $Z \in \mathbb{R}^n$ t.q.$(*)$:

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

pour une constante numérique $C > 0$

IV - Application à l'inégalité de Hanson-Wright

Theorème:

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

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

avec $\alpha : \mathbb{R}^+ \to \mathbb{R}_+$ indépendant avec n.

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

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

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

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

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

Aller plus loin...

Retour sur le calcul

$$\alpha \boxtimes \alpha \circ \min\left(\mathrm{inc}_{\delta_1}, \frac{\mathrm{Id}}{\eta_1}\right) = \left((\alpha \circ \exp) \boxplus \left(\alpha \circ \min\left(\mathrm{inc}_{\delta_1}, \frac{\mathrm{Id}}{\eta_1}\right) \circ \exp\right)\right) \circ \log$$
$$= \alpha \circ \exp \circ \left(\mathrm{Id} \boxplus \min\left(\mathrm{inc}_{\log \delta_1}, \mathrm{Id} - \log(\eta_1)\right)\right) \circ \log,$$

$$\mathrm{inc}_u : \quad t \longmapsto \begin{cases} -\infty & \text{si} \quad t < u \\ +\infty & \text{si} \quad t \geq u, \end{cases}$$

Nouvelle notation: $\frac{\mathrm{Id} - \log \delta_1}{0}\Big|_{\cap +}$

$$\frac{\mathrm{Id} - \nu_1}{1} \boxplus \min\left(\frac{\mathrm{Id} - \theta_0}{0}\Big|_{\cap +}, \frac{\mathrm{Id} - \theta_1}{1}\right) = \min\left(\frac{\mathrm{Id} - \nu_1 - \theta_0}{1+0}, \frac{\mathrm{Id} - \nu_1 - \theta_1}{1+1}\right).$$

Aller plus loin...

$$\forall t \in \mathbb{R} : \qquad \alpha(t) = \min_{a \in A} \frac{t - \check{\alpha}_a}{a} \Big|_{\substack{\cap \\ +}} \qquad \text{et} \qquad \beta(t) = \min_{b \in B} \frac{t - \check{\beta}_b}{b} \Big|_{\substack{\cap \\ +}},$$

où $(\check{\alpha}_a)_{a \in A} \in \mathbb{R}^A$ et $(\check{\beta}_b)_{b \in B} \in \mathbb{R}^B$, pour $A, B \subset \mathbb{R}_+$

$$\alpha \boxplus \beta = \min_{(a,b) \in A \times B} \frac{t - \check{\alpha}_a - \check{\beta}_b}{a + b} \Big|_{\substack{\cap \\ +}}.$$

Aller plus loin...

$$\forall t \in \mathbb{R}, \forall k \in [n] : \qquad \alpha^{(k)}(t) = \min_{a \in A^{(k)}} \frac{t - \check{\alpha}_a^{(k)}}{a} \Big|_{\cap}^{+}$$

où $(\check{\alpha}_a^{(k)})_{a \in A^{(k)}} \in \mathbb{R}^{A^{(k)}}$, pour $A^{(k)} \subset \mathbb{R}_+$

$$\alpha^{(1)} \boxplus \cdots \boxplus \alpha^{(n)} = \min_{(a_1,\dots,a_n) \in A^{(1)} \times \cdots \times A^{(n)}} \frac{t - \check{\alpha}_{a_1}^{(1)} - \cdots - \check{\alpha}_{a_n}^{(n)}}{a_1 + \cdots + a_n} \Big|_{\cap}^{+}.$$

Indice pour la preuve : $\check{\alpha} = -(\alpha^{-1})^*$
"conjugate of the inverse"

(Pour des cas où $\|\Phi(Z) - \Phi(Z')\| \leq V_2 \cdots V_n \|Z - Z'\|$)

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

Aller plus loin...

$$\alpha^{(1)} \boxplus \cdots \boxplus \alpha^{(n)} = \min_{(a_1,\dots,a_n)\in A^{(1)}\times\cdots\times A^{(n)}} \frac{t - \check{\alpha}^{(1)}_{a_1} - \cdots - \check{\alpha}^{(n)}_{a_n}}{a_1 + \cdots + a_n}\Big|_{\substack{\cap\\+}}.$$

Inégalité de concentration
"multi-niveaux"

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

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

Plus pratique pour
les statisticiens ?