# Probabilités

## Chapitre 3 : Concentration de la mesure

Lucie Le Briquer

## Sommaire

1 **Introduction** 1

2 **Fonctions de concentration** 3

3 **Concentration ensembliste et concentration des fonctions Lipschitziennes** 4

4 **De log-Sobolev à la concentration** 6

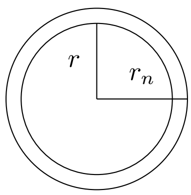
5 **Vecteurs gaussiens** 7

6 **Applications : opérateurs et matrices Gaussiennes** 9

# 1 Introduction

On travaille sur $\mathbb{R}^n$ en ayant en tête que $n$ est grand. On note $\mathcal{B}_2^n$ la boule Euclidienne de rayon 1 ($B_2^n = \{\|x\|_2 \leq 1\}$).

Un calcul (en TD) vous montrera que si on cherche le rayon $r_n > 0$ tel que $\text{Vol}(r_n\mathcal{B}_2^n) = 1$ alors $r_n \sim c\sqrt{n}$ (où $c$ est une constante nmérique.)



$\text{Vol}(\mathcal{B}) = 1$. Si $r < r_n$, $\text{Vol}(r\mathcal{B}_2^n) = \text{Vol}(\frac{r}{r_n}\mathcal{B}) = (\frac{r}{r_n})^n \xrightarrow[n\longrightarrow+\infty]{} 0$

Si par exemple $r = (1-\varepsilon)r_n$ avec $\varepsilon \in [0,1]$ on a $(\frac{r}{r_n})^n \sim (1-\varepsilon)^n \xrightarrow[n\longrightarrow+\infty]{} 0$ très vite.

La boule de rayon $(1-\varepsilon)r_n$ est de volume presque nul (en grande dimension). Toute la masse est concentrée dans une couronne finie.

Regardons $\mathcal{B}_\infty^n = [-1,1]^n$. Dans $\mathcal{B}_\infty^n$ il y a des points de la forme $(\pm 1, ..., \pm 1)$ qui sont à distance (Euclidienne) $\sqrt{n}$ de l'origine, donc très éloignés de l'origine. D'autre part, on a des points comme $(1, 0, ..., 0)$ qui sont à distance 1 de l'origine.
$\text{Vol}(\mathcal{B}_\infty^n) = 2^n$

$$r < r_n \sim \sqrt{n} \qquad \text{Vol}((r\mathcal{B}_2^n) \cap \mathcal{B}_\infty^n) \leq \text{Vol}(r\mathcal{B}_2^n) \xrightarrow[n \longrightarrow +\infty]{} 0$$

Ainsi toute la masse du cube est concentrée autour des sommets.

Soit $\mathcal{A} \subseteq \mathbb{R}^n$ compact. Définissons l' *épaissi* $t$ de $\mathcal{A}$ comme :

$$\begin{aligned} \mathcal{A}_t &= \{x \in \mathbb{R}^n | d(x, \mathcal{A}) < t\} \\ &= \mathcal{A} + t\mathcal{B}_2^n \\ &= \{x + ty | x \in \mathcal{A} \text{ et } y \in \mathcal{B}_2^n\} \end{aligned}$$

Soit $\mathcal{A} \subseteq \mathbb{R}^n$, soit $\mathcal{B}$ la boule Euclidienne ayant le même volume que $\mathcal{A}$ ($\text{Vol}(\mathcal{A}) = \text{Vol}(\mathcal{B})$)

$$\begin{aligned} \text{Vol}(\mathcal{A}_t)^{1/n} &= \text{Vol}(\mathcal{A} + t\mathcal{B}_2^n)^{1/n} \\ &\geq \text{Vol}(\mathcal{A})^{1/n} + \text{Vol}(t\mathcal{B}_2^n)^{1/n} \qquad \text{Brunn-Minkowski (TD)} \\ &= \text{Vol}(\mathcal{B})^{1/n} + \text{Vol}(t\mathcal{B}_2^n)^{1/n} \end{aligned}$$

Si $\mathcal{B} = r\mathcal{B}_2^n$ (i.e. $\mathcal{B}$ est de rayon $r$)

$$\begin{aligned} \text{Vol}(\mathcal{A}_t) &= r\text{Vol}(\mathcal{B}_2^n)^{1/n} + t\text{Vol}(\mathcal{B}_2^n)^{1/n} \\ &= (r+t)\text{Vol}(\mathcal{B}_2^n)^{1/n} \\ &= \text{Vol}(\underbrace{(r+t)\mathcal{B}_2^n}_{\mathcal{B}_t})^{1/n} \end{aligned}$$

On a montré que $\text{Vol}(\mathcal{A}_t)^{1/n} \geq \text{Vol}(\mathcal{B}_t)^{1/n}$
Si on définit $\text{Vol}(\partial\mathcal{A}) = \liminf_{t \to 0} \frac{\text{Vol}(\mathcal{A} + t\mathcal{B}_2^n) - \text{Vol}(\mathcal{A})}{t}$, on vient de montrer que $\text{Vol}(\partial\mathcal{A}) \geq \text{Vol}(\partial\mathcal{B})$ (où $\mathcal{B}$ boule Euclidienne de même volume que $\mathcal{A}$).
À volume fixé, les boules Euclidiennes sont celles qui ont la plus petite mesure de bord. On appelle ceci inégalité isopérimétrique.
Ici on a travaillé sur $\mathbb{R}^n$, avec la métrique Euclidienne et la mesure de Lebesgue, on pourrait étudier ce phénomène dans d'autres cas.

Donnons un autre exemple : prenons $S^{n-1}$ la sphère unité de $\mathbb{R}^n$. On munit $S^{n-1}$ de la métrique géodésique (i.e. $d(x,y)$ correspond au plus petit arc les reliant).
Il existe une unique mesure sur $S^{n-1}$ invariante par rotation.
On définit pour $A \subseteq S^{n-1}$ :

$$\tilde{\sigma} = \text{Vol}(\{tu | t \in [0,1], u \in A\})$$

On prend $\mu = \frac{\tilde{\sigma}}{\tilde{\sigma}(S^{n-1})}$ mesure de probabilité.
Phénomène isopérimétrique démontré par Lévy : "À mesure fixée, les coupes sphériques sont celles qui ont le plus petit périmètre."
$\forall A \subseteq S^{n-1}$, et $B$ une coupe sphérique telle que $\mu(A) = \mu(B)$ alors $\mu(A_t) \geq \mu(B_t)$.

Ainsi si $A$ est telle que $\mu(A) \geq \frac{1}{2}$ alors en prenant $B$ au moins une demi-sphère, et en calculant on trouve :

$$\mu(A_t^C) \leq e^{-\frac{(n-1)t^2}{2}}$$

On voit le lien entre inégalité isopérimétrique et phénomène de concentration.
Si $A$ a beaucoup de masse ($\geq \frac{1}{2}$), alors dès qu'on s'éloigne de $A$, la masse décroît très rapidement.

## 2 Fonctions de concentration

**Définition 1** (espace métrique de probabilité)

Un triplet $(X, d, \mu)$ est un espace métrique de probabilité (epm) si $(X, d)$ est un espace métrique et $\mu$ est une probabilité.

**Remarque.**
La tribu borélienne sur $X$ est la plus petite tribu engendrée par les ouverts de $X$.

**Exemples.**

- $S^{n-1}$ muni de la métrique géodésique et de la probabilité $\mu$ définie dans l'introduction
- $\mathbb{R}^n$ muni de la métrique Euclidienne et de la mesure Gaussienne $\gamma_n$
- $\Omega_n$ muni de la métrique de Hamming et de la mesure uniforme $\sigma_n$

**Définition 2** ($r$-voisinage)

Si $(X, d)$ est un espace métrique, on définit le $r$-voisinage de tout ensemble $A \subseteq X$ par :

$$A_r = \{x \in X \mid d(x, A) < r\}$$

et donc $A_r^C = \{x \in X \mid d(x, A^C) \geq r\}$

**Définition 3** (fonction de concentration)

Soit $(X, d, \mu)$ un espace métrique de probabilité. La fonction de concentration $\alpha_{(X,d,\mu)}$ de $(X, d, \mu)$ est donnée par :

$$\forall r > 0, \quad \alpha_{(X,d,\mu)}(r) = \sup \left\{ \mu(A_r^C) \mid A \subseteq X, \mu(A) \geq \frac{1}{2} \right\}$$

**Remarques.**

- La fonction de concentration est la meilleure (la plus petite) fonction $\alpha : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ telle que $\forall \subseteq X$ et $\forall r \geq 0, \mu(A_r^C) \leq \alpha(r), \mu(A) \geq \frac{1}{2}$
- Si $r > \text{Diam}(X,d) = \sup\{d(x,y) \mid x,y \in X\}$ alors $\alpha_{(X,d,\mu)}(r) = 0$
- Si $r \longrightarrow +\infty$ on devrait avoir $\alpha_{(X,d,\mu)}(r) \longrightarrow 0$
- $\alpha_{(X,d,\mu)}$ sert juste à majorer $\mu(A_r^C)$. On s'intéresse alors à trouver des majorations de $\alpha$ et non la calculer explicitement.
- On notera parfois $\alpha_\mu$ à la place de $\alpha_{(X,d,\mu)}$

**Définition 4** (concentration Gaussienne / exponentielle)

$(X,d,\mu)$ a une concentration Gaussienne (respectivement exponentielle) si $\exists\ c, C > 0$ (constantes) telles que :

$$\alpha_{(X,d,\mu)}(r) \leq Ce^{-cr^2} \qquad \text{(respectivement } \alpha_{(X,d,\mu)}(r) \leq Ce^{-cr}\text{)}$$

# 3 Concentration ensembliste et concentration des fonctions Lipschitziennes

**Remarque.**
Pour $(X,d)$ un EPM, $f : X \longrightarrow \mathbb{R}$ est Lipschitzienne si $\exists\ c$ telle que :

$$\forall x,y \in X, \quad |f(x) - f(y)| \leq cd(x,y)$$

On définit $\|f\|_{Lip}$ la meilleure (plus petite) constante $c$ pour laquelle on a cette relation. $f$ est 1-Lipschitzienne si $\|f\|_{Lip} \leq 1$

**Définition 5** (médiane)

Si $f$ est $\mu$-intégrable on dit que $m_f \in \mathbb{R}$ est une médiane de $f$ si :

$$\mu(\{f \leq m_f\}) \geq \frac{1}{2} \quad \text{et } \mu(\{f \geq m_f\}) \geq \frac{1}{2}$$

**Proposition 1** (conc. ensembliste $\Rightarrow$ conc. des fonctions Lip autour de la médiane)

$(X, d, \mu)$ EMP avec une fonction de concentration $\alpha_\mu$

Alors $\forall f : X \longrightarrow \mathbb{R}$ Lip de médiane $m_f$, on a :

$$\mu(\{f \leq m_f - r\}) \leq \alpha_\mu \left( \frac{r}{\|f\|_{Lip}} \right) \quad \text{et} \quad \mu(\{f \geq m_f + r\}) \leq \alpha_\mu \left( \frac{r}{\|f\|_{Lip}} \right)$$

Ainsi :

$$\mu(\{|f - m_f| \geq r\}) \leq 2\alpha_\mu \left( \frac{r}{\|f\|_{Lip}} \right)$$

**Preuve.**

On peut supposer que $\|f\|_{Lip} = 1$. Soit $A = \{f \leq m_f\}$, par définition de $m_f$, on a $\mu(A) \geq \frac{1}{2}$. Calculons $A_r$ :

$$A_r = \{x \in X | d(x, A) < r\} \subset \{x \in X | f(x) < m_f + r\}$$

Donc $A_r^c \supset \{x \in X, f(x) \geq m_f + r\}$. Concentration ensembliste $\Rightarrow \mu(A_r^c) \leq \alpha(r)$. Les autres cas s'en déduisent de manière similaire.

$\square$

**Proposition 2** (réciproque)

$(X, d, \mu)$ un e.m.p et $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ tel que $\forall f : X \to \mathbb{R}$ lipschitzienne de médiane $m_f$, $\forall r > 0$ on a :

$$\mu(\{f \geq m_f + r\}) \leq \alpha \left( \frac{r}{\|f\|_{Lip}} \right)$$

Alors pour tout $A \subseteq X$ tel que $\mu(A) \geq \frac{1}{2}$, on a

$$\forall r > 0, \mu(A_r^c) \leq \alpha(r)$$

Ainsi $\alpha_{(X,d,\mu)} \leq \alpha$

**Preuve.**

Soit $A \subseteq X$ tel que $\mu(A) \geq \frac{1}{2}$. On prend $f(x) = d(x, A)$, alors $f$ est 1-Lip (par l'inégalité triangulaire).

$$A_r = \{x | d(x, A) < r\} = \{f < r\} \quad \text{et} \quad A \subseteq \{f = 0\}$$

$\mu(A) \geq \frac{1}{2} \Rightarrow \mu(\{f = 0\}) \geq \frac{1}{2} \Rightarrow 0$ est une médiane de $f$

$$\mu(A_r^c) = \mu(\{f \geq r\}) = \mu(\{f \geq m_f + r\}) \leq \alpha(r)$$

$\square$

**Proposition 3**

$(X, d, \mu)$ e.m.p et $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ tel que $\forall f$ Lip, on a :

$$\mu\left(f \geq \int f d\mu + r\right) \leq \alpha\left(\frac{r}{\|f\|_{Lip}}\right), \forall r \geq 0$$

Alors, $\forall A \subseteq X$ tel que $\mu(A) > 0$, on a :

$$\mu(A_r^c) \leq \alpha(\mu(A)r)$$

Ainsi, si $\alpha$ décroissante, on a $\alpha_{(X,d,\mu)} \leq \alpha(\frac{r}{2})$

**Preuve.**
Soit $A \subseteq X$, $\mu(A) > 0$. Soit $r > 0$. Prenons $F_r(x) = min(d(x, A), r)$ qui est 1-Lip.

$$\int F_r d\mu = \int_{A^c} F_r d\mu \leq r\mu(A^c)$$

$$\mu(A_r^c) = \mu(\{F \geq r\}) = \mu(\{F \geq r\mu(A^c) + r\mu(A)\}) \leq \mu\left(\{F \geq \int F d\mu + r\mu(A)\}\right)$$

Donc $\mu(A^c) \leq \alpha(r\mu(A))$

□

# 4 De log-Sobolev à la concentration

**Théorème**

$(\mathbb{R}^n, \|.\|_2, \mu)$ emp satisfaisant ILS$_c$ alors :

$$\forall f \ 1-\text{lipschitzienne}, \quad \mu\left(\left\{f \geq \int f d\mu + r\right\}\right) \leq e^{-r^2/c}$$

en particulier, l'espace a une concentration Gaussienne.

**Preuve.**
Soit $f$ 1-Lipschitzienne ; on peut supposer que $f$ est différentiable et que $|\nabla f| \leq 1$. Soit $\lambda \in \mathbb{R}$ et $g(x) = \exp\left(\frac{\lambda f(x)}{2}\right)$.

$$\text{ISL}_c \text{ à } g \qquad \text{Ent}_\mu(g^2) \leq c \int |\nabla g|^2 d\mu$$

$$\begin{aligned}\text{Ent}(g^2) &= \int g^2 \ln g^2 d\mu - \int g^2 d\mu \ln \int g^2 d\mu \\ &= \int \lambda f e^{\lambda f} d\mu - \int e^{\lambda f} d\mu \ln \int e^{\lambda f} d\mu\end{aligned}$$

À finir.

$\square$

**Corollaire**

L'espace Gaussien a une concentration Gaussienne. Plus précisément :

$$\forall f \text{ } 1-\text{lipschitzienne}, \quad \mu\left(\left\{f \geq \int f d\mu + r\right\}\right) \leq e^{-r^2/2}$$

**Théorème**

$\forall f : \{-1,1\}^n \longrightarrow \mathbb{R}$, on définit :

$$v = \max_{x\in\{-1,1\}^n} \frac{1}{2}\sum_{i=1}^{n}(f(x) - f(\tau_i(x)))^2$$

On a

$$\sigma_n\left(\left\{f \geq \int f d\sigma_n + r\right\}\right) \leq e^{-r^2/v}$$

$\sigma_n$ mesure uniforme sur le cube discret.

*Cours du 31 mars*

# 5 Vecteurs gaussiens

**Définition 6** (vecteur gaussien)

On dit que $X \in \mathbb{R}^n$ est un vecteur gaussien si $\forall\theta \in S^{n-1}, < \theta|X >$ est une gaussienne.

**Proposition 4**

Si $X = (X_1, \dots, X_n)$ avec $\{X_i\}$ indépendantes gaussiennes, alors $X$ est un vecteur gaussien.

**Preuve.**

$\forall\theta \in S^{n-1}$, $< \theta|X >= \sum_{i=1}^{n} \theta_i X_i$ est une somme de gaussiennes indépendantes donc est gaussien.

$\square$

**Remarque.**
$X$ Gaussien standard $\Rightarrow$ ses coordonnées sont Gaussiennes indépendantes.

**Définition 7** (covariance)

On appelle covariance du vecteur $X = (X_1, \ldots, X_n)$ la matrice $\Sigma \in \mathcal{M}_n\mathbb{R}$ définie par $\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j)$

**Remarque.**
$\Sigma$ est symétrique et a pour diagonale les variances.

Pour simplifier, on va supposer que $X$ est centrée : $(\mathbb{E}(X_i, X_j))_{\{i,j\}} = \mathbb{E}(X^t X)$ est donc positive.

**Proposition**

$X$ vecteur aléatoire de matrice de covariance $\Sigma$ $n \times n$. $A$ $k \times n$. Alors $AX$ vecteur gaussien de matrice de covariance $A\Sigma A^k$.

**Proposition**

$A$ matrice symétrique définie positive. Alors la loi du vecteur Gaussien centré de matrice de covariance $A$ a une densité / mesure de Lesbesgue de $\mathbb{R}^n$ donnée par :

$$\frac{1}{(\sqrt{2\pi})^k \sqrt{\det A}} \exp\left(-\frac{1}{2} < A^{-1}x, x >\right)$$

**Proposition**

$X$ Gaussien centré dans $\mathbb{R}^n$, matrice de covariance $\Sigma$. Si $\theta_1, \theta_2$ sont 2 directions $\in S^{n-1}$ alors $< X, \theta_1 >$ et $< X, \theta_2 >$ sont indépendates ssi $\theta_1 \perp \theta_2$ ($\Leftrightarrow$ $\mathrm{cov}(< X, \theta_1 >, < \theta_2, X >) = 0$).

*Cours du 7 avril*

# 6 Applications : opérateurs et matrices Gaussiennes

$G$ une matrice $N \times n$ dont les entrées sont des variables aléatoires Gaussiennes i.i.d. $\mathcal{N}(0,1)$. On peut aussi voir $G$ comme un vecteur Gaussien dans $\mathbb{R}^{nN}$.

$$G : \mathbb{R}^n \longrightarrow \mathbb{R}^N$$

On cherche l'action de l'application $G$. Regardons $G$ comme un opérateur :

$$G : l_2^n \longrightarrow l_2^N$$

où $l_2^n = (\mathbb{R}^n, \|.\|_2)$

$\forall x \in S^{n-1}$ :

$$Gx = \sum_{i=1}^{N} < x, L_i(G) > e_i = \begin{pmatrix} \vdots \\ < x, L_i(G) > \\ \vdots \end{pmatrix}_N$$

$Gx$ est un vecteur Gaussien standard (i.e. $\mathcal{N}(0, Id)$) car :

$$\mathbb{E}(< L_i(G), x >) = 0$$
$$\mathbb{E}(< L_i(G), x >^2) = \|x\|_2^2 = 1$$

donc $Gx$ a des entrées indépendantes $\sim \mathcal{N}(0,1) \Rightarrow Gx \sim \mathcal{N}(0, Id_{\mathbb{R}^N})$.

Si on s'intéresse à $G$ en tant qu'opérateur de $l_2$ dans $l_2$ on aimerait trouver $\alpha$ et $\beta$ tels que :

$$\forall x \in \mathbb{R}^n, \qquad \alpha \|x\|_2 \leq \|Gx\|_2 \leq \beta \|x\|_2$$

$$\mathbb{E}(\|Gx\|_2^2) = \sum_{i=1}^{N} \mathbb{E}(< L_i(G), x >^2) = N \|x\|_2^2$$
$$\text{or} \quad \mathbb{E}(\|Gx\|_2^2) = \mathbb{E}(x^t G^t G x)$$

De plus :

$$G^t G = \begin{pmatrix} \vdots & & \vdots \\ L_1(G) & & L_N(G) \\ \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \dots & L_1(G) & \dots \\ & & \\ \dots & L_N(G) & \dots \end{pmatrix} = \sum_{i=1}^{N} L_i(G) L_i(G)^t$$

Donc :

$$\begin{aligned} \mathbb{E}(G^t G) &= \sum_{i=1}^{N} \mathbb{E}(L_i(G) L_i(G)^t) \\ &= \sum_{i=1}^{N} \mathrm{Cov}(L_i(G)) \\ &= N Id_{\mathbb{R}^n} \end{aligned}$$

On a de façon similaire : $\mathbb{E}(GG^t) = n Id_{\mathbb{R}^N}$

Le $\beta$ correspond à la norme de $G$ où :

$$\|G\| = \|G\|_{2\to 2} = \sup_{\|x\|_2\leq 1} \|Gx\|_2 = \sup_{\|x\|_2\leq 1} \sqrt{< Gx, Gx >} = \sup_{\|x\|_2\leq 1} \sqrt{< G^*Gx, x >} = \lambda_{\max}((G^*G)^{1/2})$$

De la même façon $\alpha$ correspondrait à :

$$\inf_{\|x\|_2=1} \|Gx\|_2 = \lambda_{\min}((G^*G)^{1/2})$$

**Définition 8** (valeurs singulières)

Étant donnée une matrice $A$ de $N \times n$, on définit les valeurs singulières de $A$ (et on note $s_i(A)$) les valeurs propres de $(A^*A)^{1/2}$

**Remarques.**

- si $A$ est symétrique, les valeurs singulières de $A$ sont les valeurs absolues des valeurs propres.
- les valeurs singulières sont une interprétation géométrique puisqu'on vient de voir que $\forall x \in \mathbb{R}^n$ :
$$s_{\min}(A)\,\|x\|_2 \leq \|Ax\|_2 \leq s_{\max}(A)\,\|x\|_2$$
- $A$ injective $\Leftrightarrow s_{\min}(A) > 0$
du coup si $n > N$, $s_{\min}(A) = 0$ et le nombre de valeurs singulières non nulles est égal au rang de $A$

  Supposons que $n \leq N$, on a :
$$s_{\min}(A)B_2^N \subseteq AB_2^N \subseteq s_{\max}(A)B_2^N$$
(si $y \in AB_2^n$, $y = Ax$ avec $x \in B_2^n$ donc $\|Ax\|_2 \leq s_{\max}(A)$ d'où $y = Ax \in s_{\max}(A)B_2^N$
si $s_{\min}(A) = 0$, rien à dire ; sinon $A$ est inversible, on fait comme au dessus)

**Définition 9** (nombre de conditionnement)

Le nombre de conditionnement de $A$ est :
$$\kappa(A) = \frac{s_{\max}(A)}{s_{\min}(A)}$$

**Remarque.**

- si $\kappa(A) = 1$ alors $A$ est multiple d'une isométrie (i.e. application qui conserve les normes)
- si $\kappa(A)$ est proche de 1 (ou d'une constante = ne dépend pas de la dimension) on dit que $A$ est *bien conditionnée*

Reprenons $G$. On a $\mathbb{E}(G^*G) = NId_{\mathbb{R}^n}$ donc en moyenne $G$ est une isométrie. Montrons maintenant que $G$ est une isométrie avec une grande probabilité à l'aide des inégalités de concentration.

**Proposition 5**

$G$ $N \times n$ Gaussienne. On note :

$$\begin{aligned} m &= \int_{\mathbb{R}^N} \|x\| \, d\gamma_N(x) \qquad \text{où } \|.\| \text{ est n'importe quelle norme} \\ &= \mathbb{E}(\|g\|) \quad \text{où } g \sim \mathcal{N}(0, Id_{R^N}) \\ &= \mathbb{E}(\|Gu\|) \quad \forall u \in S^{n-1} \end{aligned}$$

Soit $b > 0$ tq $\|.\| \leq b \|.\|_2$. Alors $\forall S$ ensemble fini de $\mathbb{R}^n$, on a :

$$\mathbb{P}\left(\{\forall y \in S, (1-\varepsilon)m \|y\|_2 \leq \|Gy\| \leq (1+\varepsilon)m \|y\|_2\}\right) \geq 1 - 2|S| \exp\left(-\frac{m^2\varepsilon^2}{4b^2}\right)$$

**Preuve.**

$\forall y \in S^{n-1}$, posons $E_y = \{| \|Gy\| - m \|y\|_2 | > \varepsilon m \|y\|_2\} = \{| \|Gy\| - \mathbb{E}(\|Gy\|)| > \varepsilon m\}$. On cherche à montrer que :

$$\mathbb{P}\left(\bigcap_{y \in S} \overline{E_y}\right) \geq 1 - 2|S| \exp\left(-\frac{m^2\varepsilon^2}{4b^2}\right)$$

$(\mathbb{R}^N, \gamma_N, \|.\|_2)$ est un espace Gaussien (il a une concentration Gaussienne). Soit $f : \left\{ \begin{array}{ccc} \mathbb{R}^N & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x\| \end{array} \right.$ est $b$-Lipschitzienne. Donc :

$$\gamma_N\left(\left\{\left|f(x) - \int f d\gamma_N\right| > r\right\}\right) \leq 2 \exp\left(-\frac{r^2}{2b^2}\right)$$

Donc :

$$\mathbb{P}(E_y) = \mathbb{P}(\{| \|Gy\| - m \|y\|_2 | > \varepsilon m \|y\|_2\}) \leq 2 \exp\left(-\frac{\varepsilon^2 m^2}{2b^2}\right)$$

D'où :

$$\mathbb{P}\left(\bigcup_{y \in S} E_y\right) \leq 2|S| \exp\left(-\frac{\varepsilon^2 m^2}{2b^2}\right)$$

□

**Remarques.**

- ceci redonne le lemme de Johnson-Linderstrauss
- si $\|.\| = \|.\|_2$, alors $b = 1$ et $m = \mathbb{E}\|y\|_2 \left(\leq (\mathbb{E}(\|y\|_2^2)^{1/2}\right)$ de l'ordre de $\sqrt{n}$

Retour à notre but : estimer les valeurs singulières de $G$. On aimerait que :

$$(1-\varepsilon) \leq s_{\min}(G) \leq s_{\max}(G) \leq 1+\varepsilon$$

$$(1-\varepsilon) \leq \inf_{s\in S^{n-1}} \|Gx\|_2 \leq \sup_{s\in S^{n-1}} \|Gx\|_2 \leq 1+\varepsilon$$

**Définition 10** (réseau)

Pour $\varepsilon > 0$, on dit qu'un ensemble fini $S \subseteq S^{n-1}$ est un $\delta$-réseau de $S^{n-1}$ si :

$$\forall x \in S^{n-1}, \exists y \in S \text{ tq } \|x-y\|_2 \leq \delta$$

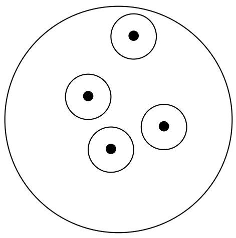
**Lemme 6**

$\forall \delta > 0$, on peut trouver $S$ un $\delta$-réseau tel que $|S| \leq \left(1+\frac{2}{\delta}\right)^n$

**Preuve.**

$S = \{y_1, ..., y_s\} \subset S^{n-1}$ un ensemble $\delta$-séparé (i.e. $\forall i \neq j, \|y_i - y_j\|_2 > \delta$) et maximal (i.e. $\forall y \in S^{n-1}$, $S \cup \{y\}$ n'est pas $\delta$-séparé).

$S$ maximal $\Rightarrow$ $S$ est un $\delta$-réseau



Les boules $\mathcal{B}(y_i, \delta/2)$ sont disjointes.

$$\bigcup_{i=1}^{s} \mathcal{B}(y_i, \delta/2) \subseteq \mathcal{B}(0, 1+\delta/2)$$

$$\mathrm{Vol}\left(\bigcup_{i=1}^{s} \mathcal{B}(y_i, \delta/2)\right) \leq \mathrm{Vol}(\mathcal{B}(0, 1+\delta/2))$$

$$\sum_{i=1}^{s} \mathrm{Vol}(\mathcal{B}(y_i, \delta/2)) \leq \mathrm{Vol}(\mathcal{B}(0, 1+\delta/2))$$

$$s\mathrm{Vol}(\mathcal{B}_2(0, \delta/2)) \leq \mathrm{Vol}(B_2(0, 1+\delta/2))$$

$\square$

**Théorème 7**

$G$ $N \times n$ avec $n < cN$ où $c \ll 1$ est une constante. On a :

$$\mathbb{P}\Big(c_1\sqrt{n} \leq s_{\min}(G) \leq s_{\max}(G) \leq c_2\sqrt{n}\Big) \geq 1 - e^{-cN}$$

où $c_1, c_2$ sont des constantes universelles. Donc $G$ est bien conditionnée.

**Preuve.**

Soit $\varepsilon \in [0,1]$.

$$(1-\varepsilon)m \leq s_{\min}(G) \leq s_{\max}(G) \leq (1+\varepsilon)m$$

où $m = \mathbb{E}(\|g\|_2)$ où $g \sim \mathcal{N}(0, Id_{\mathbb{R}_N})$

$$\Leftrightarrow |\, \|Gx\|_2 - m| \leq \varepsilon \quad \forall x \in S^{n-1}$$

On doit montrer que :

$$\Gamma = \mathbb{P}\left(\exists x \in S^{n-1} / |\, \|Gx\|_2 - m| > \varepsilon m\right) \qquad \text{est petite}$$

$$\leq \mathbb{P}\left(\exists x \in S^{n-1} / |\, \|Gx\|_2 - m| > \varepsilon m \text{ et } \ \|G\| \leq (1+\varepsilon)m\right) + \mathbb{P}(\|G\| > (1+\varepsilon)m)$$

Soit $S$ un $\delta$-réseau de $S^{n-1}$ ($\delta$ spécifié plus tard). Soit $x \in S^{n-1}$ tel que $|\, \|Gx\|_2 - m| > \varepsilon m$. Soit $y \in S$ tl que $\|x-y\|_2 \leq \delta$

$$\begin{aligned}
|\, \|Gy\|_2 - m| &= |\, \|Gx - G(x-y)\|_2 - m| \\
&\geq |\, \|Gx\|_2 - m| - \|G(x-y)\|_2 \\
&\geq \varepsilon m - (1+\varepsilon)m\delta \\
\text{d'où} \quad |\, \|Gy\|_2 - m| &> (\varepsilon - (1+\varepsilon)\delta)m
\end{aligned}$$

De plus :

$$\|G\| \leq (1+\varepsilon) \Rightarrow \exists x \in S^{n-1} \text{ tq } \ \|G\| = \|Gy\| \geq (1+\varepsilon)m$$

Soit $y \in S$ tel que $\|x-y\|_2 \leq \delta$, $\|Gy\| = \|Gx - G(x-y)\| \geq \|Gx\| - \|G\|\,\delta = (1-\delta)\,\|Gx\|$.

Donc :

$$\Gamma \leq \mathbb{P}\Big(\exists y \in S / |\, \|Gy\|_2 - m| > (\varepsilon - (1+\varepsilon)\delta)m\Big) + \mathbb{P}\Big(\exists y \in S / \ \|Gy\|_2 - m \geq (1+\varepsilon)(1-\delta)m\Big)$$

Prenons $\delta = \varepsilon/3$ (on a modifié une inégalité, prendre un bon delta)

$$\Gamma \leq 2\sum_{y\in S} \mathbb{P}\left(|\, \|Gy\|_2 - 1| \geq \frac{\varepsilon m}{3}\right)$$

Pareil que dans la proposition : on trouve $N = n(\varepsilon)$ pour avoir presque une isométrie.

□

**Théorème 8** (Dvoretsty)

$E = (\mathbb{R}^n, \|.\|)$ ; $\kappa(E) = \left(\frac{m}{b}\right)^2$ où $b$ est la constante de Lipschitz de $\|.\|$ par rapport à $\|.\|_2$ et $m = \int \|x\| \, d\gamma_n(x)$. Alors :

$\forall \varepsilon \in [0,1]$, on a :

$$l_2^\kappa \overset{1+\varepsilon}{\hookrightarrow} E \quad \text{avec} \quad \kappa = \frac{c\varepsilon^2}{\ln(1+\frac{1}{\varepsilon})}\kappa(E)$$

i.e. $\exists T : \mathbb{R}^n \longrightarrow \mathbb{R}^\kappa$ tel que :

$$(1-\varepsilon)\mathcal{B}_2^\kappa \subseteq T(\mathcal{B}_E) \subseteq (1+\varepsilon)\mathcal{B}_2^\kappa$$

**Remarque.**

$\forall K$ convexe, sym, compacte d'intérieur non vide

$$(1-\varepsilon)\mathcal{B}_2^\kappa \subseteq K \cup \underset{dim\kappa}{F} \subseteq 1+\varepsilon\mathcal{B}_2^\kappa$$

$\kappa(E) \geq \log n$