Exercices sur le problème d'acquisition comprimée

Guillaume Lecué

11 février 2016

Table des matières

1 Complexité

****

Exercice 1.1 (Argument volumique)
Soit $\|\cdot\|$ une norme sur $\mathbb{R}^n$. On note par B sa boule unité. Soit $0 < \varepsilon \leq 1$ et $\Lambda \subset B$ tels que pour tout $x, y \in \Lambda$

$$\|x - y\| \geq \varepsilon. \tag{1}$$

Alors nécessairement, le cardinal de Λ est tel que

$$|\Lambda| \leq \left(1 + \frac{2}{\varepsilon}\right)^n$$

****

****

Exercice 1.2 (Varshamov-Gilbert)
Soit $s \leq N/2$. Il existe une famille $\mathcal{S}$ d'ensembles de $\{1, \ldots, N\}$ telle que

1. **$\forall S \in \mathcal{S}$, $|S| = s$,**
2. **$\forall S_1, S_2 \in \mathcal{S}$, $S_1 \neq S_2 \Rightarrow |S_1 \cap S_2| \leq \lfloor s/2 \rfloor$,**
3. **$\log |\mathcal{S}| \geq \lfloor \frac{s}{2} \rfloor \log \left(\lfloor \frac{N}{8es} \rfloor \right)$.**

****

****

Exercice 1.3 (Calcul de l'intégrale de Dudley de B_1^N par rapport à ℓ_2^N)

On considère les boules unités $B_1^N = \{x \in \mathbb{R}^N : \sum_{i=1}^N |x_i| \leq 1\}$ **et** $B_2^N = \{x \in \mathbb{R}^N : \sum_{i=1}^N x_i^2 \leq 1\}$**. Pour tout** $\varepsilon > 0$**, le nombre minimal de translatés de la boule** εB_2^N **nécessaires pour couvrir** B_1^N **est noté** $N(B_1^N, \varepsilon B_2^N)$**. On rappelle qu'une mesure de la complexité de** B_1^N **pour la métrique** ℓ_2^N **est l'intégrale de Dudley définie par :**

$$I(B_1^N, \ell_2^N) := \int_0^\infty \sqrt{\log N(B_1^N, \varepsilon B_2^N)} d\varepsilon.$$

On va montrer qu'il existe une constante absolue $c_0 > 0$ **telle que** $I(B_1^N, \ell_2^N) \leq c_0 (\log N)^{3/2}$**. Pour cela on utilise un argument probabiliste.**

On calcul d'abord le nombre d'entropie de B_1^N **par rapport à** ℓ_2^N**. Soit** $\varepsilon > 0$**. On cherche à construire un** ε**-réseau de** B_1^N **pour** ℓ_2^N**. C'est-à-dire, on cherche un nombre minimal de points** $x_1, \ldots, x_p \in B_1^N$ **tels que pour tout** $x \in B_1^N$ **il existe** $i_0 \in \{1, \ldots, p\}$ **tel que** $\|x - x_{i_0}\|_2 \leq \varepsilon$**. Soit** $x \in B_1^N$**. On écrit** $x = \sum_{i=1}^N \lambda_i e_i$ **où** $(e_1, \ldots, e_N)$ **est la base canonique de** $\mathbb{R}^N$ **et où** $\sum_{i=1}^N |\lambda_i| \leq 1$**. On considère la variable aléatoire** X **à valeurs dans** $\{\pm e_1, \ldots, \pm e_N, 0\}$ **définie par**

$$\mathbb{P}[X = \text{Sign}(\lambda_i) e_i] = |\lambda_i|, \ \forall i = 1, \ldots, N \ \textbf{et} \ \mathbb{P}[X = 0] = 1 - \|x\|_1.$$

1. **Déterminer la moyenne de** X**.**
2. **Soit** n **un entier à déterminer plus tard et** $X_1, \ldots, X_n$ **des variables aléatoires indépendantes et de même loi que** X**. Montrer que** $\mathbb{E} \left\| n^{-1} \sum_{i=1}^n X_i - \mathbb{E}X \right\|_2^2 \leq n^{-1}$
3. **En déduire que si** $n = \lfloor \varepsilon^{-2} \rfloor$ **alors**

$$\Lambda_\varepsilon := \left\{ \frac{1}{n} \sum_{i=1}^n z_i : z_1, \ldots, z_n \in \{\pm e_1, \ldots, \pm e_N, 0\} \right\}$$

   est un ε**-réseau de** B_1^N **par rapport à** ℓ_2^N**. Déterminer alors une borne pour** $\log N(B_1^N, \varepsilon B_2^N)$**.**
4. **Montrer que pour tout** $\eta \geq \varepsilon$**, on a**

$$\log N(B_1^N, \varepsilon B_2^N) \leq \log N(B_1^N, \eta B_2^N) + N \log\left(\frac{3\eta}{\varepsilon}\right).$$

5. **En déduire que pour tout** $0 < \varepsilon \leq N^{-1/2}$**,**

$$\log N(B_1^N, \varepsilon B_2^N) \leq c_0 N \log\left(c_1 / N\varepsilon^2\right)$$

6. **Finalement, prouver que**

$$\log N(B_1^N, \varepsilon B_2^N) \leq c_0 \begin{cases} 0 & \textbf{quand } \varepsilon \geq 1, \\ \frac{\log(eN)}{\varepsilon^2} & \textbf{quand } N^{-1/2} \leq \varepsilon \leq 1, \\ N \log\left(\frac{e}{N\varepsilon^2}\right) & \textbf{quand } 0 < \varepsilon \leq N^{-1/2}. \end{cases}$$

   Conclure qu'il existe bien $c_0 > 0$ **tel que** $I(B_1^N, \ell_2^N) \leq c_0 (\log N)^{3/2}$**.**

****

****

Exercice 1.4 (Borne inférieure sur l'intégrale de Dudley de B_1^N par rapport à ℓ_2^N)
En utilisant la borne de Varshamov-Gilbert, montrer que

$$I(B_1^N, \ell_2^N) = \int_0^\infty \sqrt{\log N(B_1^N, \varepsilon B_2^N)} d\varepsilon \geq c_1 (\log N)^{3/2}$$

pour une certaine constante absolue $c_1 > 0$.

On pourra d'abord montrer que si $N(B_1^N, \varepsilon B_2^N)$ est le nombre minimal de translatées de εB_2^N nécessaires pour couvrir B_1^N et $M(B_1^N, \varepsilon B_2^N)$ est le nombre maximal de points ε-écartés dans B_1^N par rapport à ℓ_2^N alors

$$N(B_1^N, \varepsilon B_2^N) \geq M(B_1^N, \varepsilon B_2^N).$$

****
****

Exercice 1.5 (Sur la complexité du cube combinatoire par rapport à la distance de Hamming)

On considère un entier N et le cube combinatoire $\mathcal{C}_N = \{0,1\}^N$. On munit $\mathcal{C}_N$ de la distance de Hamming : pour tout $x, y \in \mathcal{C}_N$,

$$\rho(x,y) = \sum_{i=1}^N I(x_i \neq y_i)$$

où $I(x_i \neq y_i) = 1$ quand $x_i \neq y_i$ et $I(x_i \neq y_i) = 0$ quand $x_i = y_i$. On fixe un entier $1 \leq k < N/2$. On s'intéresse aux sous-ensembles de points de $\mathcal{C}_N$ qui sont k-écartés par rapport à la distance de Hamming.

1. **On considère X la variable aléatoire uniformément distribuée sur $\mathcal{C}_N$. Soit $x \in \mathcal{C}_N$. Montrer que**
$$2^N \mathbb{P}\big[\rho(X,x) \leq k\big] = |\{y \in \mathcal{C}_N : \rho(y,x) \leq k\}|$$
où pour tout ensemble E, le cardinal de E est noté $|E|$.
2. **Étant donné $x \in \mathcal{C}_N$, déterminer la loi de $\rho(X,x)$.**
3. **Étant donné un point $x \in \mathcal{C}_N$, utiliser 1), 2) et la borne de Chernoff pour majorer le nombre de points qui sont à distance au plus k de x pour la distance de Hamming.**
4. **Construire un ensemble de points k écartés de cardinal au moins**
$$\left\lceil \frac{2^N}{\left(eN/k\right)^k} \right\rceil.$$

****

2 Concentration

****

Exercice 2.1 (Estimé de Bernoulli)

Soit $(A_i)_{i\geq 1}$ une suite d'événements indépendants tel que $a := \sum_{i\geq 1} \mathbb{P}(A_i) < \infty$. Pour tout n, l'*estimé de Bernoulli* dit que

$$\mathbb{P}\Big[\sum_{i\geq 1} I_{A_i} \geq n\Big] \leq \frac{a^n}{n!} \leq \Big(\frac{ea}{n}\Big)^n.$$

****

****

Exercice 2.2 (Inégalité de concentration de Chernoff)

Soit $\delta_1, \ldots, \delta_n$ des variables i.i.d. de Bernoulli de moyenne δ (aussi appelé des *sélecteurs*) définies par $\mathbb{P}[\delta_1 = 1] = 1 - \mathbb{P}[\delta_1 = 0] = \delta$. L'inégalité de concentration de Chernoff dit que pour tout $0 \leq t < 1 - \delta$,

$$\mathbb{P}\Big[\frac{1}{n}\sum_{i=1}^n \delta_i - \delta \geq t\Big], \mathbb{P}\Big[\frac{1}{n}\sum_{i=1}^n \delta_i - \delta \leq -t\Big] \leq \exp(-nh_\delta(t)),$$

où $h_\delta(t) := (1-\delta-t)\log\Big(\frac{1-\delta-t}{1-\delta}\Big) + (\delta+t)\log\Big(\frac{\delta+t}{\delta}\Big)$.

****

****

Exercice 2.3 (Les variables bornées sont sous-gaussiennes)

Soit X une variable aléatoire réelle centrée telle que $a \leq X \leq b$ p.s.. Alors pour tout $\lambda > 0$,

$$\mathbb{E}[e^{\lambda X}] \leq \exp\Big(\frac{\lambda^2(b-a)^2}{8}\Big).$$

****

****

Exercice 2.4 (Inégalité de concentration de Hoeffding)

Soit $X_1, \ldots, X_n$ des variables aléatoires réelles centrées indépendantes telles que pour tout $i = 1, \ldots, n$, $a_i \leq X_i \leq b_i$. Pour tout $t > 0$,

$$\mathbb{P}\Big[\frac{1}{n}\sum_{i=1}^n X_i > t\Big] \leq \exp\Big(-\frac{2nt^2}{\frac{1}{n}\sum_{i=1}^n (b_i - a_i)^2}\Big).$$

****

****

Exercice 2.5 (Dénombrement et concentration)

Démontrer que pour tout N et tout $1 \leq D \leq N$, on a

$$\sum_{i=0}^D \binom{N}{i} \leq \Big(\frac{eN}{D}\Big)^D.$$

On pourra considérer N variables de Bernoulli $\delta_1, \ldots, \delta_N$ de moyenne $1/2$ et utiliser une inégalité de concentration pour borner $\mathbb{P}[\sum_{i=1}^N \delta_i \leq D]$.

****
****

Exercice 2.6 (Lemme de Johnson-Lindenstrauss)

Ce lemme est un fameux résultat de géometrie en grandes dimensions. Etant donné p points $x_1, \dots, x_p$ dans l'espace $\mathbb{R}^N$ (qu'on imagine de grande dimension), il existe une matrice $A : \mathbb{R}^N \to \mathbb{R}^k$ telle que $k \sim \log p$ et pour tout $i, j = 1, \dots, p$, $\|A(x_i - x_j)\|_2 \sim \|x_i - x_j\|_2$.

En d'autres termes, on peut projeter p points d'un espace de grande dimension dans un espace de dimension seulement $\log p$ tout en conservant les distances euclidiennes d'origines dans $\mathbb{R}^N$. C'est donc un résultat de réduction de dimension.

Plus précisément le résultat s'énonce de la manière suivante : *Il existe une constante absolue $c_0 > 0$ telle que pour tout $0 < \varepsilon < 1$, pour tout ensemble de p points $x_1, \dots, x_p$ de $\mathbb{R}^N$, et pour tout $k \geq c_0(\log p)/\varepsilon^2$, il existe un opérateur $A : \mathbb{R}^N \to \mathbb{R}^k$ tel que pour tout $i, j = 1, \dots, p$,*

$$\sqrt{1-\varepsilon}\,\|x_i - x_j\|_2 \leq \|A(x_i - x_j)\|_2 \leq \sqrt{1+\varepsilon}\,\|x_i - x_j\|_2 \,.$$

****
****

Exercice 2.7 (Le problème du collectionneur de coupons)

On dispose d'une urne contenant n coupons différents. A chaque tirage, on choisit au hasard un coupon qui est ensuite replacé dans l'urne. On se demande combien de tirages doivent être effectués pour être à peu prêt sûr (càd avec probabilité au moins $1/2$) d'avoir tiré chacun des coupons au moins une fois. La réponse est un $\mathcal{O}(n \log n)$.

On rappelle que la loi géométrique de paramètre $0 < \delta < 1$ est la loi du nombre N du premier succès dans une suite de Bernoulli indépendantes de paramètre δ. On peut la définir par

$$N = \min(n : \delta_n = 1)$$

où $(\delta_n)_n$ est une suite de Bernoulli indépendantes de paramètre δ.

1. **Déterminer la loi, l'espérance, la variance et les déviation de N : $\mathbb{P}[N > k]$ pour tout $k \in \mathbb{N}^*$.**

Soit $i \in \{1, \dots, n\}$. On suppose qu'on vient juste d'observer $i-1$ coupons différents. On se demande combien de tirages vont être nécessaires pour observer un nouveau coupon. On note par T_i ce nombre de tirages ; c'est-à-dire le premier instant où on observe un i-éme coupon différent des $i-1$ précédemment observés.

2. **Quelle est la loi de T_i ? Quelle est la moyenne de T_1 et T_n ? (Interpréter).**
3. **On note par T l'instant de la première fois où on a observé tous les coupons de l'urne. Déterminer T en fonction des T_i. Déterminer sa moyenne.**
4. **Calculer la variance de T et, grâce à l'inégalité de Chebishev, majorer la probabilité que T dévie de sa moyenne.**

****
****

Exercice 2.8 (Inégalité de Bernstein avec variance empirique)

Première partie

Soient V, $V_1, \dots, V_n$ des v.a. i.i.d. à valeurs négatives ou nulles. Soit $\bar{V} = \frac{1}{n}\sum_{i=1}^n V_i$.

1. **Montrer que $\exp(u) \leq 1 + u + \frac{u^2}{2}$ pour $u \leq 0$.**
2. **Montrer que $\log \mathbb{E} \exp(sV) \leq s\mathbb{E}V + \frac{s^2}{2}\mathbb{E}(V^2)$ pour tout $s \geq 0$.**
3. **En déduire par l'argument de Chernoff que pour tout $\lambda > 0$, on a**

$$\mathbb{P}\big(\bar{V} - \mathbb{E}V > t\big) \leq \exp\left(-\lambda t + \frac{\lambda^2 \mathbb{E}(V^2)}{2n}\right).$$

4. **En déduire qu'avec probabilité au moins $1-\varepsilon$, on a**

$$\bar{V} - \mathbb{E}V \leq \sqrt{\frac{2\mathbb{E}(V^2)\log(1/\varepsilon)}{n}}.$$

Deuxième partie

Soient X, $X_1, \dots, X_n$ des v.a. i.i.d. à valeurs dans $[0,1]$. Soient $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, $\sigma = \sqrt{\mathbb{E}[(X-\mathbb{E}X)^2]}$ l'écart type de X, $\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2}$ l'écart type empirique, $\varepsilon > 0$ et $L = \frac{\log(1/\varepsilon)}{n}$.

5. **En utilisant la question précédente, montrer qu'avec probabilité au moins $1-\varepsilon$,**

$$\sigma^2 \leq \frac{1}{n}\sum_{i=1}^n (X_i - \mathbb{E}X)^2 + \sigma\sqrt{2L}.$$

6. **Montrer qu'avec probabilité au moins $1-2\varepsilon$, on a**

$$\big|\bar{X} - \mathbb{E}X\big| \leq \sigma\sqrt{2L} + \frac{L}{3} \tag{2}$$

et qu'avec probabilité au moins $1-3\varepsilon$, on a simultanément (2) et

$$\sigma^2 \leq \hat{\sigma}^2 + \big|\bar{X} - \mathbb{E}X\big|^2 + \sigma\sqrt{2L} \tag{3}$$

7. **L'objectif de cette question est de montrer que lorsque (2) et (3) sont vraies, on a**

$$\sigma \leq \hat{\sigma} + 1,8\sqrt{L}. \tag{4}$$

7.a Montrer que (4) est vraie pour $1,8\sqrt{L} \geq \frac{1}{2}$.

7.b Pour $1,8\sqrt{L} < \frac{1}{2}$, montrer

$$\sigma^2 \leq \hat{\sigma}^2 + \frac{\sqrt{L}\sigma}{3,6} + \frac{2}{3\times(3,6)^2}\sigma\sqrt{2L} + \frac{L}{9\times(3,6)^2} + \sigma\sqrt{2L},$$

et en déduire $\sigma \leq \frac{1,77}{2}\sqrt{L} + \sqrt{\hat{\sigma}^2 + 0,8L}$.

7.c Conclure.

8 Déduire des questions précédentes qu'avec probabilité au moins $1-\varepsilon$, on a

$$\left|\bar{X}-\mathbb{E}X\right| \leq \hat{\sigma}\sqrt{\frac{2\log(3\varepsilon^{-1})}{n}}+\frac{3\log(3\varepsilon^{-1})}{n},$$

et comparer avec l'inégalité (2).

****

3 Notions en Compressed Sensing

****

Exercice 3.1 (Notion de mesures incohérentes via le problème des douze pièces)

Cet exercice a pour but de se familiariser avec les notions de mesures, parcimonie, dictionnaire et incohérence au travers du problème des douze pièces. Pour plus de détails, on renvoie le lecteur intéressé à

site web sur le problème des douze pièces.

Le problème des douze pièces s'énonce de la manière suivante : on se donne douze pièces identiques en taille, forme, etc. mais dont une seule parmi les douze a une masse différente des onze autres. On ne sait pas si cette pièce est plus lourde ou plus légère que les autres ; on sait seulement qu'il y en a une qui est de masse différente des autres. On dispose d'un balance à deux plateaux. Comment retrouver cette pièce parmi les douze en ne faisant que 3 mesures au plus.

1. **Modéliser ce problèmes dans le cadre du CS et indiqué les différences avec le cadre classique du CS ;**
2. **Construire les mesures "triviales" permettant de résoudre le problème en 7 pesées ;**
3. **Définir la notion de parcimonie et le dictionnaire associé pour ce problème ;**
4. **proposer 3 vecteurs de mesure (non-adaptatifs) qui permettent de résoudre ce problème.**

****

****

Exercice 3.2 (Principe d'incertitude discret et Dirac Comb)

La matrice de Fourier (transformée de Fourier discrète) est donnée par :

$$\Gamma : \left\{ \begin{array}{ccc} \mathbb{C}^N & \to & \mathbb{C}^N \\ x & \to & \Gamma x = \hat{x} \end{array} \right. \quad \textbf{où } \Gamma = \left(\frac{w^{(p-1)(q-1)}}{\sqrt{N}}\right)_{1\leq p,q\leq N} \quad \textbf{et } w = \exp(-2\mathbf{i}\pi/N).$$

On note par $\bar{\Gamma}_1,\dots,\bar{\Gamma}_N$ les vecteurs lignes de Γ. On a donc $\hat{x}_i = \langle \Gamma_i, x\rangle$ pour tout $i=1,\dots,N$.

Le principe d'incertitude discret dit que si $x\in\mathbb{C}^N$ est non nul alors

$$|\text{supp}(x)| \times |\text{supp}(\hat{x})| \geq N.$$

En particulier, si x est un vecteur à petit support alors nécessairement sa transformée de Fourier $\hat{x}$ a un grand support. C'est la dualité classique entre vecteurs parcimonieux et transformée de Fourier "bien étalée".

L'objectif de l'exercice est de démontrer ce principe d'incertitude et de montrer qu'il est optimal pour les signaux *Dirac comb*.

1. **Écrire l'inégalité de Plancherel discrète.**
2. **Montrer que $\|x\|_\infty \le \frac{1}{\sqrt{N}} \|\hat{x}\|_1$**
3. **Déduire le résultat grâce aux inégalités**

$$\|x\|_2^2 \le |\mathrm{supp}(x)| \, \|x\|_\infty^2 \text{ et } \|\hat{x}\|_1 \le \sqrt{|\mathrm{supp}(\hat{x})|} \, \|\hat{x}\|_2 .$$

4. **On suppose que $N = k^2$ pour un certain entier k. On définit le signal x appelé *Dirac comb* comme étant l'indicateur de l'ensemble $\{0, k, 2k, \cdots, k^2 - k\}$ (ici les coordonnées sont indexées par $\{0, 1, \ldots, N-1\}$).**
   4.a Montrer que pour le Dirac comb, on a $x = \hat{x}$
   4.b Montrer que pour le Dirac comb, on a $|\mathrm{supp}(x)| \times |\mathrm{supp}(\hat{x})| = N$

****
****

Exercice 3.3 (Parcimonie du Basis Pursuit)
Soit $A \in \mathbb{R}^{m \times N}$ et $y \in \mathbb{R}^m$. On suppose qu'il existe une seule et unique solution au problème

$$\textbf{argmin}_{At=y} \|t\|_1 = \{\hat{x}\}. \tag{5}$$

On note par $a_1, \cdots, a_N$ les vecteurs colonnes de A dans $\mathbb{R}^m$.

1. **Montrer que le système $\{a_j : j \in \mathrm{supp}(\hat{x})\}$ est linéairement indépendant.**
2. **En déduire que $\|\hat{x}\|_0 \le m$.**

****
****

Exercice 3.4 (RIP et nombre minimal de mesures)
Soit $A \in \mathbb{R}^{m \times N}$ telle que pour tout $x \in \Sigma_{2s}$, $(1/2)\|x\|_2 \le \|Ax\|_2 \le (3/2)\|x\|_2$. Alors il existe des constantes absolues $c_0, c_1 > 0$ telles que $m \ge c_0 s \log(c_1 N/s)$.

****
****

Exercice 3.5 (Robustesse et borne inférieure de RIP)
Soit $A \in \mathbb{R}^{m \times N}$ une matrice de compression et $\Delta : \mathbb{R}^m \to \mathbb{R}^N$ un algorithme de décompression (non nécessairement linéaire). On dit que (A, Δ) est C-robuste d'ordre s quand pour tout $x \in \Sigma_s$ et tout $e \in \mathbb{R}^m$, on a

$$\|\Delta(Ax + e) - x\|_2 \le C \|e\|_2 .$$

En particulier, si (A,Δ) est C-robuste d'ordre s alors pour tout $x\in\Sigma_s$, on a $\Delta(Ax)=x$. Donc Δ permet la reconstruction exacte de tout élément x de Σ_s à partir de Ax. La robustesse demande en plus que Δ s'adapte aux erreurs dans les observations. C'est-à-dire que le signal reconstruit $\Delta(Ax+e)$ à partir des observations bruitées $Ax+e$ soit proche de x à un terme de l'ordre de grandeur du bruit près.

Montrer que si (A,Δ) est C-robuste d'ordre s alors la borne inférieure de $RIP(2s)$ est nécessaire, càd, pour tout $x\in\Sigma_{2s}$

$$\left\|Ax\right\|_2 \geq \frac{1}{2}\left\|x\right\|_2.$$

****

****

Exercice 3.6 (ℓ_2-stabilité)

Soit $A\in\mathbb{R}^{m\times N}$ une matrice de mesures et $\Delta:\mathbb{R}^m\to\mathbb{R}^N$ une procédure de décompression. On dit que (A,Δ) est C-ℓ_2-stable d'ordre s quand pour tout $x\in\mathbb{R}^N$,

$$\left\|x-\Delta(Ax)\right\|_2 \leq C\sigma_{s,2}(x) \text{ où } \sigma_{s,2}(x)=\min_{z\in\Sigma_s}\left\|x-z\right\|_2. \tag{6}$$

On va montrer que s'il existe (A,Δ) C-ℓ_2-stable d'ordre s (pour n'importe quel s, même $s=1$) alors nécessairement $m\geq\big(1-\sqrt{C^2-1}/C\big)N$.

1. **Montrer que pour tout $x\in\ker(A)$ et tout $j\in\{1,\dots,N\}$, $|x_i|\leq C'\left\|x\right\|_2$ pour $C'=\sqrt{C^2-1}/C$.**
2. **Soit L un sev de $\mathbb{R}^N$ tel que pour tout $x\in L$ et pour tout $i\in\{1,\dots,N\}$, $|x_i|\leq C'\left\|x\right\|_2$. Montre que $\dim(L)\leq C'N$.**
3. **En déduire que $m\geq(1-\sqrt{C^2-1}/C)N$**

Note : La stabilité d'une procédure est importante quand les signaux à reconstruire ne sont pas exactement parcimonieux mais seulement proche de Σ_s – ces signaux sont dit *compressibles*. Le résultat de cet exercice dit qu'on ne peut pas avoir (6) sauf si on a un nombre de mesures de l'ordre de N – ce qui n'est pas envisageable en CS. C'est pour cette raison qu'on a introduit la stabilité par rapport à $\sigma_{s,1}(x)$.

****

4 Optimisation

****

Exercice 4.1 (Optimisation convexe sous contraintes linéaires)

Soit le problème d'optimisation suivant

$$\min\Big(x^2+yz+y^2+z^2 : x\leq-1, z\leq-1\Big).$$

1. **Réduire ce problème de minimisation à un problème de la forme**

$$\min_{t \in K} f(t) \tag{P}$$

où f est une fonction à valeurs réelles et

$$K = \{t : g(t) \leq 0\}$$

avec $g : \mathbb{R}^2 \to \mathbb{R}$ convexe de classe $\mathcal{C}^1$.

2. **Montrer que f est convexe.**
3. **Montrer que la contrainte K est qualifiée.**
4. **Écrire le lagrangien de (P).**
5. **Écrire les conditions KKT de (P) et résoudre (P) à partir de ces équations.**
6. **Appliquer le théorème de dualité. Écrire le problème dual de (P), et retrouver ainsi la solution de (P) obtenue à la question 5).**

****

****

Exercice 4.2 (Système d'équations linéaire et optimisation)

Soit A une matrice symétrique définie positive de $\mathbb{R}^{N \times N}$ et $y \in \mathbb{R}^N$. Montrer que les deux assertions suivantes sont équivalentes :

1. **$\hat{x}$ est solution de $Ax = y$**
2. **$\hat{x}$ minimise la fonction $x \to F(x) = \frac{1}{2}x^\top Ax - x^\top y$.**

****