

## Correction détaillée

### Sujet BCE – Mathématiques appliquées – ESSEC/HEC 2026

Théorie de l'acquisition comprimée, lemme de Johnson-Lindenstrauss et minimisation  $\ell^1$

Cette correction suit le découpage de l'énoncé et donne, pour chaque question, une rédaction complète dans l'esprit attendu en ECG.

### Préliminaire informatique

On rappelle que pour  $X = (x_1, \dots, x_d) \in M_{d,1}(\mathbb{R})$ ,

$$\|X\| = \sqrt{\sum_{i=1}^d x_i^2}.$$

1.

a) Une fonction Python convenable est :

```
def Norme(X):
    return np.sqrt(np.sum(X**2))
```

b) On a

$$X = (1, 0, 1, 1, 1), \quad \|X\| = \sqrt{1 + 0 + 1 + 1 + 1} = 2.$$

Donc

$$\left\| \frac{1}{\|X\|} X \right\| = \left\| \frac{1}{2} X \right\| = \frac{1}{2} \|X\| = 1.$$

L'affichage obtenu est donc :

```
2.0
1.0
```

2.

On a

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad X = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Alors

$$AX = \begin{pmatrix} 2 \cdot 1 + 1 \cdot (-1) \\ 0 \cdot 1 + 1 \cdot (-1) \\ 1 \cdot 1 + 0 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

D'où

$$\|AX\|^2 = 1^2 + (-1)^2 + 1^2 = 3.$$

C'est ce qui explique le premier affichage : 3.0.

Par ailleurs,

$$\|X\|^2 = 1^2 + (-1)^2 = 2, \quad \frac{\|AX\|^2}{\|X\|^2} = \frac{3}{2} = 1.5.$$

Or, avec  $\text{eps} = 0.6$ , on a

$$1 - \text{eps} = 0.4, \quad 1 + \text{eps} = 1.6,$$

et bien

$$0.4 \leq 1.5 \leq 1.6.$$

Le second affichage est donc **True**.

### 3.

Une réponse possible est :

```
def EstIsom(A, LX, eps):
    for X in LX:
        if Norme(X) == 0:
            continue
        q = Norme(np.dot(A, X))**2 / Norme(X)**2
        if q < 1 - eps or q > 1 + eps:
            return False
    return True
```

Pour le script demandé, on teste les deux vecteurs de la liste

$$X_1 = (-1, 3), \quad X_2 = (4, -2).$$

Pour  $X_1$  :

$$AX_1 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix},$$

si bien que

$$\|AX_1\|^2 = 1 + 9 + 1 = 11, \quad \|X_1\|^2 = 1 + 9 = 10, \quad \frac{\|AX_1\|^2}{\|X_1\|^2} = 1.1 \in [0.8, 1.2].$$

Pour  $X_2$  :

$$AX_2 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \\ 4 \end{pmatrix},$$

donc

$$\|AX_2\|^2 = 36 + 4 + 16 = 56, \quad \|X_2\|^2 = 16 + 4 = 20, \quad \frac{\|AX_2\|^2}{\|X_2\|^2} = 2.8 \notin [0.8, 1.2].$$

La matrice n'est donc pas une  $(0.2, \mathcal{X}')$ -isométrie.

L'affichage obtenu est

**False**

## Le lemme de Johnson-Lindenstrauss

On considère une matrice aléatoire

$$M = \frac{1}{\sqrt{m}}(G_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n},$$

où les  $G_{i,j}$  sont indépendantes et de loi  $\mathcal{N}(0, 1)$ .

On fixe ensuite  $X = (x_1, \dots, x_n) \in M_{n,1}(\mathbb{R})$  tel que  $\|X\| = 1$ . Pour tout  $i \in \llbracket 1, m \rrbracket$ , on pose

$$Y_i = \sum_{j=1}^n x_j G_{i,j}, \quad Z = \|MX\|^2.$$

4.

Une simulation d'une telle matrice s'écrit par exemple :

`M = rd.normal(0, 1, (m, n)) / np.sqrt(m)`

5.

a) Si  $x_j \neq 0$  et si  $G_{i,j} \sim \mathcal{N}(0, 1)$ , alors la variable aléatoire  $x_j G_{i,j}$  suit la loi

$$\mathcal{N}(0, x_j^2).$$

b) Les variables  $x_j G_{i,j}$  sont indépendantes (car les  $G_{i,j}$  le sont) et gaussiennes centrées. Leur somme  $Y_i$  est donc gaussienne centrée, de variance

$$\sum_{j=1}^n x_j^2 = \|X\|^2 = 1.$$

Ainsi

$$Y_i \sim \mathcal{N}(0, 1).$$

En particulier,

$$\mathbb{E}(Y_i^2) = \text{Var}(Y_i) = 1.$$

c) Le vecteur  $MX$  a pour  $i$ -ième coordonnée

$$\frac{1}{\sqrt{m}} \sum_{j=1}^n x_j G_{i,j} = \frac{1}{\sqrt{m}} Y_i.$$

Ainsi

$$Z = \|MX\|^2 = \sum_{i=1}^m \left( \frac{Y_i}{\sqrt{m}} \right)^2 = \frac{1}{m} \sum_{i=1}^m Y_i^2.$$

6.

On fixe maintenant  $t \in ]0, 1/4]$ .

a) Comme  $Y_i \sim \mathcal{N}(0, 1)$ , sa densité est

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

Donc

$$\mathbb{E}(e^{tY_i^2}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ty^2} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(1-2t)y^2/2} dy.$$

Comme  $t \leq 1/4$ , on a  $1 - 2t > 0$ , donc l'intégrale converge bien.

b) Par changement de variable  $u = \sqrt{1-2t}y$ , on obtient

$$\int_{-\infty}^{+\infty} e^{-(1-2t)y^2/2} dy = \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{+\infty} e^{-u^2/2} du = \frac{\sqrt{2\pi}}{\sqrt{1-2t}}.$$

Ainsi

$$\mathbb{E}(e^{tY_i^2}) = \frac{1}{\sqrt{1-2t}}.$$

7.

On rappelle l'inégalité de Markov : si  $U \geq 0$  admet une espérance et si  $a > 0$ , alors

$$\mathbb{P}(U \geq a) \leq \frac{\mathbb{E}(U)}{a}.$$

a) Comme

$$tmZ = t \sum_{i=1}^m Y_i^2,$$

on a

$$\mathbb{E}(e^{tmZ}) = \mathbb{E}\left(\prod_{i=1}^m e^{tY_i^2}\right) = \prod_{i=1}^m \mathbb{E}(e^{tY_i^2}),$$

par indépendance des  $Y_i$ . Cette espérance existe donc, et vaut

$$\mathbb{E}(e^{tmZ}) = \left(\frac{1}{\sqrt{1-2t}}\right)^m.$$

En appliquant Markov à  $U = e^{tmZ}$  et à  $a = e^{tm(1+\varepsilon)}$ , on obtient

$$\mathbb{P}(Z \geq 1 + \varepsilon) = \mathbb{P}(e^{tmZ} \geq e^{tm(1+\varepsilon)}) \leq \frac{\mathbb{E}(e^{tmZ})}{e^{tm(1+\varepsilon)}}.$$

Donc

$$\mathbb{P}(Z > 1 + \varepsilon) \leq \mathbb{P}(Z \geq 1 + \varepsilon) \leq \left(\frac{e^{-t(1+\varepsilon)}}{\sqrt{1-2t}}\right)^m.$$

b) On vient donc d'établir

$$\mathbb{P}(Z > 1 + \varepsilon) \leq \exp\left(m\left[-t(1+\varepsilon) - \frac{1}{2}\ln(1-2t)\right]\right).$$

c) Posons

$$f(t) = 2t^2 + t + \frac{1}{2}\ln(1-2t).$$

Alors

$$f'(t) = 4t + 1 - \frac{1}{1-2t} = \frac{2t(1-4t)}{1-2t}.$$

Sur  $[0, 1/4]$ , on a  $1-2t > 0$  et  $1-4t \geq 0$ , donc  $f'(t) \geq 0$ . Comme  $f(0) = 0$ , on en déduit que  $f(t) \geq 0$  sur  $[0, 1/4]$ , c'est-à-dire

$$-t - \frac{1}{2}\ln(1-2t) \leq 2t^2.$$

Par suite,

$$-t(1+\varepsilon) - \frac{1}{2}\ln(1-2t) = \left(-t - \frac{1}{2}\ln(1-2t)\right) - \varepsilon t \leq 2t^2 - \varepsilon t.$$

Donc

$$\mathbb{P}(Z > 1 + \varepsilon) \leq e^{m(2t^2 - \varepsilon t)}.$$

d) On choisit  $t = \varepsilon/4$ . Comme  $\varepsilon \in ]0, 1[$ , on a bien  $t \in ]0, 1/4[$ . Alors

$$2t^2 - \varepsilon t = 2\frac{\varepsilon^2}{16} - \frac{\varepsilon^2}{4} = -\frac{\varepsilon^2}{8}.$$

Ainsi

$$\boxed{\mathbb{P}(Z > 1 + \varepsilon) \leq e^{-m\varepsilon^2/8}.$$

**8.**

Par un raisonnement parfaitement analogue, en appliquant Markov à la variable  $e^{-tmZ}$ , on obtient

$$\mathbb{P}(Z < 1 - \varepsilon) \leq e^{-m\varepsilon^2/8}.$$

On en déduit

$$\mathbb{P}(Z < 1 - \varepsilon \text{ ou } Z > 1 + \varepsilon) \leq \mathbb{P}(Z < 1 - \varepsilon) + \mathbb{P}(Z > 1 + \varepsilon) \leq 2e^{-m\varepsilon^2/8}.$$

**9.**

Soit  $\mathcal{X}' = \{X_1, \dots, X_r\}$  un ensemble fini de vecteurs de  $M_{n,1}(\mathbb{R})$ . Pour tout  $i \in \llbracket 1, r \rrbracket$ , on note  $B_i$  l'événement

$$(1 - \varepsilon) \|X_i\|^2 \leq \|MX_i\|^2 \leq (1 + \varepsilon) \|X_i\|^2.$$

a) Si  $X_i \neq 0$ , on pose  $U_i = \frac{1}{\|X_i\|} X_i$ . Alors  $\|U_i\| = 1$  et

$$MX_i = M(\|X_i\| U_i) = \|X_i\| MU_i,$$

d'où

$$\|MX_i\|^2 = \|X_i\|^2 \|MU_i\|^2.$$

Par conséquent,

$$B_i = [(1 - \varepsilon) \leq \|MU_i\|^2 \leq 1 + \varepsilon].$$

b) D'après la question 8, pour chaque  $i$  tel que  $X_i \neq 0$ ,

$$\mathbb{P}(B_i^c) \leq 2e^{-m\varepsilon^2/8}.$$

L'événement «  $M$  n'est pas une  $(\varepsilon, \mathcal{X}')$ -isométrie » est exactement

$$\bigcup_{i=1}^r B_i^c.$$

Donc, par sous-additivité,

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{X}')$$
-isométrie)  $\leq \sum_{i=1}^r \mathbb{P}(B_i^c) \leq 2re^{-m\varepsilon^2/8}.$

c) Si

$$m > \frac{8 \ln(2r)}{\varepsilon^2},$$

alors

$$2re^{-m\varepsilon^2/8} < 1.$$

La probabilité que  $M$  soit une  $(\varepsilon, \mathcal{X}')$ -isométrie est donc strictement positive. Il existe donc au moins une réalisation  $A \in M_{m,n}(\mathbb{R})$  qui soit une  $(\varepsilon, \mathcal{X}')$ -isométrie.

## 10.

Une implémentation possible consiste à prendre

$$m = \left\lceil \frac{8 \ln(2r)}{\varepsilon^2} \right\rceil + 1, \quad r = \text{len}(\text{LX}),$$

puis à simuler des matrices gaussiennes jusqu'à en trouver une qui convienne :

```
def Isom(LX, n, eps):
    r = len(LX)
    m = int(8*np.log(2*r)/(eps**2)) + 1
    A = rd.normal(0, 1, (m, n)) / np.sqrt(m)
    while not EstIsom(A, LX, eps):
        A = rd.normal(0, 1, (m, n)) / np.sqrt(m)
    return A
```

### $\varepsilon$ -isométries pour les ensembles sporadiques de $M_{n,1}(\mathbb{R})$

On note :

- $S(X)$  le support de  $X$  ;
- $(X)$  le nombre de composantes non nulles de  $X$  ;
- $\mathcal{C}_k = \{X \in M_{n,1}(\mathbb{R}) ; (X) \leq k\}$  ;
- $\mathcal{T}_k$  l'ensemble des parties de  $\llbracket 1, n \rrbracket$  de cardinal  $k$  ;
- pour  $T \in \mathcal{T}_k$ ,  $\mathcal{Q}(T)$  l'ensemble des vecteurs dont le support est inclus dans  $T$ .

## 11.

Choisir un élément de  $\mathcal{T}_k$ , c'est choisir  $k$  indices distincts parmi  $n$ . Donc

$$\#\mathcal{T}_k = \binom{n}{k}.$$

## 12.

a) On a

$$\mathcal{C}_0 = \{0\}, \quad \mathcal{C}_n = M_{n,1}(\mathbb{R}).$$

b) Soit  $T = \{i_1, \dots, i_k\} \in \mathcal{T}_k$ . Alors

$$\mathcal{Q}(T) = \text{Vect}(e_{i_1}, \dots, e_{i_k}),$$

où  $(e_1, \dots, e_n)$  est la base canonique. C'est donc un sous-espace vectoriel de  $M_{n,1}(\mathbb{R})$ , de dimension  $k$ .

c) Si  $X \in \mathcal{C}_k$ , alors  $(X) \leq k$ . On peut compléter son support en un ensemble  $T \in \mathcal{T}_k$  contenant  $S(X)$ , d'où  $X \in \mathcal{Q}(T)$ . Ainsi

$$\mathcal{C}_k = \bigcup_{T \in \mathcal{T}_k} \mathcal{Q}(T).$$

Réciproquement, si  $X \in \mathcal{Q}(T)$  pour un certain  $T \in \mathcal{T}_k$ , alors  $S(X) \subset T$ , donc  $(X) \leq k$ , donc  $X \in \mathcal{C}_k$ .

Pour  $1 \leq k < n$ , l'ensemble  $\mathcal{C}_k$  n'est pas un sous-espace vectoriel : par exemple

$$e_1 \in \mathcal{C}_1, \quad e_2 \in \mathcal{C}_1, \quad e_1 + e_2 \notin \mathcal{C}_1.$$

Plus généralement, dès que  $1 \leq k < n$ , on peut additionner deux vecteurs à supports disjoints de cardinal  $k$  et sortir de  $\mathcal{C}_k$ .

### 13. Inégalité de Cauchy-Schwarz

Soient  $X = (x_1, \dots, x_n)$  et  $Y = (y_1, \dots, y_n)$  deux vecteurs non nuls.

a) Pour tout  $i$ , on a

$$(|x_i| - |y_i|)^2 \geq 0.$$

En développant,

$$2|x_i y_i| \leq x_i^2 + y_i^2.$$

b) Si  $\|X\| = \|Y\| = 1$ , alors en sommant sur  $i$ ,

$$2 \sum_{i=1}^n |x_i y_i| \leq \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 = 2.$$

Donc

$$\sum_{i=1}^n |x_i y_i| \leq 1,$$

et par conséquent

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sum_{i=1}^n |x_i y_i| \leq 1.$$

c) Dans le cas général, posons

$$X' = \frac{X}{\|X\|}, \quad Y' = \frac{Y}{\|Y\|}.$$

Alors

$$\sum_{i=1}^n x_i y_i = \|X\| \|Y\| \sum_{i=1}^n x'_i y'_i.$$

Par le cas particulier précédent,

$$\left| \sum_{i=1}^n x'_i y'_i \right| \leq \|X'\| \|Y'\| = 1.$$

En élevant au carré, on obtient

$$\boxed{\left( \sum_{i=1}^n x_i y_i \right)^2 \leq \left( \sum_{i=1}^n x_i^2 \right) \left( \sum_{i=1}^n y_i^2 \right)}.$$

Si l'un des deux vecteurs est nul, l'inégalité reste évidemment vraie, puisque les deux membres valent alors 0.

### 14.

Soit  $A = (a_{i,j}) \in M_{m,n}(\mathbb{R})$ .

a) Si  $X = (x_1, \dots, x_n)$ , alors la  $i$ -ième coordonnée de  $AX$  vaut

$$\sum_{j=1}^n a_{i,j} x_j.$$

Donc

$$\boxed{\|AX\|^2 = \sum_{i=1}^m \left( \sum_{j=1}^n a_{i,j} x_j \right)^2}.$$

b) Pour tout  $i$ , la question 13 donne

$$\left(\sum_{j=1}^n a_{i,j}x_j\right)^2 \leq \left(\sum_{j=1}^n a_{i,j}^2\right) \left(\sum_{j=1}^n x_j^2\right) = \left(\sum_{j=1}^n a_{i,j}^2\right) \|X\|^2.$$

En sommant sur  $i$ ,

$$\|AX\|^2 \leq \left(\sum_{i=1}^m \sum_{j=1}^n a_{i,j}^2\right) \|X\|^2.$$

Si l'on pose

$$\alpha = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{i,j}^2},$$

alors

$$\boxed{\|AX\| \leq \alpha \|X\| \quad \text{pour tout } X \in M_{n,1}(\mathbb{R}).}$$

### 15.

Pour tous  $X, Y$ , on a

$$AX = AY + A(X - Y).$$

Par inégalité triangulaire,

$$\|AX\| \leq \|AY\| + \|A(X - Y)\|.$$

Et par l'inégalité triangulaire écrite sous forme inverse,

$$\|AY\| = \|AX - A(X - Y)\| \leq \|AX\| + \|A(X - Y)\|,$$

soit

$$\|AY\| - \|A(X - Y)\| \leq \|AX\|.$$

Finalement,

$$\boxed{\|AY\| - \|A(X - Y)\| \leq \|AX\| \leq \|AY\| + \|A(X - Y)\|.}$$

On admet maintenant que, pour tout  $k \in \llbracket 0, n \rrbracket$ , tout  $T \in \mathcal{T}_k$  et tout  $\delta \in ]0, 1]$ , il existe un sous-ensemble fini  $\mathcal{X}_{T,\delta}$  de  $\mathcal{Q}(T)$  tel que

$$\#\mathcal{X}_{T,\delta} < \left(\frac{12}{\delta}\right)^k,$$

et tel que pour tout  $X \in \mathcal{Q}(T)$  de norme 1, il existe  $Y \in \mathcal{X}_{T,\delta}$  de norme 1 vérifiant

$$\|X - Y\| \leq \frac{\delta}{4}.$$

### 16.

Soit  $\delta \in ]0, 1]$ . On définit  $(u_n)$  par

$$u_0 = \alpha - 1, \quad u_{n+1} = \frac{\delta}{4}(3 + u_n).$$

C'est une suite arithmético-géométrique. Si elle admet une limite  $\ell$ , alors

$$\ell = \frac{\delta}{4}(3 + \ell),$$

donc

$$\ell \left( 1 - \frac{\delta}{4} \right) = \frac{3\delta}{4} \implies \ell = \frac{3\delta}{4 - \delta}.$$

Comme  $0 < \delta \leq 1$ , on a  $\delta/4 < 1$ , donc la suite converge bien vers cette limite. Enfin,

$$\frac{3\delta}{4 - \delta} < \delta \iff 3 < 4 - \delta,$$

ce qui est vrai car  $\delta < 1$ . Ainsi

$$\boxed{\lim_{n \rightarrow \infty} u_n = \frac{3\delta}{4 - \delta} < \delta.}$$

On fixe désormais  $k \in \llbracket 0, n \rrbracket$ ,  $T \in \mathcal{T}_k$ , et  $\varepsilon, \delta \in ]0, 1]$  liés par

$$\varepsilon = 2\delta + \delta^2.$$

### 17.

On suppose que  $A$  est une  $(\delta/2, \mathcal{X}_{T,\delta})$ -isométrie.

a) Si  $Y \in \mathcal{X}_{T,\delta}$  et  $\|Y\| = 1$ , alors

$$1 - \frac{\delta}{2} \leq \|AY\|^2 \leq 1 + \frac{\delta}{2}.$$

En prenant les racines carrées,

$$\sqrt{1 - \frac{\delta}{2}} \leq \|AY\| \leq \sqrt{1 + \frac{\delta}{2}}.$$

Comme pour  $u \in [0, 1]$ , on a  $1 - u \leq \sqrt{1 - u} \leq 1 \leq \sqrt{1 + u} \leq 1 + u$ , on obtient

$$\boxed{1 - \frac{\delta}{2} \leq \sqrt{1 - \frac{\delta}{2}} \leq \|AY\| \leq \sqrt{1 + \frac{\delta}{2}} \leq 1 + \frac{\delta}{2}.$$

b) Montrons par récurrence la propriété

$$(P_n) : \quad \forall X \in \mathcal{Q}(T), \|X\| = 1 \implies 1 - u_n \leq \|AX\| \leq 1 + u_n.$$

*Initialisation.* Soit  $X \in \mathcal{Q}(T)$  avec  $\|X\| = 1$ . Par hypothèse sur  $\mathcal{X}_{T,\delta}$ , il existe  $Y \in \mathcal{X}_{T,\delta}$ , de norme 1, tel que

$$\|X - Y\| \leq \frac{\delta}{4}.$$

La question 15 donne

$$\|AX\| \leq \|AY\| + \|A(X - Y)\|.$$

Par la question 17.a,

$$\|AY\| \leq 1 + \frac{\delta}{2},$$

et par la question 14.b,

$$\|A(X - Y)\| \leq \alpha \|X - Y\| \leq \alpha \frac{\delta}{4}.$$

Ainsi

$$\|AX\| \leq 1 + \frac{\delta}{2} + \alpha \frac{\delta}{4} = 1 + \frac{\delta}{4}(3 + u_0) = 1 + u_1.$$

De même,

$$\|AX\| \geq \|AY\| - \|A(X - Y)\| \geq 1 - \frac{\delta}{2} - \alpha \frac{\delta}{4} = 1 - u_1.$$

Donc  $(P_1)$  est vraie.

*Hérédité.* Supposons  $(P_n)$  vraie. Alors, pour tout  $W \in \mathcal{Q}(T)$ ,

$$\|AW\| \leq (1 + u_n) \|W\|,$$

car si  $W \neq 0$ , on applique  $(P_n)$  à  $W/\|W\|$ .

Soit maintenant  $X \in \mathcal{Q}(T)$ ,  $\|X\| = 1$ , et choisissons  $Y \in \mathcal{X}_{T,\delta}$ ,  $\|Y\| = 1$ , tel que  $\|X - Y\| \leq \delta/4$ . Comme plus haut,

$$\|AX\| \leq \|AY\| + \|A(X - Y)\| \leq 1 + \frac{\delta}{2} + (1 + u_n) \frac{\delta}{4} = 1 + u_{n+1},$$

et de même

$$\|AX\| \geq 1 - \frac{\delta}{2} - (1 + u_n) \frac{\delta}{4} = 1 - u_{n+1}.$$

Donc  $(P_{n+1})$  est vraie.

Par récurrence,  $(P_n)$  est vraie pour tout  $n \geq 1$ .

**c)** D'après la question 16,  $u_n \rightarrow \ell$  avec  $\ell < \delta$ . En faisant tendre  $n$  vers l'infini dans  $(P_n)$ , on obtient, pour tout  $X \in \mathcal{Q}(T)$  de norme 1,

$$1 - \ell \leq \|AX\| \leq 1 + \ell.$$

A fortiori,

$$1 - \delta \leq \|AX\| \leq 1 + \delta.$$

En élevant au carré,

$$(1 - \delta)^2 \leq \|AX\|^2 \leq (1 + \delta)^2.$$

Or

$$(1 + \delta)^2 = 1 + 2\delta + \delta^2 = 1 + \varepsilon,$$

et

$$(1 - \delta)^2 = 1 - 2\delta + \delta^2 \geq 1 - (2\delta + \delta^2) = 1 - \varepsilon.$$

Donc

$$1 - \varepsilon \leq \|AX\|^2 \leq 1 + \varepsilon.$$

**d)** En homogénéisant, on en déduit immédiatement que pour tout  $X \in \mathcal{Q}(T)$ ,

$$(1 - \varepsilon) \|X\|^2 \leq \|AX\|^2 \leq (1 + \varepsilon) \|X\|^2.$$

Ainsi

$$\boxed{A \text{ est une } (\varepsilon, \mathcal{Q}(T))\text{-isométrie.}}$$

## 18.

On suppose désormais  $k \in \llbracket 1, n \rrbracket$ .

Pour  $T \in \mathcal{T}_k$ , si  $M$  n'est pas une  $(\varepsilon, \mathcal{Q}(T))$ -isométrie, alors d'après la question 17,  $M$  n'est pas une  $(\delta/2, \mathcal{X}_{T,\delta})$ -isométrie. Ainsi

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{Q}(T))\text{-isométrie}) \leq \mathbb{P}(M \text{ n'est pas une } (\delta/2, \mathcal{X}_{T,\delta})\text{-isométrie}).$$

Par la question 9.b,

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{Q}(T))\text{-isométrie}) \leq 2 \# \mathcal{X}_{T,\delta} e^{-m(\delta/2)^2/8} < 2 \left( \frac{12}{\delta} \right)^k e^{-m\delta^2/32}.$$

Si maintenant  $M$  est une  $(\varepsilon, \mathcal{Q}(T))$ -isométrie pour tout  $T \in \mathcal{T}_k$ , alors  $M$  est une  $(\varepsilon, \mathcal{C}_k)$ -isométrie, puisque tout  $X \in \mathcal{C}_k$  appartient à au moins un des  $\mathcal{Q}(T)$  avec  $T \in \mathcal{T}_k$ . D'où, par union finie,

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{C}_k)\text{-isométrie}) \leq \sum_{T \in \mathcal{T}_k} \mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{Q}(T))\text{-isométrie}).$$

Comme  $\#\mathcal{T}_k = \binom{n}{k}$ , on obtient

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{C}_k)\text{-isométrie}) < 2 \binom{n}{k} \left(\frac{12}{\delta}\right)^k e^{-m\delta^2/32}.$$

**19.**

a) On a

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!}.$$

b) En développant  $e^k$  en série entière,

$$e^k = \sum_{j=0}^{+\infty} \frac{k^j}{j!}.$$

Les termes d'indices  $k-1$  et  $k$  valent respectivement

$$\frac{k^{k-1}}{(k-1)!} = \frac{k^k}{k!} \quad \text{et} \quad \frac{k^k}{k!}.$$

Donc

$$e^k \geq 2 \frac{k^k}{k!}.$$

Par suite,

$$\frac{1}{k!} \leq \frac{e^k}{2k^k}.$$

En combinant avec a),

$$2 \binom{n}{k} \leq 2 \frac{n^k}{k!} \leq \left(\frac{en}{k}\right)^k.$$

c) Posons

$$a = \frac{m}{k}, \quad b = \frac{n}{k}.$$

D'après la question 18 puis la question 19.b,

$$\mathbb{P}(M \text{ n'est pas une } (\varepsilon, \mathcal{C}_k)\text{-isométrie}) < \left(\frac{12eb}{\delta}\right)^k e^{-ak\delta^2/32} = \exp\left(k \left[\ln\left(\frac{12eb}{\delta}\right) - \frac{a\delta^2}{32}\right]\right).$$

Si

$$a > \frac{32 \ln(12eb/\delta)}{\delta^2},$$

alors le crochet est négatif et la probabilité précédente est strictement inférieure à 1. Il existe donc au moins une réalisation  $A \in M_{m,n}(\mathbb{R})$  qui soit une  $(\varepsilon, \mathcal{C}_k)$ -isométrie.

## L'acquisition comprimée

On note désormais, pour  $X = (x_1, \dots, x_n)$ ,

$$|X| = \sum_{i=1}^n |x_i|,$$

qui est la norme  $\ell^1$  de  $X$ .

## 20. Quelques propriétés utiles

a) Pour tous  $X = (x_i)$  et  $Y = (y_i)$ ,

$$|X + Y| = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) = |X| + |Y|.$$

On en déduit par récurrence que, pour tous  $X_1, \dots, X_r$ ,

$$\left| \sum_{i=1}^r X_i \right| \leq \sum_{i=1}^r |X_i|.$$

b) Si les supports des  $X_1, \dots, X_r$  sont deux à deux disjoints, alors, coordonnée par coordonnée, il n'y a jamais de compensation. Ainsi

$$\left| \sum_{i=1}^r X_i \right| = \sum_{i=1}^r |X_i|.$$

c) Si  $X \in \mathcal{C}_k$ , alors  $X$  possède au plus  $k$  composantes non nulles. En appliquant Cauchy-Schwarz aux  $k$  nombres non nuls  $|x_i|$ , on obtient

$$|X| = \sum_{i \in S(X)} |x_i| \leq \sqrt{\#S(X)} \sqrt{\sum_{i \in S(X)} x_i^2} \leq \sqrt{k} \|X\|.$$

Et comme toujours  $\|X\| \leq |X|$ , on conclut

$$\boxed{\|X\| \leq |X| \leq \sqrt{k} \|X\| \quad (X \in \mathcal{C}_k).}$$

## 21.

Soit  $A \in M_{m,n}(\mathbb{R})$  une  $(\varepsilon, \mathcal{C}_k)$ -isométrie, avec  $\varepsilon \in ]0, 1[$ .

a) Si  $X \in \mathcal{C}_k$  et  $AX = 0$ , alors

$$0 = \|AX\|^2 \geq (1 - \varepsilon) \|X\|^2.$$

Comme  $1 - \varepsilon > 0$ , on obtient  $\|X\| = 0$ , donc  $X = 0$ .

b) Considérons les  $k$  premiers vecteurs de la base canonique  $e_1, \dots, e_k$ . Si

$$\lambda_1 Ae_1 + \dots + \lambda_k Ae_k = 0,$$

alors

$$A(\lambda_1 e_1 + \dots + \lambda_k e_k) = 0.$$

Le vecteur  $\lambda_1 e_1 + \dots + \lambda_k e_k$  appartient à  $\mathcal{C}_k$ , donc, d'après a), il est nul. Les vecteurs  $Ae_1, \dots, Ae_k$  sont donc libres. Ainsi

$$\boxed{\text{rg}(A) \geq k.}$$

## 22. Première caractérisation

On suppose ici que  $A$  est une  $(\varepsilon, \mathcal{C}_{2k})$ -isométrie.

a) Si  $X, Y \in \mathcal{C}_k$  et  $AX = AY$ , alors

$$A(X - Y) = 0.$$

Comme  $X - Y \in \mathcal{C}_{2k}$ , la question 21.a appliquée à  $2k$  donne

$$X - Y = 0.$$

Donc

$$\boxed{X = Y.}$$

b) Soit  $Y \in \mathcal{C}_k$  et  $B = AY$ . Le vecteur  $Y$  est une solution de  $AX = B$ . Soit maintenant  $X$  une solution de  $AX = B$  qui minimise  $(X)$  parmi toutes les solutions. Comme  $Y$  est solution,

$$(X) \leq (Y) \leq k,$$

donc  $X \in \mathcal{C}_k$ . Or  $AX = AY$ , donc la question a) donne  $X = Y$ .

Ainsi  $Y$  est l'unique solution de l'équation  $AX = B$  qui minimise le nombre de composantes non nulles.

### 23–27. Deuxième caractérisation

On suppose maintenant que  $k \leq n/3$  et que  $A$  est une  $(\varepsilon, \mathcal{C}_{3k})$ -isométrie avec  $\varepsilon < 1/3$ .

Soit  $Y \in \mathcal{C}_k$  et  $B = AY$ . On considère un vecteur  $X \in M_{n,1}(\mathbb{R})$  tel que

$$AX = B \quad \text{et} \quad |X| \leq |Y|.$$

On pose

$$Z = Y - X.$$

Alors

$$AZ = AY - AX = 0.$$

On note  $S$  le support de  $Y$  et  $\bar{S}$  l'ensemble des indices des composantes nulles de  $Y$ . Pour tout sous-ensemble non vide  $I \subset \llbracket 1, n \rrbracket$ , on note  $Z_I$  le vecteur obtenu à partir de  $Z$  en annulant les composantes d'indice hors de  $I$ .

### 23.

Si  $(Z) \leq 3k$ , alors  $Z \in \mathcal{C}_{3k}$  et  $AZ = 0$ . La question 21.a appliquée à  $3k$  donne immédiatement

$$\boxed{Z = 0.}$$

Dans toute la suite, on suppose donc

$$3k + 1 \leq (Z).$$

### 24.

a) Comme  $Z = Z_S + Z_{\bar{S}}$ , on a

$$X = Y - Z = Y - Z_S - Z_{\bar{S}},$$

donc

$$X + Z_S = Y - Z_{\bar{S}}.$$

Autrement dit,

$$\boxed{Y - Z_{\bar{S}} = X + Z_S.}$$

De plus, les supports de  $Y$  et de  $Z_{\bar{S}}$  sont disjoints, donc par la question 20.b,

$$\boxed{|Y - Z_{\bar{S}}| = |Y| + |Z_{\bar{S}}|.}$$

b) En utilisant a) puis l'inégalité triangulaire pour la norme  $\ell^1$ ,

$$|Y| + |Z_{\bar{S}}| = |Y - Z_{\bar{S}}| = |X + Z_S| \leq |X| + |Z_S| \leq |Y| + |Z_S|.$$

D'où

$$\boxed{|Z_{\bar{S}}| \leq |Z_S| .}$$

c) Le vecteur  $Z_S$  appartient à  $\mathcal{C}_k$  car  $S$  a au plus  $k$  éléments. Par la question 20.c,

$$\|Z_S\| \geq \frac{1}{\sqrt{k}} |Z_S| .$$

Avec b), on obtient

$$\boxed{\|Z_S\| \geq \frac{1}{\sqrt{k}} |Z_S| \geq \frac{1}{\sqrt{k}} |Z_{\bar{S}}| .}$$

## 25.

On pose

$$r = \left\lfloor \frac{(Z_{\bar{S}})}{2k} \right\rfloor .$$

On partitionne  $\bar{S}$  en réunion disjointe

$$T_1 \cup \dots \cup T_{r+1},$$

avec  $(Z_{T_i}) = 2k$  pour  $1 \leq i \leq r$  et  $(Z_{T_{r+1}}) < 2k$ , les composantes de  $Z_{T_i}$  étant classées par valeurs absolues décroissantes d'un bloc au suivant. On note  $\alpha_i$  la plus petite valeur absolue des composantes non nulles de  $Z_{T_i}$ .

a) Pour  $i \in \llbracket 1, r \rrbracket$ , toutes les composantes non nulles de  $Z_{T_i}$  ont une valeur absolue au moins égale à  $\alpha_i$ , et elles sont au nombre de  $2k$ . Donc

$$|Z_{T_i}| \geq 2k\alpha_i = \sqrt{2k} \sqrt{2k\alpha_i^2} .$$

Par ailleurs, toutes les composantes non nulles de  $Z_{T_{i+1}}$  ont une valeur absolue au plus égale à  $\alpha_i$ , et il y en a au plus  $2k$ , donc

$$\|Z_{T_{i+1}}\|^2 \leq 2k\alpha_i^2 .$$

Ainsi

$$\boxed{|Z_{T_i}| \geq \sqrt{2k} \sqrt{2k\alpha_i^2} \geq \sqrt{2k} \|Z_{T_{i+1}}\| .}$$

b) On a

$$|Z_{\bar{S}}| = \sum_{i=1}^{r+1} |Z_{T_i}| .$$

D'après a), pour  $i \geq 2$ ,

$$|Z_{T_{i-1}}| \geq \sqrt{2k} \|Z_{T_i}\| .$$

Donc

$$|Z_{\bar{S}}| \geq \sqrt{2k} \sum_{i=2}^{r+1} \|Z_{T_i}\| .$$

En combinant avec la question 24.c,

$$\|Z_S\| \geq \frac{1}{\sqrt{k}} |Z_{\bar{S}}| \geq \sqrt{2} \sum_{i=2}^{r+1} \|Z_{T_i}\| .$$

Ainsi

$$\boxed{\|Z_S\| \geq \sqrt{2} \sum_{i=2}^{r+1} \|Z_{T_i}\| .}$$

26.

a) On décompose

$$Z = Z_{S \cup T_1} + \sum_{i=2}^{r+1} Z_{T_i}.$$

Donc

$$AZ = AZ_{S \cup T_1} + \sum_{i=2}^{r+1} AZ_{T_i}.$$

Par inégalité triangulaire,

$$\|AZ\| \geq \|AZ_{S \cup T_1}\| - \left\| \sum_{i=2}^{r+1} AZ_{T_i} \right\| \geq \|AZ_{S \cup T_1}\| - \sum_{i=2}^{r+1} \|AZ_{T_i}\|.$$

Or

$$(S \cup T_1) \leq k + 2k = 3k, \quad (T_i) \leq 2k \leq 3k,$$

donc, puisque  $A$  est une  $(\varepsilon, \mathcal{C}_{3k})$ -isométrie,

$$\|AZ_{S \cup T_1}\| \geq \sqrt{1-\varepsilon} \|Z_{S \cup T_1}\|, \quad \|AZ_{T_i}\| \leq \sqrt{1+\varepsilon} \|Z_{T_i}\|.$$

Ainsi

$$\boxed{\|AZ\| \geq \sqrt{1-\varepsilon} \|Z_{S \cup T_1}\| - \sqrt{1+\varepsilon} \sum_{i=2}^{r+1} \|Z_{T_i}\|.$$

b) Les supports de  $Z_S$  et de  $Z_{T_1}$  sont disjoints, donc

$$\|Z_{S \cup T_1}\|^2 = \|Z_S\|^2 + \|Z_{T_1}\|^2 \geq \|Z_S\|^2,$$

ainsi  $\|Z_{S \cup T_1}\| \geq \|Z_S\|$ . D'autre part, la question 25.b donne

$$\sum_{i=2}^{r+1} \|Z_{T_i}\| \leq \frac{1}{\sqrt{2}} \|Z_S\|.$$

Donc

$$\|AZ\| \geq \left( \sqrt{1-\varepsilon} - \frac{\sqrt{1+\varepsilon}}{\sqrt{2}} \right) \|Z_S\|.$$

Autrement dit,

$$\boxed{\|AZ\| \geq \left( \sqrt{1-\varepsilon} - \frac{\sqrt{1+\varepsilon}}{\sqrt{2}} \right) \|Z_S\|.$$

c) Or  $AZ = 0$ , donc le membre de gauche vaut 0. Comme  $\varepsilon < 1/3$ ,

$$\sqrt{1-\varepsilon} - \frac{\sqrt{1+\varepsilon}}{\sqrt{2}} > 0,$$

car cela équivaut à

$$1 - \varepsilon > \frac{1 + \varepsilon}{2} \iff 1 > 3\varepsilon.$$

On en déduit nécessairement

$$\|Z_S\| = 0, \quad \text{donc} \quad Z_S = 0.$$

Puis la question 24.b donne

$$0 \leq |Z_{\bar{S}}| \leq |Z_S| = 0,$$

donc  $Z_{\bar{S}} = 0$ . Finalement  $Z = 0$ , donc

$$\boxed{X = Y.}$$

27.

On vient de montrer que toute solution  $X$  de l'équation  $AX = B$  vérifiant

$$|X| \leq |Y|$$

est forcément égale à  $Y$ . Comme  $Y$  lui-même est solution de  $AX = B$ , il est donc l'unique solution qui minimise  $|X|$ .

Ainsi,

$Y$  est l'unique solution de  $AX = B$  qui minimise  $|X|$ .

---

**Bilan.** Le sujet établit successivement :

- l'existence, avec grande probabilité, de projections aléatoires qui préservent presque la norme euclidienne sur des ensembles finis (lemme de Johnson-Lindenstrauss) ;
- l'extension de cette propriété aux ensembles de vecteurs  $k$ -sporadiques ;
- deux principes de reconstruction en acquisition comprimée :
  - la minimisation du nombre de composantes non nulles sous une hypothèse de type  $(\varepsilon, \mathcal{C}_{2k})$  ;
  - la minimisation de la norme  $\ell^1$  sous une hypothèse de type  $(\varepsilon, \mathcal{C}_{3k})$  avec  $\varepsilon < 1/3$ .