

Conditionnement d'une matrice et applications

Dans tout ce problème, n désigne un entier naturel non nul, et on rappelle que $\mathcal{M}_n(\mathbb{R})$ désigne l'ensemble des matrices carrées à n lignes et n colonnes. On note $\mathcal{D}_n(\mathbb{R})$ le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ des matrices diagonales.

On rappelle que l'on désigne par M^T la transposée d'une matrice M .

Pour alléger les notations, on identifiera les vecteurs de $\mathbb{R}^n$ aux matrices colonnes de $\mathcal{M}_{n,1}(\mathbb{R})$.

On désignera par $B = (E_1, E_2, \dots, E_n)$ la base canonique de $\mathbb{R}^n$.

On munit $\mathbb{R}^n$ de la norme $\|\cdot\|$, en posant pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ qui est la norme euclidienne

associée au produit scalaire canonique $\langle \cdot, \cdot \rangle$ de $\mathbb{R}^n$ où par définition, pour tout X et Y de $\mathbb{R}^n$, $\langle X, Y \rangle = X^T Y$.

Pour toute matrice M de $\mathcal{M}_n(\mathbb{R})$ on note $\rho(M)$ le réel défini par : $\rho(M) = \max_{\lambda \in \text{Sp}_{\mathbb{C}}(M)} |\lambda|$.

On note par ailleurs $S_n^+(\mathbb{R})$ l'ensemble des matrices symétriques positives de $\mathcal{M}_n(\mathbb{R})$ et par $S_n^{++}(\mathbb{R})$ l'ensemble des matrices symétriques définies positives de $\mathcal{M}_n(\mathbb{R})$.

Partie A – Construction d'une norme sur $\mathcal{M}_n(\mathbb{R})$

On se propose dans cette partie de montrer que l'application N donnée sur $\mathcal{M}_n(\mathbb{R})$ par :

$$N : A \longmapsto \sup_{\|X\|=1} \|AX\|$$

est une norme sur $\mathcal{M}_n(\mathbb{R})$ et d'en étudier quelques propriétés.

I – Étude de l'application N

Dans toute cette partie, on considère A une matrice quelconque de $\mathcal{M}_n(\mathbb{R})$ dont on note $L_1, L_2, \dots, L_n$ les n lignes et $C_1, C_2, \dots, C_n$ les n colonnes, que l'on pourra identifier à des éléments de $\mathbb{R}^n$.

Q1. Soit $X \in \mathbb{R}^n$ tel que $\|X\| = 1$. En notant $M = \max_{1 \leq i \leq n} \|L_i\|$, montrer que :

$$\|AX\| \leq M\sqrt{n}.$$

On pourra au préalable s'intéresser à la i^{e} ligne de la matrice AX et utiliser l'inégalité de Cauchy-Schwarz pour les vecteurs de $\mathbb{R}^n$.

Q2. En déduire que l'application N est bien définie, puis que : $N(A) = \sup_{X \neq 0} \frac{\|AX\|}{\|X\|}$.

Q3. Montrer que l'application N ainsi définie est une norme sur $\mathcal{M}_n(\mathbb{R})$.

Q4. En est-il de même pour l'application $S : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \rightarrow \mathbb{R}_+ \\ M & \mapsto \rho(M) \end{cases}$?

Q5. Soit $\Delta \in \mathcal{D}_n(\mathbb{R})$ dont on note $\delta_1, \dots, \delta_n$ les termes diagonaux.

Vérifier que $N(\Delta) = \max_{1 \leq i \leq n} |\delta_i|$.

Q6. À l'aide de l'application $X \mapsto \|AX\|$, démontrer que : $N(A) = \max_{\|X\|=1} \|AX\|$.

Q7. Établir que : $\forall X \in \mathbb{R}^n, \|AX\| \leq N(A) \|X\|$.

Q8. Soit B une autre matrice quelconque de $\mathcal{M}_n(\mathbb{R})$. Montrer que :

$$N(AB) \leq N(A) N(B).$$

Q9. Montrer que : $\max_{1 \leq i \leq n} \|C_i\| \leq N(A)$.

Q10. Déterminer $N(A)$ dans le cas où toutes les colonnes de A sont nulles, sauf la dernière.

En déduire $N(A)$ dans le cas où $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

II – Cas des matrices orthogonales et symétriques

Dans cette partie, A désigne une matrice quelconque de $\mathcal{M}_n(\mathbb{R})$ et U une matrice orthogonale de $\mathcal{M}_n(\mathbb{R})$.

Q11. Déterminer $N(U)$.

Q12. Démontrer que $N(UA)$ et $N(A)$ sont égales.

Q13. En considérant $X_0 \in \mathbb{R}^n$ où $\|X_0\| = 1$ tel que $\|AX_0\| = N(A)$, démontrer que $N(AU) = N(A)$.

Q14. On suppose de plus dans cette question uniquement que la matrice A est une matrice symétrique réelle de $\mathcal{M}_n(\mathbb{R})$.

Montrer que : $N(A) = \rho(A)$.

Q15. Déterminer $N(A)$ dans le cas où $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Partie B – Conditionnement d'une matrice pour la norme N

On définit sur $\text{GL}_n(\mathbb{R})$ l'application notée cond par : $\text{cond} : \begin{cases} \text{GL}_n(\mathbb{R}) & \rightarrow \mathbb{R} \\ A & \mapsto N(A) N(A^{-1}) \end{cases}$

I – Quelques résultats sur le conditionnement

Dans toute cette sous-partie, A désigne une matrice inversible de $\mathcal{M}_n(\mathbb{R})$ et U une matrice orthogonale de $\mathcal{M}_n(\mathbb{R})$.

Q16. Montrer que : $1 \leq \text{cond}(A)$.

Q17. Quel lien a-t-on entre $\text{cond}(A)$ et $\text{cond}(\alpha A)$ pour $\alpha \in \mathbb{R}^*$?

Q18. Démontrer que $\text{cond}(U) = 1$.

Q19. Que dire de $\text{cond}(UA)$, $\text{cond}(AU)$ et de $\text{cond}(A)$?

II – Un exemple de minoration du conditionnement d'une matrice

On suppose dans cette partie uniquement que $A = (a_{i,j})_{\substack{1 \leq i,j \leq n}} \text{ où : } a_{i,j} = \begin{cases} 1 & \text{si } i=j \\ 2 & \text{si } j=i+1 \\ 0 & \text{sinon} \end{cases}$.

Q20. On considère le vecteur X de $\mathbb{R}^n$ donné par : $X = \sum_{k=1}^n (-1)^{n-k} 2^{n-k} E_k$.

Montrer que $AX = E_n$.

Q21. Dédurre de ce qui précède que $N(A^{-1}) \geq 2^{n-1}$.

Q22. Justifier $\|AE_2\| > 2$, pour en déduire que $\text{cond}(A) > 2^n$.

Partie C – Conditionnement pour une matrice réelle inversible

Q23. Soit S une matrice de $S_n^+(\mathbb{R})$.

On considère $\mathcal{C} = (V_1, \dots, V_n)$ une base diagonalisante orthonormée de $\mathbb{R}^n$ où pour tout $i \in \{1, \dots, n\}$, V_i est un vecteur propre associé à la valeur propre notée λ_i et où l'on suppose que $\lambda_1 \leq \dots \leq \lambda_n$ sont les valeurs propres de S comptées avec leur ordre de multiplicité.

Montrer que : $N(S) = \max_{\|X\|=1} |(SX, X)|$.

Q24. Soit $A \in M_n(\mathbb{R})$ non nulle.

Démontrer que la matrice $A^T A$ appartient à $S_n^+(\mathbb{R})$ pour établir que $N(A^T A) = N(A)^2$.

Q25. Dédurre de ce qui précède que pour $A \in M_n(\mathbb{R})$ non nulle : $N(A) = \sqrt{\rho(A^T A)}$.

Q26. On suppose dans cette question que A est une matrice $M_n(\mathbb{R})$ inversible.

En remarquant que $A^T A = A^{-1} A A^T A$, démontrer que les matrices $A A^T$ et $A^T A$ ont exactement les mêmes valeurs propres.

Q27. Soit $A \in M_n(\mathbb{R})$ inversible. On note μ_m et μ_M respectivement la plus petite et la plus grande des valeurs propres de la matrice $A^T A$ et où l'on suppose que l'on a $0 < \mu_m \leq \mu_M$.

Montrer que : $\text{cond}(A) = \sqrt{\frac{\mu_M}{\mu_m}}$.

Q28. Exprimer $\text{cond}(A)$ lorsque A appartient à $S_n^{++}(\mathbb{R})$ à l'aide des valeurs propres de A en remarquant que $A^T A = A^2$.

Partie D – Calcul explicite de conditionnement

Dans toute cette partie, on désigne par T la matrice de $M_n(\mathbb{R})$ donnée par : $T = \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix}$.

Le but de cette partie est de déterminer la valeur de $\text{cond}(T)$ en commençant par déterminer les éléments propres de la matrice T .

Q29. Montrer que les valeurs propres de T sont réelles.

Q30. Soit $k \in \mathbb{N}$ tel que $k \notin (n+1)\mathbb{Z}$. On considère le vecteur U_k de $\mathbb{R}^n$ donné par :

$$U_k = \left(\sin\left(\frac{k\pi}{n+1}\right), \sin\left(\frac{2k\pi}{n+1}\right), \dots, \sin\left(\frac{(n-1)k\pi}{n+1}\right), \sin\left(\frac{nk\pi}{n+1}\right) \right).$$

Montrer que U_k est un vecteur propre de T et préciser la valeur propre associée.

Q31. En déduire l'ensemble des valeurs propres de T .

Q32. Déterminer alors la valeur de $\text{cond}(T)$.

Partie E – Inégalité de Kantorovich

Dans toute cette partie, A désigne une matrice de $S_n^{++}(\mathbb{R})$ et on désigne par $\lambda_1, \dots, \lambda_n$ l'ensemble de ses valeurs propres où l'on suppose que $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ et comptées avec leur ordre de multiplicité, et on désigne par $\mathcal{C} = (V_1, \dots, V_n)$ une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de A .

On se propose d'établir le résultat suivant, appelée inégalité de Kantorovich :

$$(K) : \quad \forall X \in \mathbb{R}^n, \|X\|^4 \leq \langle AX, X \rangle \langle A^{-1}X, X \rangle \leq \frac{1}{4} \left(\frac{1}{\sqrt{\text{cond}(A)}} + \sqrt{\text{cond}(A)} \right)^2 \|X\|^4.$$

I – Une première démonstration

On désigne par P le polynôme de $\mathbb{R}[X]$ donné par $P = X^2 - (\lambda_1 + \lambda_n)X + \lambda_1\lambda_n$.

Q33. Exprimer $\text{cond}(A)$ à l'aide des valeurs propres de A .

Q34. On admet que l'application $(\cdot, \cdot)_A : \begin{cases} \mathbb{R}^n \times \mathbb{R}^n & \rightarrow \mathbb{R} \\ (X, Y) & \mapsto \langle AX, Y \rangle \end{cases}$ est un produit scalaire sur $\mathbb{R}^n$.

À l'aide de l'inégalité de Cauchy-Schwarz, démontrer que : $\forall X \in \mathbb{R}^n, \|X\|^4 \leq \langle AX, X \rangle \langle A^{-1}X, X \rangle$.

Q35. Montrer que : $\forall k \in \{1, \dots, n\}, P(\lambda_k) \leq 0$.

Q36. Déterminer les valeurs propres de la matrice $B = A^{-1}P(A)$ et en déduire que $\langle BX, X \rangle \leq 0$ pour tout $X \in \mathbb{R}^n$.

Q37. Pour $X \in \mathbb{R}^n$ fixé, on désigne par f la fonction polynôme de degré 2 définie par :

$$f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ \lambda & \mapsto \langle AX, X \rangle \lambda^2 - (\lambda_1 + \lambda_n) \|X\|^2 \lambda + \lambda_1 \lambda_n \langle A^{-1}X, X \rangle \end{cases}$$

Vérifier que $f(1) = \langle BX, X \rangle$, montrer que $f(0)f(1) \leq 0$, puis établir que :

$$(*) : (\lambda_1 + \lambda_n)^2 \|X\|^4 - 4 \langle AX, X \rangle \langle A^{-1}X, X \rangle \lambda_1 \lambda_n \geq 0.$$

Q38. Dédurre de ce qui précède l'inégalité de Kantorovich.

II – Une deuxième démonstration

On admet que, pour établir la relation (K), il suffit de la vérifier pour un vecteur X de norme 1.

Dans toute cette partie, $X = (x_1, \dots, x_n)$ désigne donc un vecteur de $\mathbb{R}^n$ de norme 1 dont les coordonnées sont données dans la base $\mathcal{C}$.

On considère alors un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et on définit la variable aléatoire Z par :

$$Z(\Omega) = \{\lambda_1, \dots, \lambda_n\} \text{ et : } \forall i \in \{1, \dots, n\}, \mathbb{P}(|Z = \lambda_i|) = x_i^2$$

Q39. Justifier que l'on définit bien une loi de probabilité pour Z .

Q40. Justifier que Z et $\frac{1}{Z}$ admettent une espérance, puis les exprimer en fonction de $\langle AX, X \rangle$ et de $\langle A^{-1}X, X \rangle$.

Q41. En remarquant que la variable aléatoire $(Z - \lambda_1)(Z - \lambda_n)$ est négative, établir l'inégalité suivante :

$$\frac{1}{Z} \leq \frac{\lambda_1 + \lambda_n - Z}{\lambda_1 \lambda_n}.$$

Q42. En déduire alors que : $\mathbb{E}(Z) \mathbb{E}\left(\frac{1}{Z}\right) \leq -\frac{1}{\lambda_1 \lambda_n} \left(\mathbb{E}(Z) - \frac{\lambda_1 + \lambda_n}{2} \right)^2 + \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n}$.

Q43. Dédurre de ce qui précède la seconde partie de l'inégalité de Kantorovich.

♦ Fin ♦

Partie A. Construction d'une norme sur $\mathcal{M}_n(\mathbb{R})$

I. Etude de l'application N sur $\mathcal{M}_n(\mathbb{R})$

1. Soit $X \in \mathbb{R}^n$ unitaire. La ligne i du vecteur AX s'écrit :

$$(AX)_i = \sum_{j=1}^n a_{i,j} x_j = \langle L_i, X \rangle$$

D'après Cauchy-Schwartz, on a :

$$\langle L_i, X \rangle \leq \|L_i\| \|X\| \leq M$$

Puis :

$$\|AX\| = \sqrt{\sum_{i=1}^n (AX)_i^2} \leq \sqrt{\sum_{i=1}^n M^2} = M\sqrt{n}$$

2. D'après la question précédente, pour tout A on a la quantité $\|A\|$ qui est finie donc l'application N est bien définie sur $\mathcal{M}_n(\mathbb{R})$. On pose S^1 la sphère unité sur $\mathbb{R}^n$, on a $N(A) = \sup_{X \in S^1} \|AX\|$. Notons $T = \{\frac{X}{\|X\|} \mid X \in \mathbb{R}^n \setminus \{0\}\}$. En effet $X \neq 0$ équivaut à $\|X\| \neq 0$. Si $Y \in T$, alors il est clair que $\|Y\| = 1$ et $Y \in S^1$ donc $T \subset S^1$. De plus, si $X \in S^1$, alors $\|X\| = 1$ donc $X/\|X\| = X$ et $X \in T$ d'où $T = S^1$ et donc $N(A) = \sup_{Y \in T} \|AY\| = \sup_{X \in \mathbb{R}^n \setminus \{0\}} \|A \frac{X}{\|X\|}\|$ puis par homogénéité de $X \mapsto \|X\|$ sur $\mathbb{R}^n$, on obtient le résultat. (un peu bizarre comme question parce que ça paraît vraiment trivial comme résultat)
3. On le fait brièvement :
- **Homogénéité** : On utilise l'homogénéité de $\|\cdot\|$ sur $\mathbb{R}^n$ et l'homogénéité du sup (je sais pas si c'est au programme).
 - **Séparabilité** : si $N(A) = 0$, alors pour tout $X \neq 0$ (d'après la question précédente) $\|AX\|/\|X\| = 0$ or $\|X\| \neq 0$ donc $\|AX\| = 0$ et donc $AX = 0$ pour tout $X \neq 0$. On en déduit $A = 0$.
 - **Inégalité triangulaire** : On utilise l'inégalité triangulaire de $\|\cdot\|$. $\|(A+B)X\| = \|AX + BX\| \leq \|AX\| + \|BX\| \leq N(A) + N(B)$ pour $\|X\| = 1$ et en passant au sup sur S^1 on obtient le résultat.
4. Non, une matrice triangulaire supérieure avec que des 0 sur la diagonale vérifie $S(M) = 0$ car $\text{Sp}(M) = \{0\}$ mais elle n'est pas forcément nulle.
5. Notons $\delta = \max_i \delta_i$. On a pour X tel que $\|X\| = 1$:

$$\|\Delta X\|^2 = \sum_{i=1}^n \delta_i^2 x_i^2 \leq \delta^2 \|X\|^2 = \delta^2$$

donc $\|\Delta X\| \leq \delta > 0$ (on ne s'occupe pas du cas où $\Delta = 0$ qui n'a pas d'intérêt) et donc $N(\Delta) \leq \delta$. Puisque Δ est supposée non nulle, il existe i_0 tel que $\delta = |\delta_{i_0}|$. Ainsi en prenant le vecteur X_{i_0} qui n'a que des 0 sauf 1 sur sa $i_0^{ème}$ composante (et donc $\|X_{i_0}\| = 1$), on obtient :

$$\|\Delta X_{i_0}\| = |\delta_{i_0}| = \delta$$

D'où $N(\Delta) = \delta$.

6. On sait que l'application $X \mapsto \|AX\|$ est continue sur $\mathbb{R}^n$ car polynomiale en les composantes de X . Comme S^1 est un compact (fermé borné), $X \mapsto \|AX\|$ est bornée et atteint ces bornes sur S^1 . En particulier, il existe $X_0 \in S^1$ tel que pour tout $X \in S^1$, $\|AX\| \leq \|AX_0\|$. Ainsi par définition de $N(A)$, $N(A) = \sup_{X \in S^1} \|AX\| = \|AX_0\| = \max_{S^1} \|AX\|$.
7. Si $X = 0$, OK. Sinon, on pose $\tilde{X} = X/\|X\|$. Puisque $\tilde{X} \in S^1$, on a par définition (et d'après la question précédente), $\|A\tilde{X}\| \leq N(A)$. Puis on applique l'homogénéité de $\|\cdot\|$ et on fait passer le $\|X\|$ de l'autre coté et on obtient le résultat.

8. Soit $A, B \in \mathcal{M}_n(\mathbb{R})$, soit $X \in \mathbb{R}^n$ différent de 0. D'après Q7, on a :

$$\|A(BX)\| \leq N(A)\|BX\| \leq N(A)N(B)\|X\|$$

puis en divisant par $\|X\|$, on obtient :

$$\|AB\tilde{X}\| \leq N(A)N(B)$$

vraie pour tout $X \in \mathbb{R}^n$ donc sur S^1 (car lorsque X parcourt $\mathbb{R}^n$, $\tilde{X}$ parcourt S^1). Puis en passant au sup sur S^1 , on obtient le résultat.

9. Il suffit d'appliquer Q7 à $X = e_i$ pour $i \in [1, n]$ puis de passer au max. En effet $C_i = Ae_i$.
10. On a pour $X = (x_1, x_2, x_3)^T \in S$: $AX = x_3(0, -1, 1)^T$. Ainsi $\|AX\| = |x_3|\sqrt{2}$. Or $|x_3| \leq 1$ donc $N(A) \leq \sqrt{2}$. De plus, $\|Ae_3\| = \sqrt{2}$ donc $N(A) = \sqrt{2}$ car $N(A)$ correspond au max des $\|AX\|$ sur S^1 .

II. Cas particulier des matrices orthogonales et symétriques

11. On a $\|UX\|^2 = (UX)^T UX = X^T U^T UX = X^T X = \|X\|^2 = 1$ sur S^1 (car U est orthogonale : $U^T U = I_n$). Ainsi pour tout X , $\|UX\| = 1$ donc son sup vaut 1 et $N(U) = 1$.
12. On applique le même raisonnement : pour tout $X \in S^1$, on a $\|UAX\|^2 = (UAX)^T UAX = (AX)^T U^T UAX = (AX)^T AX = \|AX\|^2$ donc leurs sup sont égaux sur S^1 d'où $N(UA) = N(A)$.
13. Attention au sens, $AU \neq UA$. Puisque U est une matrice orthogonale, c'est une isométrie bijective (car inversible). Ainsi lorsque X parcourt S^1 , $Y = UX$ parcourt S^1 également donc :

$$\sup_{X \in S^1} \|A(UX)\| = \sup_{Y \in S^1} \|AY\|$$

d'où $N(AU) = N(A)$.

14. Si A est symétrique réelle, d'après le théorème spectral il existe $P \in O_n(\mathbb{R})$ et $D \in D_n(\mathbb{R})$ tel que $A = PDP^T$. En appliquant Q12, on a $N(A) = N(PDP^T) = N(P(DP^T)) = N(DP^T)$ puis en appliquant Q13, $N(DP^T) = N(D)$. Ainsi d'après Q5, $N(A) = N(D) = \max_i |\lambda_i| = \rho(A)$. (on appelle λ_i les éléments diagonaux de D soit les valeurs propres de A).
15. Soit on calcule χ_A soit on remarque que 4 est valeur propre associée à $(1, 1, 1)^T$ et 1 est valeur propre double associée à $(1, -1, 0)^T$ et $(1, 0, -1)^T$. Donc d'après la question précédente, $N(A) = \rho(A) = 4$.

Partie B. Conditionnement d'une matrice pour la norme N

I. Quelques résultats sur le conditionnement

16. On applique Q8.
17. Si $\alpha \in \mathbb{R}^*$, $(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$. Puis en appliquant l'homogénéité de N , on obtient $\text{cond}(\alpha A) = \text{cond}(A)$.
18. On a vu Q11 que $N(U) = 1$ pour $U \in O_n(\mathbb{R})$ donc $\text{cond}(U) = 1$. (On reprend le même raisonnement que Q11 pour montrer que $N(U^T) = 1$ ou dit juste que U^T est une isométrie donc préserve la norme de la même manière que U).

19. On a $\text{cond}(UA) = N(UA)N((UA)^{-1})$. En appliquant Q12, $N(UA) = N(A)$ et $N((UA)^{-1}) = N(A^{-1}U^{-1}) = N(A^{-1}U^T) = N(A^{-1})$ en appliquant Q13. Ainsi $\text{cond}(UA) = \text{cond}(A)$. De la même manière, $\text{cond}(AU) = \text{cond}(A)$.

II. Un exemple de minoration du conditionnement d'une matrice

20. On a pour $1 \leq i \leq n-1$:

$$\begin{aligned}(AX)_i &= \sum_{j=1}^n a_{ij}X_j = \sum_{j=1}^n a_{ij}(-1)^{n-j}2^{n-j} \\ &= 1(-1)^{n-i}2^{n-i} + 2(-1)^{n-(i+1)}2^{n-(i+1)} \\ &= 0\end{aligned}$$

puis pour $i = n$, $(AX)_n = 1$.

21. D'après ce qui précède $X = A^{-1}E_n$ avec $E_n \in S^1$ donc par définition, $N(A^{-1}) \geq \|X\|$ or :

$$\|X\|^2 = \sum_{k=1}^n \left((-1)^{n-k} 2^{n-k} \right)^2 = \sum_{k=1}^n 2^{2(n-k)} \geq 2^{2(n-1)}$$

donc $\|X\| \geq 2^{n-1}$ (il s'agit d'une somme de terme positif donc supérieur à chaque terme, en particulier celui pour $k = 1$). D'où le résultat.

22. On note C_j la $j^{\text{ème}}$ colonne de A . On a :

$$AE_2 = C_2 \implies \|AE_2\| = \|C_2\| = \sqrt{2^2 + 1^2} = \sqrt{5} > 2$$

On en déduit en particulier que $N(A) > 2$ par définition de $N(A)$ puis en appliquant la question précédente, on obtient le résultat.

Partie C. Conditionnement pour une matrice réelle inversible

23. On a vu Q14 que pour S symétrique réelle : $N(S) = \rho(S) = |\lambda_n| = \lambda_n$ ($S \in S_n^+(\mathbb{R})$). De plus, soit $X \in S^1$. En écrivant X dans la base $\mathcal{C}$, on a :

$$X = \sum_{i=1}^n X_i V_i$$

avec $|X_i| \leq 1$ car $\|X\| = 1$ (facile à montrer par l'absurde). Puis :

$$SX = \sum_{i=1}^n \lambda_i X_i V_i$$

et par bilinéarité du produit scalaire :

$$\langle SX, X \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i X_i X_j \underbrace{V_i^T V_j}_{=\delta_{i,j}} = \sum_{i=1}^n \lambda_i X_i^2 \leq \lambda_n \sum_{i=1}^n X_i^2 = \lambda_n \|X\|^2 = \lambda_n$$

car $X_i^2 \geq 0$ et $\lambda_i \geq 0$ d'où :

$$|\langle SX, X \rangle| \leq \lambda_n$$

pour tout $X \in S^1$ or $|\langle SV_n, V_n \rangle| = \lambda_n$ avec $\|V_n\| = 1$ donc $V_n \in S^1$. Ainsi :

$$N(S) = \lambda_n = \max_{S^1} |\langle SX, X \rangle|$$

24. On a clairement $A^T A$ qui est symétrique réelle. De plus, on a pour $X \in \mathbb{R}^n$: $\langle A^T A X, X \rangle = (A^T A X)^T X = (AX)^T (AX) = \|AX\|^2 \geq 0$ donc $A^T A$ est bien dans $S_n^+(\mathbb{R})$. Puis il vient d'après Q23 :

$$N(A^T A) = \max_{S^1} |\langle A^T A X, X \rangle| = \max_{S^1} \|AX\|^2$$

Or les $\|AX\|$ sont positif donc on peut sortir le carré et on obtient :

$$N(A^T A) = \left(\max_{S^1} \|AX\| \right)^2 = N(A)^2$$

25. D'après Q5, on a $N(A^T A) = \rho(A^T A)$ puis on applique la racine sur les quantités $0 \leq N(A^T A) = N(A)^2$.
26. On a $A^T A = P A A^T P^{-1}$ avec $P = A^{-1}$ donc $A^T A$ et $A A^T$ sont semblables. Or deux matrices semblables ont le même polynôme caractéristique donc les mêmes valeurs propres.
27. D'après Q25 :
- $$\text{cond}(A) = N(A)N(A^{-1}) = \sqrt{\rho(A^T A)}\sqrt{\rho((A A^T)^{-1})}$$
- En effet $N(A^{-1}) = \sqrt{\rho((A^{-1})^T A^{-1})} = \sqrt{\rho((A A^T)^{-1})}$. Or comme $\mu_m > 0$, $A A^T$ est bien inversible et les valeurs propres de $(A A^T)^{-1}$ sont $(1/\mu_{m_1}, \dots, 1/\mu_{m_M})$ dont le max vaut $\rho((A A^T)^{-1}) = \frac{1}{\mu_m}$. Sachant $\rho(A^T A) = \mu_M$, on en déduit le résultat.
28. Si $A \in S_n^{++}(\mathbb{R})$, alors $\text{Sp}(A) \subset \mathbb{R}_+^*$ et en prenant toujours $0 < \lambda_1 \leq \dots \leq \lambda_n$ ses valeurs propres, on a les valeurs propres de $A^T A = A^2$ qui sont les λ_i^2 . On en déduit alors :

$$\text{cond}(A) = \sqrt{\frac{\lambda_n^2}{\lambda_1^2}} = \frac{\lambda_n}{\lambda_1}$$

Partie D. Calcul explicite de conditionnement

29. T symétrique réelle donc diagonalisable dans une BON avec un spectre réel.
30. On a :

$$(TU_k)_1 = 2 \sin\left(\frac{\pi k}{n+1}\right) - \sin\left(\frac{2\pi k}{n+1}\right)$$

or $\sin(2x) = 2 \sin(x) \cos(x)$. Ainsi :

$$(TU_k)_1 = 2 \left(1 - \cos\left(\frac{\pi k}{n+1}\right)\right) \sin\left(\frac{\pi k}{n+1}\right) = 2 \left(1 - \cos\left(\frac{\pi k}{n+1}\right)\right) (U_k)_1$$

On calcule de même les $(TU_k)_n$ et pour les $(TU_k)_i$ pour $i \in [2, n-1]$, on applique $\sin(a+b) + \sin(a-b) = 2 \sin(a) \cos(b)$. Finalement on obtient que U_k est un vecteur propre associé à :

$$\lambda_k = 2 \left(1 - \cos\left(\frac{\pi k}{n+1}\right)\right)$$

31. On en déduit l'ensemble des valeurs propres qui est $\{\lambda_k\}_{k \in [1, n]}$ car les λ_k sont distincts. De plus :

$$\lambda_{k+1} - \lambda_k = 2 \left(\cos\left(\frac{k\pi}{n+1}\right) - \cos\left(\frac{(k+1)\pi}{n+1}\right) \right)$$

or $\cos(x)$ est strictement décroissante sur $[0, \pi]$ donc pour $\frac{(k+1)\pi}{n+1} > \frac{k\pi}{n+1}$, on a :

$$\cos\left(\frac{(k+1)\pi}{n+1}\right) < \cos\left(\frac{k\pi}{n+1}\right)$$

d'où $\lambda_k < \lambda_{k+1}$ avec $\lambda_1 > 0$. Puisque l'ensemble des λ_k est de cardinal n et qu'on sait que T est diagonalisable, c'est bien l'ensemble des valeurs propres de T .

32. On a question précédente que $0 < \lambda_1 < \dots < \lambda_n$ donc $T \in S_n^{++}(\mathbb{R})$ puis d'après Q28 :

$$\text{cond}(T) = \frac{\lambda_n}{\lambda_1} = \frac{1 - \cos\left(\frac{n\pi}{n+1}\right)}{1 - \cos\left(\frac{\pi}{n+1}\right)}$$

or $\cos\left(\frac{n\pi}{n+1}\right) = \cos\left(\frac{(n+1)\pi}{n+1} - \frac{\pi}{n+1}\right) = -\cos\left(\frac{\pi}{n+1}\right)$ d'où :

$$\text{cond}(T) = \frac{1 + \cos\left(\frac{\pi}{n+1}\right)}{1 - \cos\left(\frac{\pi}{n+1}\right)}$$

Je pense que ce résultat est accepté mais on peut développer un petit peu encore sachant $1 + \cos(x) = 2 \cos^2(x/2)$ et $1 - \cos(\theta) = 2 \sin^2(x/2)$:

$$\text{cond}(T) = \frac{1}{\tan^2\left(\frac{\pi}{2(n+1)}\right)}$$

(les deux dernières formules trigo ne sont pas usuelles et pour les personnes qui ne les connaîtraient pas par cœur, cela ferait perdre un certain temps pour les retrouver si on est un peu maladroit dans les calculs, surtout pour obtenir une formule qui n'a pas beaucoup plus d'intérêt. Donc je pense que la première formule est suffisante).

Partie E. Inégalité de Kantorovich

I. Une première démonstration

33. Un peu nulle comme question parce que c'est le résultat de Q28 donc si on n'a pas réussi Q28 on perd automatiquement une question et on est mal parti pour cette partie.

$$\text{cond}(A) = \frac{\lambda_n}{\lambda_1}$$

34. Soit $X \in \mathbb{R}^n$. On note $\langle U, V \rangle_A = \langle AU, V \rangle$ qui est bien défini comme un produit scalaire sur $\mathbb{R}^n \times \mathbb{R}^n$. On a alors d'après C-S :

$$\langle U, V \rangle_A^2 \leq \langle U, U \rangle_A \langle V, V \rangle_A$$

En prenant $U = A^{-1}X$ et $V = X$, on a :

$$\langle X, X \rangle^2 = \|X\|^4 \leq \langle X, A^{-1}X \rangle \langle AX, X \rangle$$

puis on conclut par symétrie du produit scalaire.

35. Si on calcule les racines de P , on trouve λ_1 et λ_n . Ainsi P est du signe de son coefficient dominant (ici positif) sauf entre ses racines donc P est négatif sur $[\lambda_1, \lambda_n]$. Or puisque pour tout k , $\lambda_k \in [\lambda_1, \lambda_n]$, on en déduit le résultat.

Autre justification : $P(X) = (X - \lambda_1)(X - \lambda_k)$ donc $P(\lambda_k) = (\lambda_k - \lambda_1)(\lambda_k - \lambda_n)$ or $\lambda_k - \lambda_1 \geq 0$ et $\lambda_k - \lambda_n \leq 0$ pour tout k d'où le résultat.

36. A est diagonalisable : $A = PDP^T$ et $A^{-1} = PD^{-1}P^T$ où $D^{-1} = \text{Diag}(1/\lambda_1, \dots, 1/\lambda_n)$ ($\lambda_i > 0$). Puis :

$$P(A) = P(D^2 - (\lambda_1 + \lambda_2)D + \lambda_1\lambda_n I_n)P^T$$

d'où :

$$B = PD^{-1}P(D)P^T$$

avec $D^{-1}P(D)$ diagonale donc B est diagonalisable. De plus ses valeurs propres sont $\mu_k = \frac{P(\lambda_k)}{\lambda_k} \leq 0$. Ainsi $\text{Sp}(B) \subset \mathbb{R}_-$ donc $\forall X \in \mathbb{R}^n$, $\langle BX, X \rangle \leq 0$. (pour le justifier on prend $\tilde{B} = -B$ donc le spectre est ≥ 0 et on applique le cours puis on multiplie par -1 le tout).

37. On a :

$$\begin{aligned} f(1) &= \langle AX, X \rangle + \langle -(\lambda_1 + \lambda_n)X, X \rangle + \langle \lambda_1\lambda_n A^{-1}X, X \rangle \\ &= \langle (A - (\lambda_1 + \lambda_n)I_n + \lambda_1\lambda_n A^{-1})X, X \rangle \\ &= \langle BX, X \rangle \\ &\leq 0 \end{aligned}$$

d'après la question précédente. Puis $f(0) = \lambda_1\lambda_n \langle A^{-1}X, X \rangle > 0$ car $\lambda_1\lambda_n > 0$ et $\langle A^{-1}X, X \rangle \geq 0$ pour tout X car $\text{Sp} A^{-1} \in \mathbb{R}_+$. D'où $f(0)f(1) \leq 0$.

Notons que le discriminant du polynôme $f(\lambda)$ est :

$$\Delta = (\lambda_1 + \lambda_n)^2 \|X\|^4 - 4\lambda_1\lambda_n \langle AX, X \rangle \langle A^{-1}X, X \rangle$$

Puisque $f(0) > 0$, on a deux cas possible :

- Soit $f(1) = 0$.
- Soit $f(1) < 0$ et donc f admet une racine sur $[0, 1]$ d'après le TVI (car f est continue en tant que fonction polynomiale).

Dans les deux cas, f admet une racine réelle donc finalement son discriminant est ≥ 0 d'où :

$$(\lambda_1 + \lambda_n)^2 \|X\|^4 - 4\lambda_1\lambda_n \langle AX, X \rangle \langle A^{-1}X, X \rangle \geq 0$$

38. On a vu Q31 la première inégalité : $\|X\|^4 \leq (AX, X)(A^{-1}X, X)$. D'après la question précédente, on a :

$$(AX, X)(A^{-1}X, X) \leq \frac{1}{4} \frac{(\lambda_1 + \lambda_n)^2}{\lambda_1 \lambda_n} \|X\|^4$$

or :

$$\frac{(\lambda_1 + \lambda_n)^2}{\lambda_1 \lambda_n} = \left(\frac{\lambda_1 + \lambda_n}{\sqrt{\lambda_1} \sqrt{\lambda_n}} \right)^2 = \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^2 = \left(\frac{1}{\sqrt{\text{cond}(A)}} + \sqrt{\text{cond}(A)} \right)^2$$

d'où le résultat.

II. Une deuxième démonstration

39. Puisque X est unitaire, on a bien $0 \leq x_i^2 \leq 1$ donc $P(Z = \lambda_i) \in [0, 1]$. De plus, $Z(\Omega)$ est dénombrable et :

$$P(Z(\Omega)) = P\left(\bigcup_{i=1}^n (Z = \lambda_i)\right) = \sum_{i=1}^n P(Z = \lambda_i) = \sum_{i=1}^n x_i^2 = 1$$

donc on définit bien une loi de probabilité.

40. $Z(\Omega)$ est fini donc $E(Z)$ est une somme finie qui est bien définie. Sinon, on peut remarquer que :

$$E(Z) = \sum_{i=1}^n \lambda_i P(Z = \lambda_i) = \sum_{i=1}^n \lambda_i x_i^2 = (AX, X) < \infty$$

On a de même $1/Z$ est d'espérance finie car $1/Z(\Omega)$ est finie, puis par le théorème de transfert :

$$E(1/Z) = \sum_{i=1}^n \frac{1}{\lambda_i} P(Z = \lambda_i) = \sum_{i=1}^n \frac{1}{\lambda_i} x_i^2 = (A^{-1}X, X) < \infty$$

bien défini car $\lambda_i > 0$ pour tout i puisque $A \in S_n^+(\mathbb{R})$.

41. Il suffit de dire que Z ne prend comme valeur que les λ_i et pour λ_i , on a :

$$(\lambda_i - \lambda_1)(\lambda_i - \lambda_n) = \lambda_i^2 - (\lambda_1 + \lambda_n)\lambda_i + \lambda_1 \lambda_n \leq 0$$

d'où :

$$\lambda_1 \lambda_n \leq (\lambda_1 + \lambda_n)\lambda_i - \lambda_i^2$$

puis en multipliant par $\frac{1}{\lambda_i \lambda_n \lambda_i} > 0$, on a :

$$\frac{1}{\lambda_i} \leq \frac{\lambda_1 + \lambda_n - \lambda_i}{\lambda_1 \lambda_n}$$

ceci étant vrai pour tout i , on en déduit l'inégalité pour Z .

42. D'après ce qui précède et puisque l'inégalité précédente est vraie presque sûrement, on applique la croissances de l'espérance et :

$$E(Z)E\left(\frac{1}{Z}\right) \leq \frac{(\lambda_1 + \lambda_n)E(Z) - E(Z)^2}{\lambda_1 \lambda_n}$$

(car $E(Z) \geq 0$). On vérifie en développant le membre de droite fourni par l'énoncé que les deux expressions sont bien égales.

43. Puisque $\frac{1}{\lambda_1 \lambda_n} (E(Z) - \frac{\lambda_1 + \lambda_n}{2})^2 \geq 0$, on a :

$$E(Z)E\left(\frac{1}{Z}\right) = -\frac{1}{4\lambda_1 \lambda_n} (E(Z) - \frac{\lambda_1 + \lambda_n}{2})^2 + \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n} \underbrace{\|X\|^4}_{=1}$$

En remplaçant $E(Z) = (AX, X)$ et $E(1/Z) = (A^{-1}X, X)$, on est ramené à la même conclusion qu'à la question 38.