Devoir sur table n° 11

(type Centrale, corrigé)

Table des matières

1 Commentaires

Ce sujet est une adaptation de l'épreuve de Mathématiques II du Concours Centrale-Supélec, filière MP, année 2013. J'ai enlevé quelques indications en début de sujet pour les « questions de cours ».

La première partie est ultra-classique et en dehors de la dernière question, difficile, vous devez d'ici les écrits en maîtriser tout ce qu'on y fait (existence d'une racine carrée symétrique positive, décomposition OS, compacité de $\mathrm{O}_n(\mathbb{R})$...).

La troisième partie donne (encore) un exemple de matrice tridiagonale. Vous en aviez déjà croisé une, très générale, dans le DST n° 10. Vous y aviez démontré que ces matrices étaient toujours diagonalisables (du moins dans le cas symétrique), avec n valeurs propres distinctes et des sous-espaces propres de dimension 1. Ici, en plus de retrouver ce résultat dans un cas particulier, vous explicitez les sous-espaces propres. Il s'avère qu'ils ont exactement la même forme même lorsqu'on remplace 2 et -1 par des réels α et β ! Le moyen le plus économe de le démontrer est de traiter le cas particulier :

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

et de s'y ramener ensuite en notant que pour toute matrice B tridiagonale, l'équation $BX = \lambda X$ équivaut à une équation aux éléments propres impliquant la matrice A. Ainsi les valeurs propres et vecteurs propres de B se déduisent de ceux de A. Exercice.

La quatrième partie permet de comprendre une partie de l'intérêt des décompositions OT et OS : elles permettent de ramener toute matrice A à une matrice orthogonale, et en particulier de mesurer sa distance à $\mathrm{O}_n(\mathbb{R})$. C'est implicitement le genre de raisonnement à l'œuvre lorsqu'on calcule le maximum de $f : M \mapsto \mathrm{tr}(AM)$ sur $\mathrm{O}_n(\mathbb{R})$ en écrivant $A = OS$ avec O orthogonale.

En plus de couvrir une bonne partie du chapitre de géométrie, ce sujet a l'avantage de vous faire travailler la trigonométrie (deux dernières parties) et de vous forcer à une compréhension pratique des outils mobilisés (ce qui est peut-être éclairant) : calcul explicite d'une matrice de passage orthogonale pour réduire une matrice symétrique, puis d'une racine carrée de matrice, observation que $\mathrm{tr}(AM)$ est un produit scalaire très banal (somme des produits des coefficients), etc.

Ce qu'on retiendra en bref. Existence et unicité d'une racine carrée de matrice symétrique positive. Interpolation de Lagrange pour montrer qu'un endomorphisme est un polynôme en un autre, pour en déduire des relations de commutation. Projecteurs spectraux. Utilisation de la compacité (ici celle de $\mathrm{O}_n(\mathbb{R})$) pour faire converger « de force » des suites. Extension d'un résultat par densité. Utilisation de la caractérisation séquentielle de la continuité (ou par les images réciproques de fermés) pour montrer la continuité d'une bijection réciproque. Des matrices réelles semblables sur $\mathbb{C}$ sont semblables sur $\mathbb{R}$. Réduction des matrices tridiagonales. Théorème de représentation de Riesz. La matrice $A^\top A$ est symétrique positive. Inverse de $\sum_{k=0}^{n-1} N^k$ quand N est nilpotente.

Questions faciles ou classiques à retravailler

- PREMIÈRE PARTIE : questions 1 à 4, questions 6 à 7 ;
- DEUXIÈME PARTIE : questions ;
- TROISIÈME PARTIE : toutes les questions ;
- QUATRIÈME PARTIE : questions 23 à 25.

1.1 Rapport officiel de l'épreuve

1.2 Présentation du sujet

Le problème porte sur la décomposition polaire et diverses applications.

La partie I est consacrée à la décomposition polaire. D'abord l'existence et l'unicité de la décomposition polaire d'un élément de $\mathrm{GL}_n(\mathbb{R})$ sont établies. Par des arguments topologiques, on montre ensuite l'existence d'une telle décomposition pour une matrice réelle.

La deuxième partie traite deux applications. Dans la première, on montre que deux matrices réelles unitairement semblables sont orthogonalement semblables. Dans la seconde, il faut utiliser la décomposition polaire pour résoudre un système matriciel.

Dans la troisième partie, les valeurs propres et les vecteurs propres d'une matrice tridiagonale donnée sont étudiées (intervenant naturellement en analyse numérique).

Enfin, dans la quatrième partie, il faut déterminer le maximum de l'application qui à M associe la trace de AM sur le groupe orthogonal.

1.3 Analyse globale des résultats

Ce texte, de longueur raisonnable, permet de couvrir largement le programme : matrice symétriques réelles et théorème spectral, algèbre linéaire et réduction, topologie dans les espaces de matrices, analyse asymptotique.

La première partie est assez facile et classique. Les correcteurs se sont donc particulièrement attachés à la qualité de la rédaction. Les trois parties suivantes ont été abordées par la majorité des candidats.

Le sujet a permis un bon étalement des notes. En valorisant le travail approfondi du cours et la maîtrise d'objets et de notions fondamentaux du programme de mathématiques, il a permis d'évaluer de manière satisfaisante les diverses qualités des candidats.

1.4 Commentaires sur les réponses apportées

(...) Dans la question 2, le caractère diagonalisable car autoadjoint de l'induit est souvent omis. La question 4 est en général bien traitée. La fin de la partie I est souvent abordée. On y relève beaucoup d'arguments sommaires voire faux (caractérisations des matrices symétriques positives et des matrices orthogonales uniquement à l'aide du déterminant par exemple). La continuité de la réciproque, dans la question 8, n'est correctement traitée que par quelques candidats.

La question 9 est le plus souvent bien traitée. L'argument d'inversibilité faisant intervenir le déterminant est souvent vu. Cependant, la justification est parfois incomplète (caractère non nul du polynôme). La suite est en général correcte. Le début des questions 16 à 18 fait appel aux mêmes arguments que ceux de la partie I et est souvent bien traité. En revanche, la question 18, délicate, n'est que rarement résolue.

La relation de récurrence de la question 19 est établie dans une grande majorité de copies. En revanche, l'obtention du terme général de la suite récurrente et la recherche des vecteurs propres de la matrice A_p sont très discriminants.

Dans la question 23, le lien avec la représentation des formes linéaires est perçu de manière satisfaisante par d'assez nombreux candidats. D'autres exhibent une telle matrice en regardant l'action de f sur la base canonique. Les questions 24 à 27, 28 et 29, sont bien traitées quand elles sont abordées. En revanche, les questions 30 à 32 n'est menée à son terme que par une poignée de candidats.

1.5 Conclusions

Une fois de plus, ce sujet montre qu'il est indispensable de maîtriser les définitions, les démonstrations et les méthodes du cours. La rédaction et la présentation d'une copie de concours doivent être irréprochables. Les copies peu lisibles et mal présentées ont été systématiquement sanctionnées.

2 Corrigé

Dans tout ce corrigé, nous noterons $\langle\cdot,\cdot\rangle$ et $\|\cdot\|$ le produit scalaire et la norme euclidienne usuels de $\mathrm{M}_{n,1}(\mathbb{R})$.

Première partie – décomposition polaire

1. Comme S est définie positive, ses valeurs propres sont strictement positives et donc 0 n'est pas valeur propre de S. Cela équivaut à l'inversibilité de S. Montrons que S^{-1} est symétrique définie positive. On a : $(S^{-1})^\top = (S^\top)^{-1} = S^{-1}$, donc S^{-1} est symétrique. De plus :
$$\mathrm{Sp}_{\mathbb{R}}\left(S^{-1}\right) = \left\{\lambda^{-1} \mid \lambda \in \mathrm{Sp}_{\mathbb{R}}(S)\right\}.$$
Une façon parmi d'autres de le justifier : si $\mu \in \mathbb{R}$ et $X \in \mathrm{M}_{n,1}(\mathbb{R})$ est non nul, alors μ doit être non nul puisque S^{-1} est inversible, et $S^{-1}X = \mu X$ si et seulement si $X = \mu S X$, si et seulement si $SX = \frac{1}{\mu}X$, si et seulement $\mu^{-1} \in \mathrm{Sp}_{\mathbb{R}}(S)$.
Comme les valeurs propres de S sont strictement positives, l'égalité ci-dessus montre que les valeurs propres de S^{-1} le sont aussi, d'où : $S^{-1} \in \mathrm{S}_n^{++}(\mathbb{R})$.

> **Questions à se poser, réflexes à acquérir.**
> — Peut-on trouver une démonstration de l'inversibilité des matrices symétriques définies positives ne nécessitant pas le recours au puissant théorème spectral ?
> — Trouver une autre façon de justifier le lien entre spectres de S et S^{-1}. Se demander quel est, en toute généralité, l'avantage de la démonstration ci-dessus par rapport à d'autres (j'en vois deux).

2. On traite principalement le cas positif, le cas défini positif étant presque en tous points analogue. La démonstration de ce résultat est dans votre cours. Pour varier les plaisirs, je vous propose une démonstration matricielle de l'existence, plus facile à rédiger, et une autre démonstration de l'unicité (mais toujours en passant par des endomorphismes : là, les matrices manquent de souplesse pour y parvenir).

   Existence par les matrices. Comme u est autoadjoint, par le théorème spectral il existe une base orthonormée $\mathcal{B}$ de $\mathbb{R}^n$ dans laquelle la matrice de u est de la forme : $D = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$.

   Comme u est de plus positif, on sait que l'on a : $\forall i \in [\![1, n]\!]$, $\lambda_i \geqslant 0$. Considérons donc l'endomorphisme v de $\mathbb{R}^n$ tel que :
$$\mathrm{M}_{\mathcal{B}}(v) = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}.$$
   L'endomorphisme v est autoadjoint car sa matrice dans la base orthonormée $\mathcal{B}$ est symétrique (elle est même diagonale), et positif parce que son spectre est $\left\{\sqrt{\lambda_i} \mid i \in [\![1, n]\!]\right\}$, qui est inclus dans $\mathbb{R}_+$. De plus on a clairement : $\mathrm{M}_{\mathcal{B}}(v)^2 = D = \mathrm{M}_{\mathcal{B}}(u)$, donc : $v^2 = u$. D'où l'existence.

   Dans le cas défini positif, ce n'est guère différent : le spectre de v est inclus dans $\mathbb{R}_+^*$ dans ce cas.

   Existence par les endomorphismes (comme dans le cours, hormis pour la justification de la positivité de v : comme d'habitude, on peut se dispenser des matrices). Comme u est autoadjoint, par le théorème spectral il est diagonalisable et on a : $u = \sum_{\lambda \in \mathrm{Sp}(u)} \lambda p_\lambda$, où les p_λ sont les projecteurs spectraux, qui sont des endomorphismes autoadjoints car les sous-espaces propres de u sont orthogonaux, de sorte que les caractéristiques géométriques des projecteurs spectraux soient supplémentaires orthogonales (on sait que pour des projecteurs, être orthogonal équivaut à être autoadjoint).

   Toutes les valeurs propres de u étant positives par hypothèse sur A, on peut poser :
$$v = \sum_{\lambda \in \mathrm{Sp}(u)} \sqrt{\lambda} p_\lambda.$$
   Alors v est autoadjoint car $\mathrm{S}(\mathbb{R}^n)$ est un espace vectoriel, et les propriétés des projecteurs spectraux impliquent :
$$v^2 = \sum_{\lambda \in \mathrm{Sp}(u)} (\sqrt{\lambda})^2 p_\lambda = \sum_{\lambda \in \mathrm{Sp}(u)} \lambda p_\lambda = u.$$
   Il reste à justifier qu'il est positif. On a pour tout $\vec{x} \in \mathbb{R}^n$:
$$\langle v(\vec{x}), \vec{x} \rangle = \sum_{\lambda \in \mathrm{Sp}(u)} \sqrt{\lambda} \langle p_\lambda(\vec{x}), \vec{x} \rangle \geqslant 0,$$
   parce que les $\sqrt{\lambda}$ sont positifs et un projecteur orthogonal est toujours positif (on le constate soit en notant que son spectre est inclus dans $\{0,1\}$ et donc dans $\mathbb{R}_+$, soit en notant que pour un projecteur orthogonal on a : $\langle p(\vec{x}), \vec{x} \rangle = \|p(\vec{x})\|^2 \geqslant 0$). Donc v est autoadjoint positif, ce qu'il fallait démontrer.

   Dans le cas défini positif, notons que si $\langle v(\vec{x}), \vec{x} \rangle = 0$ alors $\langle p_\lambda(\vec{x}), \vec{x} \rangle = 0$ pour tout $\lambda \in \mathrm{Sp}(u)$, étant donné qu'une somme de réels est nulle seulement si chaque terme est nul. On aurait alors : $\langle u(\vec{x}), \vec{x} \rangle = \sum_{\lambda \in \mathrm{Sp}(u)} \lambda \langle p_\lambda(\vec{x}), \vec{x} \rangle = 0$, donc $\vec{x} = \vec{0}$ par définie positivité de u.

   Démontrons l'unicité. Soit w un endomorphisme autoadjoint positif tel que : $w^2 = u$.

   Unicité démontrée sans codiagonaliser (du moins, pas explicitement). Nous allons expliciter w en l'explicitant sur les sous-espaces propres de u, et constater qu'on obtient ainsi

$w = v$. Comme u est un polynôme en w, les endomorphismes u et w commutent, donc les sous-espaces propres de u sont stables par w. Soient $\lambda \in \mathrm{Sp}(u)$ et $E_\lambda = \ker(u - \lambda \mathrm{Id}_{\mathbb{R}^n})$. Comme w est autoadjoint, par le théorème spectral w est diagonalisable, et donc l'endomorphisme induit par w sur E_λ, que l'on note w_λ dans ce qui suit, est aussi diagonalisable. Nous allons montrer que w_λ est une homothétie en montrant qu'il admet une unique valeur propre : soient $\mu \in \mathbb{R}$ une valeur propre de w_λ et $\vec{x} \in E_\lambda$ un vecteur propre associé. On a : $w_\lambda(\vec{x}) = \mu \vec{x}$, donc : $w_\lambda^2(\vec{x}) = \mu^2 \vec{x}$; mais on a aussi, comme $w^2 = u$ et $\vec{x} \in E_\lambda$:

$$w_\lambda^2(\vec{x}) = w^2(\vec{x}) = u(\vec{x}) = \lambda \vec{x}.$$

Comme $\vec{x}$ est non nul, ces deux égalités imposent : $\mu^2 = \lambda$, et la positivité de μ implique : $\mu = \sqrt{\lambda}$. Ainsi l'unique valeur propre possible de w_λ est $\sqrt{\lambda}$, et comme w_λ est diagonalisable cela veut dire : $w_\lambda = \sqrt{\lambda}\mathrm{Id}_{E_\lambda}$. On conclut grâce à la somme directe $E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda$:

$$w = \sum_{\lambda \in \mathrm{Sp}(u)} w_\lambda \circ p_\lambda = \sum_{\lambda \in \mathrm{Sp}(u)} \sqrt{\lambda} p_\lambda = v.$$

D'où le résultat.

Unicité démontrée comme dans le cours. On veut montrer : $w = v$. Pour cela, notons que w et v sont diagonalisables en tant qu'endomorphismes autoadjoints et ils commutent (en effet w commute avec u et u est un polynôme en v, comme on le montre par interpolation de Lagrange ou par le lemme des noyaux : voir le détail dans la question suivante), donc ils sont codiagonalisables : soit $(\vec{e}_1, \ldots, \vec{e}_n)$ une base commune de vecteurs propres de u et v, et pour tout $i \in [\![1, n]\!]$, soit $(\lambda_i, \mu_i) \in (\mathbb{R}_+)^2$ tel que : $v(\vec{e}_i) = \lambda_i \vec{e}_i$, $w(\vec{e}_i) = \mu_i \vec{e}_i$. On a : $v^2 = w^2 = u$, donc : $\lambda_i^2 = \mu_i^2$. Comme λ_i et μ_i sont positifs : $\lambda_i = \mu_i$. Ceci vaut pour tout $i \in [\![1, n]\!]$, donc v et w coïncident sur une base, ce qui montre : $v = w$, comme attendu.

Tout ceci montre aussi que l'endomorphisme induit par v sur $\ker(u - \lambda \mathrm{Id}_{\mathbb{R}^n})$ est $\sqrt{\lambda}\mathrm{Id}_{\ker(u - \lambda \mathrm{Id}_{\mathbb{R}^n})}$.

Questions à se poser, réflexes à acquérir. Il y a beaucoup de choses à méditer.

— Bien noter que je n'introduis pas la matrice de u relativement à la base canonique de $\mathbb{R}^n$, écrivant directement la matrice dans une base orthonormée de vecteurs propres : allusion à l'obsession de certains, parmi vous, qui veulent toujours se ramener à la base canonique (alors que la philosophie de la réduction devrait vous habituer au fait que ce ne soit qu'une base parmi d'autres, et rarement la plus pertinente), qui vous aurait ici fait introduire une matrice de passage P inutile, rendant la vérification de la positivité de v moins directe.

— Est-ce que les deux démonstrations de l'existence sont « les mêmes », simplement formulées différemment ? Pour répondre à cette question, il faut bien comprendre ce qu'est un projecteur spectral (indice : rien de bien fou). Vous pouvez par exemple raisonner à la place en écrivant : $\vec{x} = \sum_\lambda \vec{x}_\lambda$, avec les $\vec{x}_\lambda$ dans les sous-espaces propres de u, et calculer $v(\vec{x})$, $v^2(\vec{x})$, etc., avec cette expression. Faire le lien avec ce que je fais avec des projecteurs spectraux.

— Se convaincre de l'identité $v^2 = \sum_\lambda (\sqrt{\lambda})^2 p_\lambda$. Plus généralement, a-t-on $P(v) = \sum_\lambda P(\sqrt{\lambda}) p_\lambda$? Est-ce que cela utilise l'aspect orthogonal des p_λ ? Éventuellement réviser le chapitre V.

— Pourquoi $\langle p(\vec{x}), \vec{x} \rangle = \|p(\vec{x})\|^2$ est-elle vraie ? Est-ce propre aux projecteurs orthogonaux ? (Un dessin permet de comprendre comment construire un contre-exemple.)

— Dans le cas matriciel, on parvient à montrer le caractère défini positif de v seulement grâce au spectre, sans revenir à u. Ce n'est pas le cas dans la seconde démonstration que je propose (avec les endomorphismes). Pouvait-on effectivement montrer que $\langle v(\vec{x}), \vec{x} \rangle = 0$ implique $\vec{x} = \vec{0}$ sans utiliser u, mais seulement la stricte positivité des $\sqrt{\lambda}$?

— Si l'identité $w = \sum_\lambda w_\lambda \circ p_\lambda$ vous paraît obscure : la rendre plus concrète en révisant dans le chapitre V comment on peut définir une application linéaire à l'aide de ses restrictions sur des sous-espaces supplémentaires. Si les projecteurs vous ennuient encore, écrire $\vec{x} = \sum_\lambda \vec{x}_\lambda$ et calculer $w(\vec{x})$.

— Revoir si besoin pourquoi le fait d'avoir une unique valeur propre n'implique pas forcément qu'on est une homothétie, hormis si l'endomorphisme est diagonalisable.

— Revoir comment on codiagonalise des endomorphismes diagonalisables et qui commutent.

3. **Par les matrices.** Soit $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ dans laquelle la matrice de u est de la forme : $D = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$. Par la question précédente, on a : $\mathrm{M}_{\mathcal{B}}(v) = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$.
Notons D' cette matrice. Nous allons construire un polynôme Q tel que $D' = Q(D)$ par interpolation de Lagrange : la fonction racine carrée étant injective sur $\mathbb{R}_+$, les $\sqrt{\lambda}$ sont tous distincts quand λ parcourt le spectre de u. On en déduit qu'il existe un polynôme $Q \in \mathbb{R}[X]$ tel que : $\forall \lambda \in \mathrm{Sp}(u)$, $Q(\lambda) = \sqrt{\lambda}$. On a alors :
$$\mathrm{M}_{\mathcal{B}}(v) = \begin{pmatrix} Q(\lambda_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & Q(\lambda_n) \end{pmatrix} = Q\left(\begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}\right) = Q(D) = Q(\mathrm{M}_{\mathcal{B}}(u)),$$
d'où : $v = Q(u)$.
Par les endomorphismes. D'après la question précédente, v est une combinaison linéaire des projecteurs liés à la décomposition en somme directe $E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} \ker(u - \lambda \mathrm{Id}_E)$. Par le lemme des noyaux, ces projecteurs sont des polynômes en u, donc v est un polynôme en u : d'où le résultat.

> **Questions à se poser, réflexes à acquérir.**
> — Noter comment j'ai rédigé pour éviter le point épineux que certains $\sqrt{\lambda_i}$ peuvent être égaux (s'il y a des valeurs propres multiples) alors que l'interpolation de Lagrange doit être en des points distincts.
> — Est-ce que ces deux démonstrations fournissent le même polynôme Q ? Rappelons que nous avons explicité au chapitre V les polynômes décrivant les projecteurs spectraux (dans le cas diagonalisable).

4. Nous allons montrer l'existence et unicité du couple (O, S) par analyse-synthèse.
Analyse. Supposons qu'il existe $(O, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ tel que : $A = OS$. Alors on a : $A^\top = S^\top O^\top = SO^\top$, et donc :
$$A^\top A = SO^\top OS = S^2.$$
On observe donc que si S convient, alors S est l'unique matrice symétrique définie positive telle que : $S^2 = A^\top A$ (nous justifions que $A^\top A$ est bien symétrique définie positive dans la synthèse ci-dessous). La matrice S étant déterminée, on obtient O en posant : $O = AS^{-1}$, la matrice S étant inversible par la question 1.
Synthèse. Montrons que $A^\top A$ est dans $\mathrm{S}_n^{++}(\mathbb{R})$, de sorte à pouvoir utiliser la question 2 pour définir S. On a : $(A^\top A)^\top = A^\top \left(A^\top\right)^\top = A^\top A$, donc $A^\top A$ est symétrique, et elle est définie positive puisque :
$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \left\{0_{\mathrm{M}_{n,1}(\mathbb{R})}\right\}, \quad X^\top A^\top A X = (AX)^\top AX = \|AX\|^2 > 0$$
(comme A est inversible et X non nul, on a $AX \neq 0_{\mathrm{M}_{n,1}(\mathbb{R})}$). Par la question 2, il existe donc une matrice (unique) $S \in \mathrm{S}_n^{++}(\mathbb{R})$ telle que : $S^2 = A^\top A$. Elle est inversible par la question 1, ce qui permet de poser : $O = AS^{-1}$. On a alors, évidemment : $A = OS$, et si l'on montre que O est orthogonale l'existence du couple (O, S) vérifiant les hypothèses de l'énoncé est prouvée.
On a :
$$O^\top O = \left(AS^{-1}\right)^\top \left(AS^{-1}\right) = \left(S^{-1}\right)^\top A^\top A S^{-1} = S^{-1}\left(A^\top A\right) S^{-1} = S^{-1} S^2 S^{-1} = \mathrm{I}_n,$$
donc O est orthogonale.
Ainsi la synthèse montre l'existence de $(O, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ tel que : $A = OS$, et l'analyse montre l'unicité.

Questions à se poser, réflexes à acquérir.

— Si l'on n'avait pas démontré l'unicité de la racine carrée symétrique définie positive, comment aurait-on pu malgré tout répondre à cette question ?

— Est-ce que O et S sont des polynômes en A ?

5. On a : $A^\top A = \begin{pmatrix} 10 & 0 & -6 \\ 0 & 36 & 0 \\ -6 & 0 & 10 \end{pmatrix}$. En développant par rapport à la seconde colonne, on a aisément : $\chi_{A^\top A} = (X-36)\left((X-10)^2 - 6^2\right) = (X-36)(X-16)(X-4)$. L'ensemble des valeurs propres de $A^\top A$ est donc $\{4{,}16{,}36\}$. Ses valeurs propres étant simples, les sous-espaces propres associés sont des droites et donc un vecteur propre suffit à engendrer chacun d'entre eux. Or $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ est clairement vecteur propre de $A^\top A$ associé à 36 (par simple inspection de la seconde colonne) et la somme des coefficients des première et troisième ligne égale 4, ce qui permet de montrer aisément que $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ est un vecteur propre associé à 4. Donc :

$$\ker\left(A^\top A - 4\mathrm{I}_3\right) = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right), \quad \ker\left(A^\top A - 36\mathrm{I}_3\right) = \mathrm{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right).$$

Pour trouver le sous-espace propre associé à 16, on peut le voir à l'œil nu par une inspection identique à celle faite pour la valeur propre 4 ; si ce n'est pas le cas, on peut le trouver facilement en se souvenant que pour une matrice symétrique, les sous-espaces propres sont deux à deux orthogonaux (et leur somme donne l'espace entier par le théorème spectral). Il suffit donc de trouver un vecteur non nul orthogonal aux deux trouvés ci-dessus pour qu'il engendre $\ker\left(A^\top A - 16\mathrm{I}_3\right)$. On trouve ainsi :

$$\ker\left(A^\top A - 16\mathrm{I}_3\right) = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right).$$

On en déduit une base orthonormée de vecteurs propres en rendant unitaires les vecteurs ci-dessus (on a déjà l'orthogonalité par l'argument ci-dessus). En appliquant la formule de changement de base à l'endomorphisme $X \mapsto AX$ entre la base canonique et la base :

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right),$$

on obtient :

$$A^\top A = P\begin{pmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 36 \end{pmatrix}P^{-1}, \quad \text{où :} \quad P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}.$$

On prend alors, suivant la question précédente et la construction explicite des racines carrées dans la question 2 :

$$S = P\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}P^\top = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 6 & 0 \\ -1 & 0 & 3 \end{pmatrix},$$

puis :

$$O = AS^{-1} = \begin{pmatrix} 3 & 0 & -1 \\ 1/\sqrt{2} & 3\sqrt{2} & -3/\sqrt{2} \\ -1/\sqrt{2} & 3\sqrt{2} & 3/\sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{3}{8} & 0 & \frac{1}{8} \\ 0 & \frac{1}{6} & 0 \\ \frac{1}{8} & 0 & \frac{3}{8} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Questions à se poser, réflexes à acquérir.

- Bien noter que S est symétrique positive ; pourtant ses coefficients ne sont pas tous positifs. Ne pas éviter d'interpréter n'importe comment la notion de positivité.
- Observer comment j'ai produit une matrice de passage orthogonale. Ici tout se passe très bien car il y a trois sous-espaces propres, orthogonaux entre eux, donc concaténer les bases des sous-espaces propres donne automatiquement une famille orthogonale. Mais si l'on n'avait eu que deux sous-espaces propres (toujours pour une matrice d'ordre 3), il aurait éventuellement fallu faire quelque chose avant de concaténer, pour avoir une famille orthogonale : quoi donc ?

6. L'application $A \mapsto (A^\top, A)$ est linéaire et $\varphi_2 : (A, B) \mapsto AB$ bilinéaire sur des espaces vectoriels de dimension finie, donc elles sont continues. Par conséquent leur composition $f : A \mapsto A^\top A$ est continue sur $\mathrm{M}_n(\mathbb{R})$, donc $\mathrm{O}_n(\mathbb{R}) = f^{-1}\{\mathrm{I}_n\}$ est une partie fermée de $\mathrm{M}_n(\mathbb{R})$ en tant qu'image réciproque du fermé $\{\mathrm{I}_n\}$ par une application continue.
   Montrons qu'elle est également bornée. On munit $\mathrm{M}_n(\mathbb{R})$ de la norme euclidienne usuelle. On a : $\forall A \in \mathrm{O}_n(\mathbb{R})$, $\|A\| = \sqrt{\mathrm{tr}\,(A^\top A)} = \sqrt{\mathrm{tr}(\mathrm{I}_n)} = \sqrt{n}$, donc $\mathrm{O}_n(\mathbb{R})$ est une partie bornée de $\mathrm{M}_n(\mathbb{R})$.
   En tant que partie fermée et bornée d'un espace vectoriel de dimension finie, $\mathrm{O}_n(\mathbb{R})$ est compact : d'où le résultat.

Questions à se poser, réflexes à acquérir. Qu'obtient-on comme majorant pour les autres normes habituelles sur $\mathrm{M}_n(\mathbb{R})$? (Normes subordonnées, norme 1, norme $\|\cdot\|_\infty$...).

7. On raisonne par densité. Posons : $\forall p \in \mathbb{N} \setminus \{0\}$, $A_p = A - \frac{1}{p}\mathrm{I}_n$. On a clairement :
$$\lim_{p\to+\infty} A_p = A.$$
   Justifions que $(A_p)_{p\geqslant 1}$ est à valeurs dans $\mathrm{GL}_n(\mathbb{R})$ au voisinage de $+\infty$. Comme A admet un nombre fini de valeurs propres, il existe $N \in \mathbb{N} \setminus \{0\}$ tel que pour tout entier $p \geqslant N$, on ait : $\frac{1}{p} \notin \mathrm{Sp}_{\mathbb{R}}(A)$ (prendre N tel que : $\frac{1}{N} < \min\left(\mathrm{Sp}_{\mathbb{R}}(A) \cap \mathbb{R}_+^*\right)$). Par conséquent A_p est inversible pour tout entier $p \geqslant N$. Par commodité rédactionnelle, indexons la suite $(A_p)_{p\geqslant N}$ à partir de $N = 0$ (cela revient simplement à composer cette suite par $p \mapsto p + N$).
   Par la question 4, pour tout entier naturel p il existe un couple $(O_p, S_p) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^+(\mathbb{R})$ tel que : $A_p = O_p S_p$. On aimerait en déduire que A est égale au produit des limites des suites $(O_p)_{p\in\mathbb{N}}$ et $(S_p)_{p\in\mathbb{N}}$; le problème est qu'elles n'ont pas de raison de converger. Heureusement on sait que $\mathrm{O}_n(\mathbb{R})$ est compact par la question 6, donc il existe une suite extraite $(O_{\varphi(p)})_{p\in\mathbb{N}}$ qui converge vers une matrice orthogonale $O \in \mathrm{O}_n(\mathbb{R})$. On en déduit que $\left(S_{\varphi(p)}\right)_{p\in\mathbb{N}}$ converge également, puisque :
$$\forall p \in \mathbb{N}, \quad S_{\varphi(p)} = O_{\varphi(p)}^{-1} A_{\varphi(p)} = O_{\varphi(p)}^\top A_{\varphi(p)} \underset{p\to+\infty}{\longrightarrow} O^\top A$$
   par continuité de la transposition (c'est une application linéaire définie sur un espace vectoriel de dimension finie) et du produit matriciel. De plus la limite $S = O^\top A$ est symétrique positive parce que $\mathrm{S}_n^+(\mathbb{R})$ est un fermé de $\mathrm{M}_n(\mathbb{R})$ en tant qu'intersection de fermés :
$$\mathrm{S}_n^+(\mathbb{R}) = \mathrm{S}_n(\mathbb{R}) \cap \bigcap_{X\in\mathrm{M}_{n,1}(\mathbb{R})} f_X^{-1}(\mathbb{R}_+),$$
   où, pour tout vecteur colonne X, l'application f_X est définie sur $\mathrm{M}_n(\mathbb{R})$ par $M \mapsto X^\top M X$ et continue par linéarité. Ainsi $f_X^{-1}(\mathbb{R}_+)$ est fermé pour tout $X \in \mathrm{M}_{n,1}(\mathbb{R})$ en tant qu'image

réciproque d'un fermé par une application continue et $\mathrm{S}_n(\mathbb{R})$ est fermé en tant qu'espace vectoriel de dimension finie, donc l'égalité ci-dessus montre bien que $\mathrm{S}_n^+(\mathbb{R})$ est fermé et que : $S \in \mathrm{S}_n^+(\mathbb{R})$.
Concluons. On a : $\forall p \in \mathbb{N}$, $A_{\varphi(p)} = O_{\varphi(p)}S_{\varphi(p)}$. Quand on prend la limite dans cette égalité quand $p \to +\infty$, la continuité du produit matriciel implique : $A = OS$. Or O est orthogonale et S symétrique positive par ce qui précède : d'où le résultat.
Il n'y a en revanche plus unicité, comme le démontre la décomposition suivante :

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{\in \mathrm{O}_2(\mathbb{R})} \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{\in \mathrm{S}_2^+(\mathbb{R})} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\in \mathrm{O}_2(\mathbb{R})} \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{\in \mathrm{S}_2^+(\mathbb{R})}.$$

Un autre exemple, trivial, est fourni par la matrice nulle, pour laquelle il suffit de prendre $S = 0_{\mathrm{M}_2(\mathbb{R})} \in \mathrm{S}_2^+(\mathbb{R})$ et pour O n'importe quelle matrice orthogonale.

Questions à se poser, réflexes à acquérir.

- Revoir si besoin les détails omis, pour la construction de la suite de matrices inversibles.
- Voici une illustration typique de l'emploi de la compacité pour faire converger des suites qui n'ont pas lieu d'être convergentes. À avoir en tête !
- Pourquoi ai-je changé O_p^{-1} en $O_p^\top$? Qu'aurais-je dû changer à ma rédaction, sans cela ?
- Est-ce que $\mathrm{S}_n^{++}(\mathbb{R})$ est fermé ? (Noter que je n'ai même pas fait mention que S est définie positive.)
- On note que la matrice symétrique positive choisie est la même dans les deux exemples : nous avons utilisé la partie orthogonale pour contredire l'unicité. Aurait-il été possible de contredire l'unicité en proposant deux matrices symétriques positives différentes ?
- A-t-on toujours $S^2 = A^\top A$, comme dans le cas défini positif ?

8. L'application φ est continue par continuité du produit matriciel et bijective par la question 4. La subtilité est donc la continuité de sa bijection réciproque $\varphi^{-1} : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$. Nous proposons deux démonstrations, qui sont toutes les deux déroutantes au premier abord (vous n'avez jamais démontré la continuité d'une application par ce moyen jusqu'à présent, sans doute), mais qui sont toutes les deux justifiées naturellement par une inspection méthodique de la situation : c'est φ qui a une expression simple, et non φ^{-1} (même si l'on pourrait me rétorquer qu'on a construit explicitement les matrices O et S de la décomposition $A = OS$; mais cette construction fait intervenir la racine carrée de la matrice $A^\top A$, et la continuité de cette application ne va pas de soi – elle semble se ramener à un problème de continuité des valeurs propres, affaire ô combien épineuse comme vous avez pu en juger à la fin du DST n° 6). Il est donc préférable de montrer la continuité de φ^{-1} par une caractérisation qui ne fait intervenir que φ, et non φ^{-1} : la caractérisation de la continuité par les images réciproques d'ouverts ou de fermés est une telle caractérisation ! C'est pourquoi je vais m'en servir ici ; le fait de privilégier les fermés aux ouverts est pour utiliser la compacité de $\mathrm{O}_n(\mathbb{R})$ (qui se traduit séquentiellement, comme les fermés, au contraire des ouverts).
Démonstration par les images réciproques des fermés. On sait que $\varphi^{-1} : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ est continue si et seulement si l'image réciproque par φ^{-1} (qui est en fait l'image directe par φ) de tout fermé relatif de $\mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ est un fermé relatif de $\mathrm{GL}_n(\mathbb{R})$. Considérons donc une partie F fermée dans $\mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ et montrons que $\varphi(F)$ est une partie fermée de $\mathrm{GL}_n(\mathbb{R})$. Soit $(A_p)_{p\in\mathbb{N}}$ une suite à valeurs dans $\varphi(F)$ et convergeant vers une matrice A inversible. Montrons : $A \in \varphi(F)$. Comme A est inversible, on sait qu'il existe $(O, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ tel que : $A = OS = \varphi(O, S)$. La subtilité est de montrer que (O, S) appartient à F. Nous allons essayer d'obtenir (O, S) comme limite d'éléments de F *via* les matrices A_p : on sait en effet qu'elles appartiennent à $\varphi(F)$, donc pour tout $p \in \mathbb{N}$ il existe $(O_p, S_p) \in F$ tel que : $A_p = O_pS_p$. Si l'on montre que $((O_p, S_p))_{p\in\mathbb{N}}$ converge vers (O, S), alors c'est gagné puisque F est un fermé. Comme la convergence se montre composante par composante, il suffit de montrer que $(O_p)_{p\in\mathbb{N}}$ et $(S_p)_{p\in\mathbb{N}}$ convergent respectivement vers O et S.

Faisons. On sait déjà que $(O_p)_{p\in\mathbb{N}}$ admet au moins une valeur d'adhérence, puisqu'elle est à valeurs dans le compact $\mathrm{O}_n(\mathbb{R})$. Notons $Q \in \mathrm{O}_n(\mathbb{R})$ cette valeur d'adhérence et montrons que $Q = O$. Comme :

$$S_p = O_p^{-1} A_p = O_p^\top A_p$$

pour tout $p \in \mathbb{N}$, et que $(A_p)_{p\in\mathbb{N}}$ converge, on en déduit aisément (par continuité du produit matriciel et de la transposition) que $(S_p)_{p\in\mathbb{N}}$ admet pour valeur d'adhérence $Q^\top A$. On a : $A = Q(Q^\top A)$; si l'on justifie que $Q^\top A$ est une matrice symétrique définie positive, l'unicité de la décomposition OS nous donnera $Q = O$ et $Q^\top A = S$; nous serons alors tout près de conclure. La matrice $Q^\top A$ est symétrique positive puisqu'elle est limite de matrices symétriques positives et parce que $\mathrm{S}_n^+(\mathbb{R})$ est un fermé (on l'a démontré dans la question 7), et elle est inversible puisqu'elle est le produit de deux matrices inversibles : on en déduit que son spectre ne contient pas 0, et comme il est inclus dans $\mathbb{R}_+$ par ce qui précède on a même : $\mathrm{Sp}_{\mathbb{R}}(Q^\top A) \subseteq \mathbb{R}_+^*$. C'est donc une matrice symétrique définie positive : ce qu'on voulait justifier. La décomposition $A = Q(Q^\top A)$ implique comme annoncé : $Q = O$, et : $Q^\top A = S$.

Ainsi $\left(\left(O_{\varphi(p)}, S_{\varphi(p)}\right)\right)_{p\in\mathbb{N}}$ est à valeurs dans F et converge vers (O, S). Comme F est fermé, on a : $(O, S) \in F$, puis : $A = \varphi(O, S) \in \varphi(F)$. Ceci montre que $\varphi(F)$ est fermé.

On a montré que l'image réciproque par φ^{-1} de tout fermé est un fermé, donc φ^{-1} est continue : d'où le résultat.

Démonstration par la caractérisation séquentielle de la continuité. C'est la compacité de $\mathrm{O}_n(\mathbb{R})$ qui invite à traduire séquentiellement la continuité de φ^{-1}, encore une fois. C'est-à-dire : soit $(A_p)_{p\in\mathbb{N}}$ une suite à valeurs dans $\mathrm{GL}_n(\mathbb{R})$ et convergeant vers une matrice inversible A. On veut montrer que si l'on note, pour tout $p \in \mathbb{N}$, l'unique couple $(O_p, S_p) \in \mathrm{O}_p(\mathbb{R}) \times \mathrm{S}_n^{++}(\mathbb{R})$ tel que : $A_p = O_p S_p$, alors $(\varphi^{-1}(A_p))_{p\in\mathbb{N}} = ((O_p, S_p))_{p\in\mathbb{N}}$ converge vers $\varphi^{-1}(A) = (O, S)$, c'est-à-dire : les suites $(O_p)_{p\in\mathbb{N}}$ et $(S_p)_{p\in\mathbb{N}}$ convergent respectivement vers O et S.

Par le même argument de compacité que ci-dessus, on sait que $(O_p)_{p\in\mathbb{N}}$ admet au moins une valeur d'adhérence et on en déduit que $(S_p)_{p\in\mathbb{N}}$ aussi ; grâce à l'unicité de la décomposition $A = OS$, on montre que ces valeurs d'adhérence sont nécessairement O et S respectivement (même argument que ci-dessus). Comme $(O_p)_{p\in\mathbb{N}}$ est à valeurs dans un compact et admet exactement une valeur d'adhérence, elle converge vers cette valeur d'adhérence, qui est O. On en déduit aisément que $(S_p)_{p\in\mathbb{N}} = \left(O_p^\top A_p\right)_{p\in\mathbb{N}}$ converge vers $O^\top A = S$. Ceci démontre ce qu'on voulait : $(\varphi^{-1}(A_p))_{p\in\mathbb{N}} = ((O_p, S_p))_{p\in\mathbb{N}}$ converge vers $\varphi^{-1}(A) = (O, S)$. Par caractérisation séquentielle de la continuité, φ^{-1} est continue sur $\mathrm{GL}_n(\mathbb{R})$.

Questions à se poser, réflexes à acquérir.

— Pourquoi tous ces efforts ? N'est-il pas évident que la réciproque d'une bijection continue est continue ? N'a-t-on pas un théorème de 1re année à ce sujet ?

— Ayant assimilé ces deux démonstrations, comprendre dans quelles configurations analogues il est pertinent de songer à ces caractérisations. Je préfère la première : pourquoi ? Que simplifie-t-elle d'un peu de vue mathématique ou rédactionnel ?

Vous pouvez vérifier si vous avez compris la démarche en montrant ce résultat général : si A et B sont des parties non vides de deux espaces vectoriels normés, avec A compact, et si $f : A \to B$ est une bijection continue, montrer que f^{-1} est continue sur B.

— Autre argument rarement utilisé (même s'il fut déjà présent dans le DM n° 7) : la convergence d'une suite à valeurs dans un compact grâce à l'unicité de la valeur d'adhérence. Que ce soit ici ou dans le DM cité, pourquoi cette idée ?

— Arriveriez-vous à démontrer qu'une matrice symétrique positive inversible est définie positive, sans raisonner sur le spectre comme ci-dessus ?

— Essayer de montrer la continuité de $A \mapsto \sqrt{A^\top A}$ directement (où $\sqrt{\cdot}$ désigne l'application qui à une matrice symétrique positive associe son unique racine carrée définie positive), et comprendre pourquoi c'est difficile, ce qui justifia cette autre démonstration.

Deuxième partie – deux applications

9. On a : $\overline{A} = \overline{UBU^{-1}}$, or les matrices A et B sont à cœfficients réels, donc : $A = \overline{U}B\overline{U}^{-1}$. On en déduit : $A^\top = \left(\overline{U}^\top\right)^{-1} B^\top \overline{U}^\top$, et comme : $\overline{U}^\top = U^{-1}$, on conclut : $A^\top = UB^\top U^{-1}$.
10. Le raisonnement est classique. Posons : $Q = \det(A + XB)$. C'est un polynôme puisque le déterminant est polynomial en les coefficients d'une matrice, et il est non nul puisque l'inversibilité de U assure : $Q(i) = \det(U) \neq 0$. Par conséquent Q admet un nombre fini de racines réelles : il existe $\mu \in \mathbb{R}$ tel que : $Q(\mu) \neq 0$, et la matrice $A + \mu B$ est alors inversible à coefficients réels. D'où le résultat.
11. Comme $A = UBU^{-1}$, on a : $AU = UB$, c'est-à-dire : $AX + iAY = XB + iYB$. Il suffit alors d'identifier les parties réelles et imaginaires, pour obtenir $AX = XB$ et $AY = YB$.
12. L'égalité de la question précédente implique : $AX + \mu AY = XB + \mu YB$. C'est-à-dire :

$$A(X + \mu Y) = (X + \mu Y)B.$$

Or $P = X + \mu Y$ est inversible et à coefficients réels par la question 10, donc : $A = PBP^{-1}$. En échangeant les rôles de A et $A^\top$, et de B et $B^\top$ (les deux matrices transposées vérifient la même relation que A et B d'après la question 9), on obtient : $A^\top = PB^\top P^{-1}$, d'où le résultat.

> **Questions à se poser, réflexes à acquérir.** Comme d'habitude, apprécier la puissance du recours au déterminant pour caractériser (topologiquement ou algébriquement : les polynômes couvrent tout) des propriétés matricielles, ici l'inversibilité.

13. Par ce qui précède : $PBP^{-1} = A = \left(P^\top\right)^{-1} BP^\top$, donc : $P^\top PB = BP^\top P$. Par définition de S (ou simplement en faisant le calcul de $S^2 = S^\top S = \left(O^{-1}P\right)^\top O^{-1}P$, si on n'a pas réussi à traiter la question 4), on en déduit : $S^2B = BS^2$. Par une récurrence facile, on en déduit :

$$\forall k \in \mathbb{N}, \quad \left(S^2\right)^k B = B\left(S^2\right)^k.$$

Ainsi les applications définies sur $\mathbb{R}[X]$ par $Q \mapsto Q(S^2)B$ et $Q \mapsto BQ(S^2)$ coïncident sur la base canonique, donc par linéarité on a : $\forall Q \in \mathbb{R}[X]$, $Q\left(S^2\right)B = BQ\left(S^2\right)$. Or S est la racine carrée symétrique positive de S^2, donc la question 3 assure qu'il existe $Q \in \mathbb{R}[X]$ tel que : $S = Q\left(S^2\right)$; en considérant un tel polynôme, on a : $SB = BS$. D'où le résultat.

> **Questions à se poser, réflexes à acquérir.**
> - Illustration de l'interpolation de Lagrange pour montrer une relation de commutation : premier exemple du devoir (et c'est loin d'être le seul : c'est un classique pour montrer des commutativités, songez-y !).
> - Faire la récurrence que j'ai omise, si vous n'êtes pas convaincus.
> - Apprécier le fait de raisonner sur une partie génératrice pour s'épargner la lourdeur rédactionnelle de l'introduction d'un polynôme $Q = \sum_k a_k X^k$, etc.

14. Il suffit de considérer la matrice O de la décomposition $P = OS$ fournie par l'énoncé. On a alors : $A = PBP^{-1} = OSBS^{-1}O^{-1} = OBO^{-1}$, l'égalité $SBS^{-1} = B$ résultant de la question précédente.
15. Supposons que $X \in \mathrm{GL}_n(\mathbb{R})$ est une solution de $(*)$. Par un raisonnement analogue à celui de la question 4 (voir la synthèse), on montre que la matrice $A^\top A$ est symétrique positive, donc ses valeurs propres sont positives. Montrons que ses valeurs propres sont strictement inférieures à 1. Soit λ une valeur propre de $A^\top A$, il existe $Y \in \mathrm{M}_{n,1}(\mathbb{R})$ non nul tel que : $A^\top AY = \lambda Y$. La première équation du système $(*)$ entraîne :

$$Y^\top A^\top AY + Y^\top X^\top XY = Y^\top Y,$$

c'est-à-dire : $\lambda\|Y\|^2 + \|XY\|^2 = \|Y\|^2$, ce qui s'écrit encore :

$$\frac{\|XY\|^2}{\|Y\|^2} = 1 - \lambda,$$

où l'on note que $\|XY\| > 0$ et $\|Y\| > 0$ car Y est non nul en tant que vecteur propre et X est inversible (donc XY est non nul également). On en déduit : $1-\lambda > 0$, d'où le résultat : $\lambda \in [0,1[$.

16. On cherche une solution X inversible : c'est la question 4 qui assure qu'une telle matrice peut s'écrire $X = UH$ avec U orthogonale et H symétrique définie positive : d'où le résultat.

    Remarque culturelle. Les notations U et H ne sont pas par hasard : le U renvoie à « unitaire » (l'équivalent complexe des matrices orthogonales) et H à « hermitien » (l'équivalent complexe des matrices symétriques réelles dans un contexte géométrique). Il s'avère que toute matrice complexe inversible peut s'écrire comme le produit d'une matrice unitaire et d'une matrice hermitienne définie positive.

17. Cherchons X sous la forme $X = UH$ avec U orthogonale et H symétrique définie positive. On a : $X^\top X = H^\top U^\top UH = H^2$. Par conséquent, si X vérifie $(*)$, alors la première équation donne :

$$\mathrm{I}_n - A^\top A = H^2,$$

ce qui conduit à prendre pour H l'unique racine carrée symétrique définie positive de $\mathrm{I}_n - A^\top A$, qui existe par la question 2 à condition que $\mathrm{I}_n - A^\top A$ soit elle-même symétrique définie positive. La symétrie ne pose pas de problème : justifions l'aspect défini positif. On a clairement :

$$\mathrm{Sp}_{\mathbb{R}}(\mathrm{I}_n - A^\top A) = \left\{1 - \lambda \mid \lambda \in \mathrm{Sp}(A^\top A)\right\}$$

(cela vaut pour toute matrice à laquelle on ajoute un multiple de l'identité : exercice facile). Par hypothèse de l'énoncé :

$$\mathrm{Sp}_{\mathbb{R}}(\mathrm{I}_n - A^\top A) \subseteq \mathbb{R}_+^*,$$

donc $\mathrm{I}_n - A^\top A$ est une matrice symétrique dont les valeurs propres sont strictement positives : c'est une matrice définie positive et la matrice H ci-dessus existe bien. Ce doit être l'unique racine carrée symétrique définie positive de $\mathrm{I}_n - A^\top A$.

Questions à se poser, réflexes à acquérir.

— Faire ce que je prétends : pourquoi les valeurs propres de $\alpha\mathrm{I}_n + \beta M$ sont de la forme $\alpha + \beta\lambda$ avec λ valeur propre de M ? Y a-t-il aussi un lien avec les sous-espaces propres ?

— A-t-on aussi $\mathrm{Sp}_{\mathbb{R}}(M - N) = \{\lambda - \mu \mid \lambda \in \mathrm{Sp}_{\mathbb{R}}(M), \mu \in \mathrm{Sp}_{\mathbb{R}}(N)\}$ plus généralement ?

18. La question précédente nous permet de déterminer H dans la décomposition $X = UH$ d'une solution éventuelle. Il reste à déterminer ce que doit être U, et à établir une réciproque. Soit $(O, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^+(\mathbb{R})$ tel que : $A = OS$. L'existence de cette décomposition découle de la question 7. La deuxième équation de $(*)$ nous enseigne que si $X = UH$ est solution, alors :

$$SO^\top UH = HU^\top OS.$$

On note que H est un polynôme en $\mathrm{I}_n - A^\top A$ par la question 3 : déduisons-en que H et S commutent. L'égalité $A = OS$ implique encore : $S^2 = A^\top A$, donc H est en fait un polynôme en S et on en déduit : $HS = SH$. Ce constat permet de montrer que $U = O$ vérifie la seconde égalité du système, puisque :

$$SO^\top OH = HO^\top OS \Longleftrightarrow SH = HS.$$

Réciproquement, soit H l'unique matrice symétrique définie positive telle que : $H^2 = \mathrm{I}_n - A^\top A$, dont l'existence fut justifiée dans la question précédente. Posons : $X = OH$. On a alors, comme H et S commutent par ce qui vient d'être justifié et comme O est orthogonale :

$$\begin{cases} A^\top A + X^\top X & = A^\top A + H^\top O^\top O H & = A^\top A + H^2 & = \mathrm{I}_n, \\ A^\top X - X^\top A & = (OS)^\top (OH) - H^\top O^\top (OS) & = SH - HS & = 0_{\mathrm{M}_n(\mathbb{R})}. \end{cases}$$

D'où le résultat : si $A^\top A$ est à valeurs propres dans $[0,1[$ alors $(*)$ admet une solution.

Questions à se poser, réflexes à acquérir.

- Deuxième occurrence de l'argument polynomial (par interpolation de Lagrange) pour montrer que deux matrices commutent. Y accorder de l'attention.
- Pouvait-on rédiger cette question et les précédentes de sorte à montrer directement l'équivalence voulue, sans faire une double implication ?

Troisième partie – valeurs propres d'une matrice

19. On trouve aisément après calculs : $P_1 = X - 2$ et $P_2 = (X-2)^2 - 1$. Ensuite, pour tout entier $p \geqslant 3$, un développement suivant la première colonne de P_p puis suivant la première ligne du déterminant obtenu à la deuxième étape, donne :

$$P_p = \begin{vmatrix} X-2 & 1 & 0 & \dots & 0 \\ 1 & X-2 & 1 & \ddots & \vdots \\ 0 & 1 & X-2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & X-2 \end{vmatrix} = (X-2)P_{p-1} - \begin{vmatrix} 1 & 0 & \dots & 0 \\ 1 & X-2 & \ddots & 0 \\ & \ddots & \ddots & 1 \\ \mathbf{0} & & 1 & X-2 \end{vmatrix}$$
$$= (X-2)P_{p-1} - P_{p-2},$$

ce qui fournit la relation de récurrence demandée.

20. Le cosinus réalise une bijection de $]0,\pi[$ dans $]-1,1[$, or $\frac{2-x}{2} \in]-1,1[$ par hypothèse sur x, donc il existe un unique $\theta \in]0,\pi[$ tel que : $2 - x = 2\cos\theta$.

On constate que $(P_p(x))_{p\geqslant 1}$ vérifie une relation de récurrence d'ordre 2 à coefficients constants, dont l'équation caractéristique :

$$r^2 - (x-2)r + 1 = 0 \iff r^2 + 2\cos(\theta) r + 1 = 0 \iff \left| r + e^{i\theta} \right|^2 = 0$$

admet pour racines $-e^{i\theta}$ et $-e^{-i\theta}$. Elles sont distinctes car $\theta \in]0,\pi[$. On a donc l'existence de $(\alpha,\beta) \in \mathbb{C}^2$ tel que :

$$\forall p \in \mathbb{N} \setminus \{0\}, \quad P_p(x) = \alpha \left(-e^{i\theta}\right)^p + \beta \left(-e^{-i\theta}\right)^p = (-1)^p \left(\alpha e^{ip\theta} + \beta e^{-ip\theta}\right).$$

Comme on sait que $P_p(x)$ est un nombre réel (en effet P_p est le polynôme caractéristique d'une matrice réelle et on a $x \in \mathbb{R}$), on a : $\forall p \in \mathbb{N} \setminus \{0\}$, $\overline{P_p(x)} = P_p(x)$, donc par indépendance linéaire des suites $\left((-e^{i\theta})^p\right)_{p\in\mathbb{N}}$ et $\left((-e^{-i\theta})^p\right)_{p\in\mathbb{N}}$ (ce sont des vecteurs propres de l'endomorphisme $T : (u_p)_{p\in\mathbb{N}} \mapsto (u_{p+1})_{p\in\mathbb{N}}$ associés aux valeurs propres différentes $-e^{i\theta}$ et $-e^{-i\theta}$) cela implique : $\overline{\alpha} = \beta$. Autrement dit :

$$\forall p \in \mathbb{N} \setminus \{0\}, \quad P_p(x) = \alpha(-1)^p e^{ip\theta} + \overline{\alpha}(-1)^p e^{-ip\theta} = 2(-1)^p \mathrm{Re}\left(\alpha e^{ip\theta}\right).$$

Grâce aux valeurs de $P_1(x) = x - 2 = -2\cos(\theta) = -e^{i\theta} - e^{-i\theta}$ et $P_2(x) = (2\cos(\theta))^2 - 1 = e^{2i\theta} + e^{-2i\theta} + 1$, on obtient :

$$\begin{cases} \alpha e^{i\theta} + \overline{\alpha} e^{-i\theta} & = & e^{i\theta} + e^{-i\theta} \\ \alpha e^{2i\theta} + \overline{\alpha} e^{-2i\theta} & = & e^{2i\theta} + e^{-2i\theta} + 1 \end{cases}$$

L'opération $L_2 \leftarrow L_2 - e^{-i\theta} L_1$ donne :

$$\alpha = \frac{e^{2i\theta}}{e^{2i\theta}-1} = \frac{e^{i\theta}}{2i\sin(\theta)},$$

donc :

$$\forall p \in \mathbb{N}\setminus\{0\}, \quad P_p(x) = \frac{(-1)^p}{\sin(\theta)}\mathrm{Re}\left(\frac{e^{i(p+1)\theta}}{i}\right) = (-1)^p \frac{\sin((p+1)\theta)}{\sin(\theta)}.$$

Questions à se poser, réflexes à acquérir.

— Comment ai-je reconnu $\left|r+e^{i\theta}\right|^2$ dès que j'ai vu $r^2+2\cos(\theta)+1$? Se souvenir de mes commentaires disséminés çà et là sur l'identité remarquable $|z \pm z'|^2 = |z|^2 \pm 2\mathrm{Re}(z\bar{z}') + |z'|^2$, en particulier dans le cas $z \in \mathbb{R}$ et $z' = \rho e^{i\theta}$.

— Pourquoi ai-je préféré exprimer $P_p(x)$ avec des exponentielles complexes, plutôt que comme une combinaison linéaire de $(-1)^p\cos(p\theta)$ et $(-1)^p\sin(p\theta)$? Au vu de la question posée, cela paraissait pourtant avisé. J'ai pourtant une bonne raison de ne pas avoir procédé ainsi !

— Peut-on aussi expliciter P_p, pas seulement évalué en des réels x convenables ? Faire éventuellement le lien avec les polynômes de Tchebychev.

— Qu'est-ce qui, ici, pouvait conduire à l'idée de calculer le polynôme caractéristique à travers une relation de récurrence, alors qu'on le fit rarement voire jamais ? La réponse se situe à la fois dans la taille et la forme de la matrice.

— Est-ce que cette approche aurait aussi fonctionné pour expliciter le polynôme caractéristique, dans le cas très général des matrices tridiagonales vues dans le DST n° 10 ?

21. Chercher les valeurs propres de A_p revient à déterminer les racines de P_p. Commençons par les chercher parmi les réels x qui vérifient $|2-x| < 2$. Par la question précédente, et avec les mêmes notations, on a :

$$\begin{aligned} P_p(x) = 0 &\iff \sin((p+1)\theta) = 0 \iff \exists k \in [\![1,p]\!], \quad (p+1)\theta = k\pi \qquad (\text{car } \theta \in]0,\pi[) \\ &\iff \exists k \in [\![1,p]\!], \quad \theta = \frac{k\pi}{p+1}. \end{aligned}$$

Cela fournit p racines $x_k = 2 - 2\cos\left(\frac{k\pi}{p+1}\right)$. Elles sont distinctes par injectivité du cosinus sur $]0,\pi[$. Comme P_p est de degré p, on les a toutes déterminées. On en déduit :

$$\mathrm{Sp}_{\mathbb{R}}(A_p) = \left\{2 - 2\cos\left(\frac{k\pi}{p+1}\right) \mid k \in [\![1,p]\!]\right\}.$$

22. La matrice A_p est symétrique réelle donc elle est diagonalisable par le théorème spectral (autre argument : elle admet p valeurs propres distinctes ; cet argument assure en plus que les sous-espaces propres sont des droites). Déterminons les sous-espaces propres.

Soit $k \in [\![1,p]\!]$. Posons : $\lambda_k = 2 - 2\cos\left(\frac{k\pi}{p+1}\right)$. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathrm{M}_{p,1}(\mathbb{R})$. Alors :

$$A_p X = \lambda_k X \iff \left\{ \begin{array}{rcl} (2-\lambda_k)x_1 - x_2 &=& 0, \\ \forall \ell \in [\![2,p-1]\!], \quad -x_{\ell-1} + (2-\lambda_k)x_\ell - x_{\ell+1} &=& 0, \\ -x_{p-1} + (2-\lambda_k)x_p &=& 0. \end{array}\right.$$

En posant la convention $x_0 = x_{p+1} = 0$, on a simplement :

$$A_p X = \lambda_k X \iff \forall \ell \in [\![1,p]\!], \quad x_{\ell-1} - (2-\lambda_k)x_\ell + x_{\ell+1} = 0.$$

On remarque que $(x_\ell)_{\ell\in[\![0,p+1]\!]}$ vérifie presque la même relation de récurrence que $(P_p(x))_{p\geqslant 1}$ à la question 20. En reprenant cette étude (cette fois-ci l'équation caractéristique est : $\left|r-e^{i\theta}\right|^2=0$, dont les racines distinctes sont $e^{i\theta}$ et $e^{-i\theta}$) et en notant que l'on a : $x_0=0$, on trouve d'abord qu'il existe $\alpha\in\mathbb{C}$ tel que :

$$\forall \ell\in[\![0,p+1]\!],\quad x_\ell=2\mathrm{Re}\left(\alpha e^{\frac{ik\ell\pi}{p+1}}\right),$$

et ensuite que l'on a : $\alpha=-\overline{\alpha}$, donc : $\alpha\in i\mathbb{R}$. Prenons par exemple $\alpha=-\frac{i}{2}$. Alors le vecteur X de coefficients :

$$\forall \ell\in[\![1,p]\!],\quad x_\ell=\mathrm{Re}\left(ie^{\frac{ik\ell\pi}{p+1}}\right)=\sin\left(\frac{k\ell\pi}{p+1}\right),$$

est non nul et est un vecteur propre de A_p associé à λ_k d'après l'étude qui précède, et comme les sous-espaces propres de A_p sont des droites c'est un vecteur directeur de $\ker(A_p-\lambda_k)\mathrm{I}_n)$. On a montré :

$$\forall k\in[\![1,p]\!],\quad \ker(A_p-\lambda_k\mathrm{I}_p)=\mathrm{Vect}_{\mathbb{R}}\left(\begin{pmatrix}\sin\left(\frac{k\pi}{p+1}\right)\\ \vdots\\ \sin\left(\frac{kp\pi}{p+1}\right)\end{pmatrix}\right).$$

Questions à se poser, réflexes à acquérir.

- Songer à avoir l'initiative d'introduire des notations conventionnelles pour simplifier un calcul et éviter des distinctions de cas (autre occurrence banale : la transformation d'Abel).
- On note que les x_k et $(P_p)_{p\geqslant 1}$ vérifient presque la même relation de récurrence. Est-ce explicable ? fréquent ?
- Je n'ai pas utilisé la relation $x_{p+1}=0$ pour déterminer les x_k. Est-ce possible ? Utile ?
- Obtient-on un résultat analogue en remplaçant les 2 et -1 de la matrice par des réels a et b quelconques ?

Quatrième partie

23. On applique le théorème de représentation de Riesz avec le produit scalaire usuel de $\mathrm{M}_n(\mathbb{R})$, défini par $(M,N)\mapsto \mathrm{tr}(M^\top N)$: l'application $B\mapsto \big(M\mapsto \mathrm{tr}(B^\top M)\big)$ est un isomorphisme de $\mathrm{M}_n(\mathbb{R})$ dans $\mathrm{L}(\mathrm{M}_n(\mathbb{R}),\mathbb{R})$. Comme f est une forme linéaire, il existe donc $B\in\mathrm{M}_n(\mathbb{R})$ telle que : $\forall M\in\mathrm{M}_n(\mathbb{R})$, $f(M)=\mathrm{tr}(B^\top M)$. Il suffit de poser $A=B^\top$ pour avoir le résultat voulu.

Questions à se poser, réflexes à acquérir. Même si je n'ai pas utilisé le théorème de Riesz à cet effet dans le cours, noter *via* cette question et les suivantes qu'il a l'utilité sur $\mathrm{M}_n(\mathbb{R})$ de nous ramener à la trace (pour profiter de ses propriétés agréables, qui ne sont pas propres à toutes les formes linéaires *a priori* puisqu'elles concernent le produit).

24. L'application f est linéaire sur $\mathrm{M}_n(\mathbb{R})$ qui est de dimension finie, donc f est continue. Or $\mathrm{O}_n(\mathbb{R})$ est un compact non vide d'après la question 6, donc par le théorème des bornes atteintes la restriction de f à $\mathrm{O}_n(\mathbb{R})$ est bornée et atteint ses bornes : d'où l'existence de $M_n=\sup\limits_{O\in\mathrm{O}_n(\mathbb{R})} f(O)$.
25. On a déjà dit à plusieurs reprises (questions 4 et 15) que $A^\top A$ est une matrice symétrique positive, donc elle est diagonalisable par le théorème spectral et elle admet n valeurs propres positives comptées avec leurs multiplicités.
26. La définition de D incite à considérer la racine carrée de A : la construction explicite de cette matrice dans la question 4 montre en effet que la partie symétrique positive de la décomposition $A=OS$ est égale à la racine carrée de $A^\top A$, et la construction de racine carrée de la question 2

montre que cette racine carrée fait intervenir la matrice D de coefficients diagonaux $\sqrt{\mu_i}$, où les μ_i sont les valeurs propres de $A^\top A$.

Reproduisons tout cela ici, dans une démarche auto-suffisante : considérons deux matrices $P \in \mathrm{O}_n(\mathbb{R})$ et $D' \in \mathrm{M}_n(\mathbb{R})$ diagonale, de coefficients diagonaux les valeurs propres $\mu_1, \ldots, \mu_n$ de $A^\top A$, telles que : $A^\top A = PD'P^{-1}$. Leur existence est assurée par le théorème spectral. Soit $(U, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{S}_n^+(\mathbb{R})$ tel que : $A = US$. On a :

$$S^2 = S^\top S = \left(U^{-1}A\right)^\top \left(U^{-1}A\right) = A^\top A,$$

donc S est l'unique racine carrée de $A^\top A$, et la construction explicite de la question 2 montre que l'on a : $S = PDP^{-1}$, où D est la matrice diagonale de coefficients diagonaux $\sqrt{\mu_1}, \ldots, \sqrt{\mu_n}$. Alors :

$$M_n = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OA) = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OUS),$$

et comme $O \mapsto OU$ est une permutation de $\mathrm{O}_n(\mathbb{R})$ (car c'est un groupe : il est stable par produit) de réciproque $O \mapsto OU^{-1}$, on a :

$$M_n = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OS) = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OPDP^{-1}) = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(P^{-1}OPD).$$

Même argument : comme P est orthogonale, l'application $O \mapsto P^{-1}OP$ est une permutation de $\mathrm{O}_n(\mathbb{R})$, donc :

$$M_n = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OD),$$

d'où le résultat.

Questions à se poser, réflexes à acquérir.

— Constater le retour de la permutation $g \mapsto gh$ fréquemment utilisée dans les groupes, pour simplifier des sommes et produits. Le contexte change mais on fait au fond la même chose, toujours : bien relever la commodité apportée en bien des endroits par cette permutation, pour créer de l'invariance, simplifier des calculs, etc.

— Exemple d'utilisation de la décomposition OS : on va pouvoir mesurer l'écart de f (représentée par A) au groupe orthogonal plus simplement.

27. Comme I_n est orthogonale, on a :

$$M_n = \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}(OD) \geqslant \mathrm{tr}(D) = \sum_{i=1}^n \sqrt{\mu_i}.$$

Montrons l'inégalité en sens contraire. Si $O = ((\omega_{i,j}))_{1 \leqslant i,j \leqslant n}$, alors : $\mathrm{tr}(OD) = \sum_{i=1}^n \omega_{i,i}\sqrt{\mu_i}$. Or O est orthogonale, donc ses colonnes sont unitaires et cela impose : $\forall (i,j) \in [\![1,n]\!]^2$, $\omega_{i,j} \leqslant 1$. Donc :

$$\mathrm{tr}(OD) \leqslant \sum_{i=1}^n \sqrt{\mu_i}.$$

Cette majoration est indépendante de O. Par la propriété de plus petit majorant de la borne supérieure, on a : $M_n \leqslant \sum_{i=1}^n \sqrt{\mu_i}$. D'où le résultat :

$$M_n = \sum_{i=1}^n \sqrt{\mu_i}.$$

Questions à se poser, réflexes à acquérir.

— Comme souvent, une égalité avec une borne supérieure se calcule par double inégalité. Une majoration indépendante de la variable pour majorer la borne supérieure, puis une évaluation en un cas particulier donnant le cas d'égalité, pour l'inégalité en sens contraire (si l'on ne parvient pas à obtenir le cas d'égalité, passer à la limite ; mais ici, un argument théorique assure que le cas d'égalité est possible, pourquoi ?).

— Si l'on avait eu $\sup\limits_{g \in \mathrm{O}(E)} \operatorname{tr}(f \circ g)$ à exprimer à l'aide des valeurs propres de $f^* \circ f$ (traduction géométrique de l'énoncé que nous avons démontré), comment auriez-vous procédé sans passer par les matrices ? Y parvenir, puis se convaincre que le raisonnement matriciel est au fond la même chose, mais en plus simple (nul besoin de se coltiner une matrice de passage puisqu'on peut directement raisonner dans la « bonne » base : sempiternelle raison qui rend beaucoup de raisonnements géométriques plus souples que les matriciels).

28. La matrice A provient du théorème de représentation de Riesz, et nous avons dit en remarque dans le cours qu'en présence d'une base orthonormée, on pouvait expliciter l'isomorphisme réciproque de ce théorème. Comme la base « canonique » de $\mathrm{M}_n(\mathbb{R})$ est orthonormée pour le produit scalaire usuel, nous allons nous en sortir sans difficulté.

Posons $A = \big((a_{i,j})\big)_{1 \leqslant i,j \leqslant n}$. Comme le produit scalaire usuel $(M,N) \mapsto \operatorname{tr}(M^\top N)$ de $\mathrm{M}_n(\mathbb{R})$ est simplement la somme des produits des coefficients de M et N, on a pour toute matrice $M = \big((m_{i,j})\big)_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{R})$:

$$f(M) = \operatorname{tr}(AM) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{j,i} m_{i,j}.$$

Mais on a aussi d'après l'énoncé : $f(M) = \sum\limits_{j=1}^{n} \sum\limits_{i=j}^{n} m_{i,j}$. Il suffit donc de prendre A telle que :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad a_{j,i} = \begin{cases} 1 & \text{si } i \geqslant j, \\ 0 & \text{si } i < j, \end{cases}$$

c'est-à-dire :

$$A = \begin{pmatrix} 1 & \cdots & \cdots & 1 \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Questions à se poser, réflexes à acquérir.

— Retenir cette observation pourtant triviale, mais dont peu d'élèves ont conscience : nul besoin de calculer AM puis d'en prendre la trace pour obtenir $\operatorname{tr}(AM)$. Il suffit de faire la somme des produits des coefficients de $A^\top$ et M, indice par indice !

— Si l'on ne « voit » pas la matrice qui convient, comment la trouver *méthodiquement* ?

— Revoir si besoin le cours, concernant l'explicitation du vecteur représentant une forme linéaire quand on connaît une base orthonormée.

29. On peut passer par une banale inversion de système, mais il est plus élégant de remarquer que si l'on pose :

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix},$$

alors $N^n = 0_{\mathrm{M}_n(\mathbb{R})}$ et on a : $A = \sum_{k=0}^{n-1} N^k$. On en déduit :

$$(\mathrm{I}_n - N)\, A = \sum_{k=0}^{n-1} \left(N^k - N^{k+1}\right) = \mathrm{I}_n - N^n = \mathrm{I}_n.$$

Dans $\mathrm{M}_n(\mathbb{R})$, un inverse à gauche est aussi un inverse à droite, donc :

$$A^{-1} = \mathrm{I}_n - N = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Questions à se poser, réflexes à acquérir.

— Obtenir cet inverse par d'autres approches (méthode du pivot ?). Comparer les mérites.

— Pourquoi ai-je songé à introduire cette matrice N pour trivialiser le calcul de l'inverse ? Qu'est-ce qui m'y incitait, dans la forme de A ?

30. Posons : $J = A^{-1}(A^{-1})^\top = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$ (pour effectuer rapidement un calcul du type $MM^\top$ ou $M^\top M$ à la main, il est utile de constater que le coefficient (i,j) de $MM^\top$ est toujours $\langle L_i, L_j\rangle$ et celui de $M^\top M$ est $\langle C_i, C_j\rangle$, où les L_i et C_i désignent les lignes et colonnes de M respectivement).

Attention, ce n'est pas la matrice A_n de la partie III, à cause du coefficient d'indice (n,n). Pas grave, nous allons nous y ramener : si l'on note $E_{n,n} \in \mathrm{M}_n(\mathbb{R})$ la matrice dont tous les coefficients sont nuls sauf celui d'indice (n,n), on a par linéarité du déterminant par rapport à la dernière colonne :

$$\begin{aligned}
\chi_J &= \begin{vmatrix} X-2 & 1 & 0 & \dots & 0 \\ 1 & X-2 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & X-2 & 1 \\ 0 & \dots & 0 & 1 & X-1 \end{vmatrix} \\
&= \begin{vmatrix} X-2 & 1 & 0 & \dots & 0 \\ 1 & X-2 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & X-2 & 1 \\ 0 & \dots & 0 & 1 & X-2 \end{vmatrix} + \begin{vmatrix} X-2 & 1 & 0 & \dots & 0 \\ 1 & X-2 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & X-2 & 0 \\ 0 & \dots & 0 & 1 & 1 \end{vmatrix} \\
&= P_n + \begin{vmatrix} X-2 & 1 & & \mathbf{0} \\ 1 & X-2 & \ddots & \\ & 1 & \ddots & 1 \\ \mathbf{0} & & \ddots & X-2 \end{vmatrix} \\
&= P_n + P_{n-1}.
\end{aligned}$$

On en déduit, pour tout $x \in \mathbb{R}$ tel que : $|2-x| < 2$, que l'on écrit $2 - x = 2\cos(\theta)$:

$$\begin{aligned}
\chi_J(x) = P_n(x) + P_{n-1}(x) &\overset{(q.20)}{=} \frac{(-1)^n}{\sin(\theta)}\left(\sin((n+1)\theta) - \sin(n\theta)\right) \\
&= (-1)^n \frac{2\cos\left(\left(n+\frac{1}{2}\right)\theta\right)\sin\left(\frac{\theta}{2}\right)}{2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)} \\
&= (-1)^n \frac{\cos\left(\left(n+\frac{1}{2}\right)\theta\right)}{\cos\left(\frac{\theta}{2}\right)}.
\end{aligned}$$

On en déduit :

$$\chi_J(x) = 0 \iff \exists k \in [\![0, n-1]\!],\ \left(n + \frac{1}{2}\right)\theta = \frac{\pi}{2} + k\pi \iff \exists k \in [\![0, n-1]\!],\ \theta = \frac{(2k+1)\pi}{2n+1}.$$

L'injectivité du cosinus sur $]0, \pi[$ assure que les n racines $x_k = 2\left(1 - \cos\left(\frac{(2k+1)\pi}{2n+1}\right)\right)$ de χ_J ainsi produites sont toutes distinctes. Comme χ_J est de degré n, on les a toutes déterminées. Donc :

$$\mathrm{Sp}_{\mathbb{R}}\left(A^{-1}\left(A^{\top}\right)^{-1}\right) = \left\{2\left(1 - \cos\left(\frac{(2k+1)\pi}{2n+1}\right)\right) \mid k \in [\![0, n-1]\!]\right\}.$$

Par anticipation sur la question suivante, nous allons réécrire commodément ces valeurs propres. On a :

$$\forall k \in [\![1, n]\!], \quad 1 - \cos\left(\frac{(2k+1)\pi}{2n+1}\right) = 2\left[\sin\left(\frac{(2k+1)\pi}{2(2n+1)}\right)\right]^2,$$

donc :

$$\begin{aligned}
\mathrm{Sp}_{\mathbb{R}}\left(A^{-1}\left(A^{\top}\right)^{-1}\right) &= \left\{4\left[\sin\left(\frac{(2k+1)\pi}{2(2n+1)}\right)\right]^2 \mid k \in [\![0, n-1]\!]\right\} \\
&= \left\{4\left[\sin\left(\frac{(2(n-k)+1)\pi}{2(2n+1)}\right)\right]^2 \mid k \in [\![1, n]\!]\right\} \\
&= \left\{4\left[\sin\left(\frac{-2k\pi}{2(2n+1)} + \frac{\pi}{2}\right)\right]^2 \mid k \in [\![1, n]\!]\right\} \\
&= \left\{4\left(\cos\left(\frac{k\pi}{2n+1}\right)\right)^2 \mid k \in [\![1, n]\!]\right\}.
\end{aligned}$$

Questions à se poser, réflexes à acquérir.

- Vérifier mon observation sur le produit $MM^{\top}$ ou $M^{\top}M$. Faire le lien avec la caractérisation des matrices orthogonales et sur l'écriture des matrices de Gram comme un produit de cette forme.
- Reprendre les commentaires de la question 20 qui s'appliquent à χ_J.
- À quelles symétries visibles graphiquement correspond le passage de $\sin\left(\frac{(2k+1)\pi}{2(2n+1)}\right)$ à $\cos\left(\frac{k\pi}{2n+1}\right)$, dans l'écriture du spectre ?

31. On a : $A^{\top}A = J^{-1}$, et les valeurs propres de J^{-1} sont les inverses des valeurs propres de J par la question 1, donc d'après la question précédente : $\forall k \in [\![1, n]\!]$, $\mu_k^{-1} = 4\left(\cos\left(\frac{k\pi}{2n+1}\right)\right)^2$ (quitte à réordonner les μ_k). On en déduit, par la question 27 :

$$M_n = \sum_{k=1}^{n} \sqrt{\mu_k} = \sum_{k=1}^{n} \frac{1}{2\cos\left(\frac{k\pi}{2n+1}\right)}.$$

Questions à se poser, réflexes à acquérir.

- Pourquoi l'énoncé nous a-t-il fait passer par J pour obtenir les valeurs propres de $A^\top A$? Qu'est-ce qui a été simplifié en passant par $A^{-1}\left(A^\top\right)^{-1}$?
- Après avoir vu le chapitre XIII : pouvait-on obtenir M_n par des techniques de calcul différentiel ?

32. En remontant les calculs effectués à la fin de la question 30, on sait montrer :

$$\forall k \in [\![1,n]\!], \quad \sum_{k=1}^{n} \frac{1}{\cos\left(\frac{k\pi}{2n+1}\right)} = \sum_{k=0}^{n-1} \frac{1}{\sin\left(\frac{(2k+1)\pi}{2(2n+1)}\right)}$$

et on sait que l'on a : $\forall x \in \mathbb{R}_+$, $x - \frac{x^3}{6} \leqslant \sin(x) \leqslant x$. La première inégalité se démontre rapidement *via* la formule de Taylor avec reste intégral :

$$\forall x \in \mathbb{R}_+, \quad \sin(x) = x + x^3 \int_0^1 \frac{(1-t)^2}{2} \sin''(tx)\mathrm{d}t = x - x^3 \int_0^1 \frac{(1-t)^2}{2} \underbrace{\sin(tx)}_{\leqslant 1} \mathrm{d}t \geqslant x - \frac{x^3}{6}.$$

On en déduit l'encadrement :

$$\frac{1}{2}\sum_{k=0}^{n-1} \frac{1}{\frac{(2k+1)\pi}{2(2n+1)}} \leqslant M_n \leqslant \frac{1}{2}\sum_{k=1}^{n} \frac{1}{\frac{(2k+1)\pi}{2(2n+1)} - \frac{1}{6}\left(\frac{(2k+1)\pi}{2(2n+1)}\right)^3}.$$

Cherchons des équivalents des deux extrémités de l'encadrement, que nous notons $S_{n,1}$ et $S_{n,2}$ ci-dessous. La première est très facile à étudier par une comparaison série-intégrale. On a en effet, suivant un raisonnement très classique :

$$\sum_{k=0}^{n-1} \frac{1}{2k+1} \underset{n\to+\infty}{\sim} \int_0^n \frac{\mathrm{d}x}{2x+1} \underset{n\to+\infty}{\sim} \frac{\ln(n)}{2},$$

donc : $S_{n,1} \underset{n\to+\infty}{\sim} \dfrac{(2n+1)\ln(n)}{2\pi} \underset{n\to+\infty}{\sim} \dfrac{n\ln(n)}{\pi}$. Passons à la seconde somme. Nous n'arriverons pas à nous « débarrasser » de la présence de n dans la somme cette fois-ci, à cause des puissances différentes au dénominateur. Procédons autrement, en montrant que la différence entre $S_{n,1}$ et $S_{n,2}$ est négligeable devant $S_{n,1}$ au voisinage de $+\infty$. On a :

$$0 \leqslant S_{n,2} - S_{n,1} = \frac{1}{12}\sum_{k=0}^{n-1} \frac{\frac{(2k+1)\pi}{2(2n+1)}}{1 - \frac{1}{6}\left(\frac{(2k+1)\pi}{2(2n+1)}\right)^2} = \frac{1}{12}\sum_{\substack{\ell=0\\ \ell \text{ impair}}}^{2n} \frac{\frac{\ell\pi}{2(2n+1)}}{1 - \frac{1}{6}\left(\frac{\ell\pi}{2(2n+1)}\right)^2} \leqslant \frac{1}{12}\sum_{\ell=0}^{2n} \frac{\frac{\ell\pi}{2(2n+1)}}{1 - \frac{1}{6}\left(\frac{\ell\pi}{2(2n+1)}\right)^2}.$$

On reconnaît, à un facteur multiplicatif près, la somme de Riemann de pas régulier $\frac{\pi}{2(2n+1)}$ associée à la fonction $x \mapsto \frac{x}{1-\frac{x^2}{6}}$ sur $\left[0, \frac{\pi}{2}\right]$. Elle est continue sur $\left[0, \frac{\pi}{2}\right]$ (en effet $1 - \frac{x^2}{6} = 0$ pour $x = \sqrt{6} \notin \left[0, \frac{\pi}{2}\right]$, puisque : $\sqrt{6} \leqslant \frac{\pi}{2} \iff 24 \leqslant \pi^2$, ce qui est faux). On a donc :

$$\lim_{n\to+\infty} \frac{\pi}{2(2n+1)}\sum_{\ell=0}^{2n} \frac{\frac{\ell\pi}{2(2n+1)}}{1 - \frac{1}{6}\left(\frac{\ell\pi}{2(2n+1)}\right)^2} = \int_0^{\frac{\pi}{2}} \frac{x}{1-\frac{x^2}{6}}\mathrm{d}x = \left[-3\ln\left(1-\frac{x^2}{6}\right)\right]_0^{\frac{\pi}{2}} = -3\ln\left(1-\frac{\pi^2}{24}\right).$$

On en déduit : $S_{n,2} - S_{n,1} = \underset{n\to+\infty}{O}(n) = \underset{n\to+\infty}{o}(n\ln(n))$, donc : $S_{n,2} = S_{n,1} + \underset{n\to+\infty}{o}(S_{n,1}) \underset{n\to+\infty}{\sim} S_{n,1}$. Par le théorème des gendarmes :

$$M_n \underset{n\to+\infty}{\sim} S_{n,1} \underset{n\to+\infty}{\sim} \frac{n\ln(n)}{2\pi}.$$

Remarque. Pour minorer M_n ci-dessus, on aurait pu vouloir utiliser l'inégalité : $\forall x \in \mathbb{R}, \cos(x) \leqslant 1 - \frac{x^2}{2}$. Cependant cette inégalité, d'excellente qualité près de 0, se dégrade au voisinage de $\frac{\pi}{2}$; pire, la fonction $x \mapsto \frac{x^2}{2}$ change de signe et est strictement négative en $\frac{\pi}{2}$, donc le passage à l'inverse ne va pas conserver l'inégalité. Ce n'est pas le seul souci : comme le cosinus s'annule en $\frac{\pi}{2}$, la quantité $\frac{1}{\cos}$ devient très grande au voisinage de $\frac{\pi}{2}$; or ce n'est pas le cas de $x \mapsto \left(1 - \frac{x^2}{2}\right)^{-1}$. Voilà pourquoi il fallut procéder autrement.

Le passage par le sinus a résolu le problème, bien qu'on semble faire la même chose que ci-dessus en utilisant la minoration : $\forall x \in \mathbb{R}_+, \sin(x) \leqslant x$. Il n'en est rien : cette inégalité est excellente au voisinage de 0 et même si elle n'est plus aussi bonne au voisinage de $\frac{\pi}{2}$, cette fois-ci $\frac{1}{\sin}$ ne devient pas arbitrairement grand au voisinage de $\frac{\pi}{2}$ (au contraire de $\frac{1}{\cos}$), donc l'écart avec $x \mapsto x^{-1}$ ne devient pas rédhibitoire. C'est ce qui a motivé mes calculs (on pouvait aussi remplacer l'approximation $\cos(x) \leqslant 1 - \frac{x^2}{2}$, où l'on compare le cosinus avec son développement limité en 0, avec une inégalité qui compare le cosinus avec son développement limité en $\frac{\pi}{2}$; c'est au fond ce que l'on fait en passant par le sinus).

C'est un raisonnement analogue mais plus fin qui justifie que l'inégalité $\cos(x) \geqslant 1 - \frac{2x}{\pi}$, valable sur $\left[0, \frac{\pi}{2}\right]$, n'aurait pas permis d'aboutir (en fait ici l'approximation est déjà trop mauvaise aux voisinages de 0 et π : comparer les tangentes en 0 et π des deux graphes pour s'en convaincre).

Questions à se poser, réflexes à acquérir.

— Faire la comparaison série-intégrale que j'ai omise.

— Toutes les sommes en présence ressemblent beaucoup à des sommes de Riemann associées à une subdivision régulière. Pourquoi ne pas avoir essayé d'y recourir ? (J'affirme que ce point de vue aurait permis de *conjecturer* que $M_n \underset{n\to+\infty}{\longrightarrow} +\infty$, pourquoi ?)

— Pouvait-on prédire cet équivalent (ou au moins son ordre de grandeur) par des arguments qualitatifs non rigoureux ?