Formes quadratiques

Marc SAGE

2 juillet 2005

Table des matières

Notations.

b.o. : base orthogonale
b.o.n. : base orthonormée
b.o.n.d. : base orthonormée directe

A^* : transconjuguée de la matrice A, qui s'identifie avec la transposée dans le cas réel

$\mathcal{O}_n$: groupe orthogonal (réel)
$\mathcal{U}_n$: groupe unitaire (complexe)

$\mathcal{S}_n$: matrices symétriques réelles
$\mathcal{S}_n^+$: matrices symétriques réelles positives
$\mathcal{S}_n^{++}$: matrices symétriques réelles définies positives

$\mathcal{H}_n$: matrices hermitiennes (complexes)
$\mathcal{H}_n^+$: matrices hermitiennes positives
$\mathcal{H}_n^{++}$: matrices hermitiennes définies positives

L'exposant $^{+*}$ (dans $\mathcal{S}_n^{+*}$ ou $\mathcal{H}_n^{+*}$) serait plus cohérent pour décrire le caractère *défini* positif (à l'instar de $\mathbb{R}^{+*}$), mais $^{++}$ est tellement plus simple à écrire...

1 Compacité des groupes orthogonal et unitaire

Montrer que $\mathcal{O}_n$ et $\mathcal{U}_n$ sont compacts.

Solution proposée.
$M_n(\mathbb{R})$ étant de dimension finie, il suffit de montrer que $\mathcal{O}_n$ est un fermé borné. Les relations d'orthogonalité des colonnes montrent qu'une matrice orthogonale est bornée par 1 en norme infinie (on peut aussi dire que sa norme 2 vaut constamment n). Par ailleurs, $\mathcal{O}_n$ est fermé comme l'image réciproque du fermé $\{I_n\}$ par l'application continue $A \mapsto AA^*$.
Évidemment, tout cela s'adapte sans soucis pour passer de $\mathcal{O}_n$ à $\mathcal{U}_n$.

2 Racine carrée dans $\mathcal{S}_n^+$ et $\mathcal{H}_n^+$

Montrer qu'un endomorphisme auto-adjoint positif admet une unique racine carrée auto-adjointe positive. Énoncé le résultat obtenu pour les matrices réelles puis complexes.

Solution proposée.
On notera bien le cas $n = 1$: un réel positif admet une unique racine carrée dans $\mathbb{R}^+$!
Soit u notre endomorphisme auto-adjoint positif. Le théorème spectral permet de diagonaliser u dans une certaine base $\mathcal{B}_d$:

$$\operatorname*{Mat}_{\mathcal{B}_d} u = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

où les λ_i sont réels positifs par hypothèse, admettant ainsi des racines carrées dans $\mathbb{R}^+$. Il est alors plus que naturel de définir une racine r par

$$\operatorname*{Mat}_{\mathcal{B}_d} r := \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}.$$

Il reste à vérifier qu'en fait le r ci-dessus est le seul possible.

Soit r une racine carrée de u auto-adjointe positive. r et u sont diagonalisables par le théorème spectral, mais comme r commute avec $r^2 = u$, r et u sont en fait co-diagonalisables, disons

$$\underset{\mathcal{B}_d}{\mathrm{Mat}}\, u = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \text{ et } \underset{\mathcal{B}_d}{\mathrm{Mat}}\, r = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}.$$

La condition $r^2 = u$ impose alors $\mu_i^2 = \lambda_i$ pour tout i, d'où $\mu_i = \sqrt{\lambda_i}$ par posivité de r. Ainsi, r vaut $\sqrt{\lambda_i}\,\mathrm{Id}$ sur le sous-espace propre de u associé à λ_i, ce qui détermine entièrement r vu que les sous-espaces propres de u recomposent tout l'espace.

Matriciellement, le résultat se traduit par :

$$\textit{tout } S \in \mathcal{S}_n^+ \textit{ admet une unique racine carrée dans } \mathcal{S}_n^+.$$

Pour le lecteur nécessiteux d'explications, notons $\mathcal{B}_c$ la base canonique de $\mathbb{R}^n$ et s l'endomorphisme canoniquement associé à S. On a donc

$$S = \underset{\mathcal{B}_c}{\mathrm{Mat}}\, s.$$

D'après ce qui précède, s admet une unique racine carrée r réelle symétrique positive, ce qui permet d'écrire

$$\left(\underset{\mathcal{B}_c}{\mathrm{Mat}}\, r\right)^2 = \underset{\mathcal{B}_c}{\mathrm{Mat}}\left(r^2\right) = \underset{\mathcal{B}_c}{\mathrm{Mat}}\, s = S.$$

En posant $R = \mathrm{Mat}_{\mathcal{B}_c}\, r$, on a la racine voulue. Pour l'unicité, si R' est une autre racine, en notant r' l'endmorphisme associé, on devrait avoir

$$\underset{\mathcal{B}_c}{\mathrm{Mat}}\left(r'^2\right) = \underset{\mathcal{B}_c}{\mathrm{Mat}}\, s \implies r'^2 = s \implies r = r'$$

par l'unicité que l'on a démontrée pour les endomorphismes.

Pour les matrices complexes, on énonce :

$$\textit{tout } H \in \mathcal{H}_n^+ \textit{ admet une unique racine carrée dans } \mathcal{H}_n^+.$$

La preuve est calquée sur le cas réel.

Remarque. Il est bon de noter que si u est auto-adjoint *défini* positif, alors la racine est aussi *définie* positive (ses valeurs propres sont des racines carrées de réels > 0). Matriciellement, cela donne

$$\textit{tout } S \in \mathcal{S}_n^{++} \textit{ admet une unique racine carrée dans } \mathcal{S}_n^{++},$$
$$\textit{tout } H \in \mathcal{H}_n^{++} \textit{ admet une unique racine carrée dans } \mathcal{H}_n^{++}.$$

3 Décomposition polaire

On sait qu'un nombre complexe se décompose comme le produit de son module, réel positif que l'on peut voir comme un élément de $\mathcal{H}_1^+$, par un élément du cercle unité, que l'on peut toujours voir comme une matrice unitaire de taille 1 étant donné que

$$e^{i\theta} \times \left(e^{i\theta}\right)^* = e^{i\theta} \times e^{-i\theta} = 1$$

(on voit d'ailleurs d'où vient la terminologie *unitaire* pour les matrices). On a même unicité de la décomposition pour les complexes non nuls.

On se propose de généraliser cela à la dimension n.

Montrer qu'un matrice complexe se décompose comme le produit d'une matrice hermitienne positive par une matrice unitaire :

$$\forall A \in M_n(\mathbb{C}),\ \exists (H, U) \in \mathcal{H}_n^+ \times \mathcal{U}_n,\ A = HU,$$

avec unicité de H dans tous les cas et unicité de U si A est inversible.

Énoncer la décomposition "polaire" pour les matrices réelles.

Solution proposée.

Raisonnons par analyse-synthèse. Partant de $A = HU$ et utilisant le caractère unitaire de U, on obtient

$$AA^* = HU\,(HU)^* = HUU^*H^* = H^2.$$

H doit donc être l'unique racine de AA^* dans $\mathcal{H}_n^+$ (cf. exercice précédent) ; vérifions que cette dernière est bien dans $\mathcal{H}_n^+$: pour un vecteur x, on a

$$\langle AA^*x \mid x\rangle = (AA^*x)^*\,x^* = x^*AA^*x = \|A^*x\|^2 \geq 0$$

(avec $=$ ssi $x = 0$ dans le cas A inversible), d'où la positivité de AA^*, le caractère hermitien étant évident. H est donc entièrement déterminé par la matrice A, et il en est alors de même pour $U = AH^{-1}$ si H est inversible, *i.e.* si AA^* est définie (d'après la remarque de l'exercice précédent), ce qui le cas si A est inversible (cf. cas d'égalité plus haut). Fin de l'analyse.

Synthèse. Soit A une matrice complexe et H l'unique racine de A^*A dans $\mathcal{H}_n^+$. On a envie de poser $U = AH^{-1}$, ce qui incite à supposer A inversible dans un premier temps. Il est alors immédiat de vérifier que U est unitaire et indigne de montrer $A = HU$. Si A n'est pas inversible, on se ramène au cas précédent en approchant A par des matrices inversibles :

$$A = \lim A_n.$$

On dispose de la décomposition polaire de chaque $A_p = H_pU_p$. Puisque $\mathcal{U}_n$ est compact, quitte à extraire, on peut supposer que (U_p) converge vers une matrice unitaire U. Il est alors clair que H_p converge vers AU^{-1}, qui est hermitienne vu que $\mathcal{H}_n^+$ est fermé (attention, on perd le caratère *défini* positif en passant à la limite) ; en appelant H, la limite des H_p, on a la décomposition $A = HU$.

Pour les matrices réelles, on reprend la même preuve en remplacant les H par des S et les U par des O :

$$\forall A \in M_n(\mathbb{R}),\ \exists\,(S,O) \in \mathcal{S}_n^+ \times \mathcal{O}_n,\ A = SO,$$

la partie symétrique étant unique, l'unicité de O ne tenant que pour les matrices inversibles.

Remarques. Il est évident que l'unicité de la partie unitaire se perd pour les matrices non inversibles : peut-on définir l'argument de 0 ? En ce sens, l'unicité obtenue pour la décomposition polaire des matrices non inversibles est optimale.

On pourait également adapter la preuve pour obtenir une décomposition $A = UH$ ou OS, selon les goûts de chacun, même s'il est plus usuel de mettre le module en premier dans $z = re^{i\theta}$... De fait, les deux décompositions ne coïncident pas en général puisque H et S sont définis par $\sqrt{AA^*}$ ou $\sqrt{A^*A}$ (observer que, dans le cas $n = 1$, on retrouve bien la formule exprimant le module $|z| = \sqrt{zz^*}$) et que A n'a aucune raison de commuter avec son adjoint. Noter par ailleurs que H et U ne commutent pas en général, comme le montre l'exemple

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \neq \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

On peut donner une autre preuve de l'existence de la partie unitaire. Une fois posé $H = \sqrt{AA^*}$, on peut remarquer que

$$\forall x \in E,\ \|Ax\|^2 = x^*A^*Ax = x^*H^2x = x^*H^*Hx = \|Hx\|.$$

L'exercice 10 de la feuille sur le produit scalaire garantit alors l'existence d'un U unitaire tel que $A = HU$.

4 Pour ne plus raconter de bêtises sur les matrices positives

Une erreur (et confusion) très fréquente est induite par la terminologie des matrices dites *positives*.

Une forme quadratique q est dite *positive* si elle ne prend que des valeurs positives, *i.e.* si pour tout vecteur x on a $q(x) \geq 0$; jusque là, la terminologie est claire. Maintenant, une matrice auto-adjointe sera dite *positive* si la forme quadratique associée l'est, *i.e.* si pour tout vecteur x on a $x^*Ax \geq 0$.

Observer bien que la définition n'a de sens que pour les matrices *auto-adjointes* (celles qui représentent une forme quadratique) ; il est donc totalement stupide d'essayer de parler d'une matrice "positive" qui ne serait pas auto-adjointe. Ainsi, tout énoncé dérivé, aussi naturel soit-il, à l'instar de "A est positive ssi tous ses coefficients sont positifs" doit être proscrit, car une telle définition pourrait s'appliquer à n'importe quelle matrice.

(noter qu'il ne s'agit là que d'un moyen mnémotechnique pour éviter l'écueil. Dans la vraie vie, les matheux cherchent très souvent à généraliser une définition en montrant qu'elle est équivalente à telle propriété pouvant s'appliquer à une classe plus large d'objets : par exemple, la proposition "*une application est continue ssi l'image réciproque de tout ouvert est ouverte*" permet de généraliser la continuité sur les evn aux espaces topologiques généraux).

Donnons à présent une "bonne" raison pour ne plus faire de confusion. Toute matrice A auto-adjointe se réduisant diagonalement en b.o., sa forme quadratique associée prend la forme suivante (dans la b.o. considérée) :

$$q(x) = \sum \lambda_i x_i^2.$$

On voit très bien sur la formule ci-dessus que le signe de $q(x)$ ne dépend que des λ_i et que la base considérée ne joue aucun rôle : ce qui importe vraiment avec ces histoires de signes, ce sont les *valeurs propres* de A.

Tout ce qu'on a fait pour obtenir A, c'est déplacer la b.o.n. canonique de K^n dans l'espace (peut-être en remplaçant quelques uns de ses vecteurs par leurs opposés), ce qui n'a strictement aucun effet sur notre problème (il ne s'agit que d'un changement de point de vue), puis dilater les vecteurs de base à l'aide de coefficients λ_i, et là on fait vriament quelque chose : ce sont les λ_i qui contiennent toute l'information de A.

Moralité, étant donné une matrice A auto-adjointe :

A est *positive* ssi toutes ses valeurs propres le sont.

En identifiant A à sa forme quadratique associée, cela s'écrit aussi

$$A \geq 0 \iff \operatorname{Sp} A \geq 0.$$

Il est alors naturel de noter "$q > 0$" pour "q définie positif", puisque cela se traduit par

$$A > 0 \iff \operatorname{Sp} A > 0,$$

soit, en d'autres termes :

A est *définie positive* ssi toutes ses valeurs propres sont > 0.

Pour finir, nous laisserons méditer le lecteur sur les quelques contre-énoncés suivants. Nous l'invitons à en chercher d'autres et à les illustrer par les fruits de sa propre réflexion.

a) *Montrer qu'une matrice définie positive n'a pas forcément tous ses coefficients* ≥ 0.

b) *Montrer qu'une matrice symétrique ayant sa diagonale* > 0 *n'est pas forcément positive. Même conclusion si tous les coefficients sont supposés* > 0.

c) *Montrer qu'avoir tous ses coefficients hors diagonale* < 0 *n'empêche pas d'être positif.*

d) *Montrer qu'une matrice symétrique à diagonale nulle n'est pas forcément dégénérée.*

e) *Montrer qu'une matrice symétrique à coefficients tous* < 0 *peut être définie positive.*

Solution proposée.

a) Considérer $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$: la forme quadratique associée est

$$q(x,y) = 2x^2 + 2y^2 - 2xy = x^2 + y^2 + (x-y)^2 \geq 0$$

avec $=$ ssi $x = y = 0$, ce qui montre que q est définie positive. On peut aussi calculer le polynôme caractéristique $X^2 - 4X + 3 = (X-1)(X-3)$, d'où un spectre $\{1, 3\}$ strictement positif.

b) Le polynôme caractéristique de $\begin{pmatrix} 1 & -4 \\ -4 & 1 \end{pmatrix}$ vaut $X^2 - 2X - 15 = (X+3)(X-5)$ d'où une valeur propre négative -3 : impossible d'être positif.

Le polynôme caractéristique de $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ vaut $X^2 - 2X - 3$, donc le produit des valeurs propres (qui sont réelles car la matrice est symétrique) est négatif (il vaut -3), d'où deux valeurs propres de signes opposés : même conclusion.

c) Reprendre l'exemple $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ du *a)*.

d) Le déterminant de $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est non nul, donc la forme quadratique est non dégénérée.

e) Le polynôme caractéristique de $\begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$ vaut $X^2 + 2X - 3 = (X-1)(X-2)$, donc le spectre étudié vaut $\{1, 2\}$, de sorte que la forme quadratique est définie positive.

Remarque. Les deux premiers points montrent bien qu'il n'y a aucun rapport entre le caractère (défini) positif d'une matrice et la (stricte) positivité de ses coefficients. L'exercice suivant propose cependant un critère satisfaisant.

5 Critère pour déterminer la positivité d'une matrice symétrique

L'exercice précédent nous laisse un peu sur notre faim. Voici enfin un critère de positivité en fonction de la positivité, non pas des coefficients, mais de certains déterminants extraits.

Pour une matrice $A \in M_n(\mathbb{R})$ et I une partie de $\{1, ..., n\}$, on notera A_I la matrice extraite $(a_{i,j})_{i,j\in I}$.

Soit S symétrique.

- *Montrer que S est définie positive ssi* $\det S_{\{1,...,r\}} > 0$ *pour tout* $r = 1, ..., n$.
- *Montrer que S est positive ssi* $\det S_I \geq 0$ *pour tout* $I \subset \{1, ..., n\}$.

Solution proposée.

Notons q la forme quadratique associée à S.

• Le sens directe est clair : q reste > 0 par restriction aux r premiers vecteurs de base, donc le déterminant de $S_{\{1,...,r\}}$ est > 0 car produit de r valeurs propres > 0.

Pour l'autre sens, on raisonne par récurrence. En se restreignant aux $n-1$ vecteurs $e_1, ..., e_{n-1}$ de la base canonique, $S_{\{1,...,n-1\}}$ vérifie les hypothèses au rang $n-1$, donc (par récurrence) q est > 0 sur $\mathrm{Vect}\{e_1, ..., e_{n-1}\}$, d'où une famille orthonormée de $n-1$ vecteurs pour q. On aimerait bien la compléter en un base orthogonale pour simplifier la matrice de q. On commence par compléter en une base $(d_1, ..., d_n)$, puis on va perturber le dernier vecteur car c'est celui qu'on a rajouté sans aucun rapport avec q (tuons l'intrus !). Cherchons donc un $d = d_n + \sum_{i<n} \lambda_i d_i$ tel que $\widetilde{q}(d, d_i) = 0$ pour tout $i < n$, ce qui s'écrit

$$\forall i < n,\ \widetilde{q}(d, d_i) + \underbrace{q(d_i)}_{=1}\lambda_i = 0,$$

d'où les λ_i cherchés.

Dans cette nouvelle base, q s'écrit

$$\begin{pmatrix} I_{n-1} & \\ & \lambda \end{pmatrix},$$

et en invoquant la seule condition encore non utilisée, à savoir $\det S > 0$, on voit que $\lambda > 0$, ce qui satisfait à notre bonheur.

• Le sens direct se fait comme pour le premier point.

Pour l'autre sens, le raisonnement précédent ne marche pas (essentiellement parce que les coefficients $q(d_i)$ devant les λ_i que l'on cherche peuvent s'annuler par isotropie de q).

La méthode est alors classique : pour montrer un résultat sur du ≥ 0, on le montrer sur du > 0, on perturbe le ≥ 0, on applique le cas > 0, puis on étouffe la perturbation.

Le cas > 0 venant d'être traité à la question précédente, perturbons. On regarde donc le déterminant de $S + \varepsilon I_n$, qui n'est autre que le polynôme caractéristique de S évalué en $\pm\varepsilon$:

$$\det(S + \varepsilon I_n) = \varepsilon^n + \sum_{0\leq i<n} \left(\sum_{|I|=n-i} \det S_I \right) \varepsilon^i$$

(cf. feuille sur les déterminants pour une preuve élémentaire de cette formule). Cela tombe bien, on a de l'information sur les $\det S_I$. Pour une perturbation $\varepsilon > 0$, on obtient quelque chose de > 0. En appliquant aux sous-matrice $S_{\{1,...,r\}}$, dont les déterminants extraits sont en fait extraits de S, on peut appliquer le premier point : $S + \varepsilon I_n > 0$, ceci tenant $\forall \varepsilon > 0$.

Il reste à étouffer la perturbation. En fixant un vecteur x, la condition trouvée implique $x^* (S + \varepsilon I_n) x \geq 0$, ce qui a le bon goût d'être conservé lorqu'on étouffe ε :

$$x^* S x \geq 0.$$

Ceci tenant pour tout x, S reste positive.

Remarque. Concernant le premier point où l'on a compléter un b.o.n. en une b.o., on montre plus généralement que si la restriction de q à un sev F de dim finie est > 0, alors $F \oplus F^{\perp} = E$; on aurait appliqué cela à $F = \text{Vect}\{e_1, ..., e_{n-1}\}$.

Attention, le premier critère fait défaut pour le cas ≥ 0 : il s'agit de chercher une matrice dont les $\det A_{\{1,...r\}}$ sont ≥ 0 (mais pas tous > 0, sinon le premier critère s'appliquerait...) et dont l'un des $\det A_I$ est < 0. La matrice $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ fait l'affaire.

6 Quotients de Rayleigh

Soit q une forme quadratique. À d fixé, déterminer

$$\min_{\dim V = d} \max_{v \in V \cap \mathbb{S}} q(v)$$

où $\mathbb{S}$ désigne la sphère unité euclidienne (on pourra ordonner les valeurs propres de q).

Solution proposée.

On peut expliciter la quantité $q(v)$ en se plaçant dans une b.o. $(e_1, ..., e_n)$ de q, disons $q(e_i) = \lambda_i$ avec $\lambda_1 \leq ... \leq \lambda_n$: si $v = \sum v_i e_i$ avec $\|v\| = 1$, on aura

$$q(v) = q\left(\sum v_i e_i\right) = \sum \lambda_i v_i^2.$$

Cette dernière quantité est clairement majorée par λ_n et minorée par λ_1 vu que $\sum v_i^2 = 1$. Mais v ne variant que dans V, l'on est pas sûr d'avoir égalité, par exemple pour $V = \text{Vect}\{e_2, e_3\}$.

De façon plus précise, on peut majorer/minorer $\sum \lambda_i v_i^2$ par le plus grand/petit λ_i tel que il y ait un $v \in V$ possédant une coordonnée selon e_i non nulle. Par exemple, pour $V = \text{Vect}\{e_1, ..., e_d\}$, on aura $q(v) \leq \lambda_d$ avec égalité pour $v = e_d$, d'où $\max_{v \in V \cap \mathbb{S}} q(v) = \lambda_d$. Mais un V quelconque de dimension d recoupe nécessairement $\text{Vect}\{e_d, ..., e_n\}$, donc contient un vecteur unitaire de la forme $\sum_{i=d}^{n} \lambda_i e_i$, d'où $\max_{v \in V \cap \mathbb{S}} q(v) \geq \lambda_d$. Finalement :

$$\min_{\dim V = d} \max_{v \in V \cap \mathbb{S}} q(v) = \lambda_d.$$

Remarque. L'appellation "quotients de Rayleigh" vient probablement de l'écriture

$$\min_{\dim V = d} \max_{v \in V \setminus \{0\}} \frac{\langle sv, v \rangle}{\langle v, v \rangle}$$

où s désigne l'endomorphisme auto-adjoint associé à q.

7 Centre de $\mathcal{O}_n$ et $\mathcal{U}_n$

Quels sont les centres des groupes orthogonal et unitaire ?

Solution proposée.
Soit ω dans le centre de $\mathcal{O}_n$ ou $\mathcal{U}_n$. ω stabilise en particulier $\operatorname{Fix}\omega$ pour tout automorphisme orthogonal/unitaire ω. Or, on peut toujours choisir pour $\operatorname{Fix}\omega$ une droite arbitraire (prendre la symétrie orthogonale par rapport à cette droite), ce qui montre que ω stabilise toutes les droites. Il est alors classique que ω est une homothétie. Comme son déterminant doit être 1 en modulo, il ne reste plus grand monde :

$$\begin{aligned} Z(\mathcal{O}_n) &= \{\pm \operatorname{Id}\}, \\ Z(\mathcal{U}_n) &= \{\lambda \operatorname{Id} \ ; \ \lambda \in \mathcal{U}_1\}. \end{aligned}$$

8 Une optimisation, pour se faire les mains

Soit S une matrice réelle symétrique. Déterminer

$$\inf \operatorname{tr}\left[(S - \lambda XX^*)^2\right]$$

où l'infimum porte sur tous les réels λ et tous les vecteurs colonne X de la sphère unité.

Solution proposée.
Commençons par développer bêtement :

$$\operatorname{tr}\left[(S - \lambda XX^*)^2\right] = \operatorname{tr}\left(S^2\right) - 2\lambda \operatorname{tr}(SXX^*) + \lambda^2 \operatorname{tr}\left[(XX^*)^2\right].$$

On va s'occuper de chacun des termes séparément, puis minorer ce trinôme du second degré en λ. Commençons par diagonaliser $S = ODO^*$ en base orthonormée et notons d_i les valeurs propres de D ainsi que d le plus grand des d_i (sans considérer les modules).

Le premier terme se calcule aisément :

$$\operatorname{tr}\left(S^2\right) = \operatorname{tr}\left(OD^2O^*\right) = \operatorname{tr}\left(D^2\right) = \sum_{i=1}^{n} d_i^2.$$

Quant à $\operatorname{tr}\left[(XX^*)^2\right]$, remarquer qu'une matrice D diagonalisable de rang 1 vérifie $\operatorname{tr}\left(D^2\right) = (\operatorname{tr} D)^2$, la valeur commune étant le carré de l'unique valeur propre non nulle ; puisque c'est le cas de XX^* (utiliser le théorème spectral), on a

$$\operatorname{tr}\left[(XX^*)^2\right] = \left[\operatorname{tr}(XX^*)\right]^2 = \left[\operatorname{tr}(X^*X)\right]^2 = \langle X \mid X\rangle^2 = \|X\|^4 = 1.$$

Notre trinôme en λ est donc unitaire.

On a envie de minorer tout de suite :

$$\begin{aligned} \operatorname{tr}\left[(S - \lambda XX^*)^2\right] &= \operatorname{tr}\left(S^2\right) - 2\lambda \operatorname{tr}(SXX^*) + \lambda^2 \underbrace{\operatorname{tr}\left[(XX^*)^2\right]}_{=1} \\ &= \operatorname{tr}\left(S^2\right) + \left[\lambda - \operatorname{tr}(SXX^*)\right]^2 - \left[\operatorname{tr}(SXX^*)\right]^2 \\ &\geq \operatorname{tr}\left(S^2\right) - \left[\operatorname{tr}(SXX^*)\right]^2. \end{aligned}$$

Il s'agit donc de majorer $\operatorname{tr}(SXX^*)$, en se souvenant que O conserve le produit scalaire et que le vecteur X est unitaire :

$$\begin{aligned} \operatorname{tr}(SXX^*) &= \operatorname{tr}(X^*SX) = \langle X \mid SX\rangle = \langle X \mid ODO^*X\rangle = \langle O^*X \mid DO^*X\rangle \\ &= \sum_{i=1}^{n} d_i \left[O^*X\right]_i^2 \leq d \|O^*X\|^2 = d \|X\|^2 = d. \end{aligned}$$

Notons que l'égalité est atteinte (entre autres) pour O^*X une colonne de 0 avec un 1 à la k-ième ligne si $d_k = d$. Pour ce choix de X, la minoration trois lignes plus haut devient une égalité pour $\lambda = \mathrm{tr}\,(SXX^*) = d$

Il en résulte

$$\inf_{\substack{\lambda\in\mathbb{R}\\ \|X\|=1}} \mathrm{tr}\left[(S-\lambda XX^*)^2\right] = \sum_{i\neq k}\lambda_i^2$$

où k désigne l'indice de la plus grande valeur valeur propre.

9 Classifications des formes quadratiques sur un corps fini

Soit K un corps fini à q éléments. On rappelle que q est une puissance de la caractéristique de K. On pourra montrer que K contient $\frac{q+1}{2}$ carrés pour q impair.

Montrer que toute matrice symétrique S à coefficients dans K est congrue à $\begin{pmatrix} I_{n-1} & \\ & \lambda \end{pmatrix}$ où λ est soit 1 (forme quadratique standard), soit n'est pas un carré dans K.

Solution proposée.

Quitte à se placer dans une b.o., on peut toujours supposer S sous forme diagonale $S = \mathrm{Diag}\,(\lambda_1, ..., \lambda_n)$. Un bête calcul matriciel permet, combiné avec une récurrence, de se ramener au cas $n = 2$:

$$\begin{aligned}
S &= \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \begin{pmatrix} P\begin{pmatrix} I_{n-2} & \\ & \lambda \end{pmatrix}{}^tP & \\ & & \lambda_n \end{pmatrix} \text{ où } \lambda \in K \\
&= P'\begin{pmatrix} I_{n-2} & & \\ & \lambda & \\ & & \lambda_n \end{pmatrix}{}^tP' \text{ avec } P' = \begin{pmatrix} P & \\ & 1 \end{pmatrix} \\
&= P'\begin{pmatrix} I_{n-2} & & \\ & Q\begin{pmatrix} 1 & \\ & \mu \end{pmatrix}{}^tQ \end{pmatrix}{}^tP' \text{ où } \mu = 1 \text{ ou } \mu \text{ non carré} \\
&= P''\begin{pmatrix} I_{n-1} & \\ & \mu \end{pmatrix}{}^tP'' \text{ avec } P'' = P'\begin{pmatrix} I_{n-2} & \\ & Q \end{pmatrix}.
\end{aligned}$$

Il suffit donc de traiter les cas $n = 1$ et 2.

Pour $n = 1$, la matrice S est un scalaire λ, qui est soit un carré c^2, auquel cas $S = (c)\,(1)\ {}^t(c)$ est la forme quadratique standard, soit un non carré, *CQFD*.

Pour $n = 2$, soit $S = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}$. On veut mettre un 1 en haut à gauche de la matrice, ce qui correspond à trouver un vecteur $u = (x, y)$ de K^2 tel que ${}^tuSu = 1$, *i.e.* à résoudre l'équation

$$\lambda x^2 + \mu y^2 = 1.$$

Distinguons deux cas selon que la caractéristique de K soit 2 ou non.

Si q est pair, tout élément x de K est un carré puisque $x^q = x$: en effet, l'ordre d'un $x \in K^*$ divise le cardinal de K^*, ce qui s'écrit $x^{q-1} = 1$, *i.e.* $x^q = x$, ce qui reste valable pour $x = 0$. S est donc congrue à l'identité, et le problème est résolu.

Pour q impair, montrons qu'il y a $\frac{q+1}{2}$ carrés dans K. Il suffit d'introduire le morphisme de groupes "élévation au carré" :

$$c : \left\{ \begin{array}{ccc} K^* & \longrightarrow & K^* \\ x & \longmapsto & x^2 \end{array} \right. .$$

Son noyau est $\{1, -1\}$ qui est de cardinal 2 puisqu'on est en caractéristique impaire. Le théorème de factorisation des morphismes permet d'écrire

$$\mathrm{Im}\, c \simeq {}^{K^*}\!\diagup_{\mathrm{Ker}\, c},$$

d'où le résultat en prenant les cardinaux et en rajoutant le carré manquant 0.

10 Théorème de Springer

Soit $q(x) = \sum a_{i,j} x_i x_j$ une forme quadratique anisotrope sur k^n. Soit $k \hookrightarrow K$ une extension de k. Montrer que q considérée comme polynôme homogène de degré 2 à n variables dans K reste anisotrope dans les cas suivants :

- *$K = k\left((X_i)_{i\in I}\right)$ est un corps des fractions rationnelles en des indéterminées X_i ;*
- *$k \hookrightarrow K$ est de dimension finie impaire en tant que k-ev.*

Solution proposée.

Quitte à se placer dans une b.o., on peut supposer $q(x) = \sum \lambda_i x_i^2$ où $\lambda_i \in k$. Considérons par l'absurde un vecteur $x \in K^n$ isotrope non nul.

• On peut supposer I fini en ne considérant que les indéterminées qui apparaissent dans les coordonnéées de x. Puis on peut toujours faire une récurrence sur le nombre d'indéterminées et suppposer $K = k(X)$. En éliminant les dénominateurs, on obtient une expression

$$\sum \lambda_i P_i^2 = 0$$

où les P_i peuvent être pris premiers entre eux quitte à diviser par leur pgcd ; en particulier, X ne peut pas tous les diviser. En évaluant en 0, on trouve un vecteur isotrope non nul dans k^n, c'est une contradiction.

• q est isotrope sur le corps $k(x_1, ..., x_n)$ engendré par les x_i, et quitte à faire une récurrence sur le nombre de générateurs, on peut supposer $K = k(\alpha)$ monogène. On précède alors par récurrence sur le degré $d := \deg \alpha = [K : k]$ de l'extension.

Les composantes de x étant des polynômes P_i en α non tous nul de degré $< d$, on peut écrire $\sum \lambda_i P_i(\alpha)^2 = 0$, ce qui signifie $\sum \lambda_i P_i^2$ multiple du polynôme minimal de α, mettons

$$\sum \lambda_i P_i^2 = A\mu_\alpha.$$

Quitte à diviser par le pgcd des P_i dans $k[X]$ (qui doit diviser A car μ_α est irréductible), on peut supposer les P_i premiers entre eux. En notant $m = \max \deg P_i < d$, on voit par anisotropie de q que le terme en X^m du terme de gauche n'est pas nul : on en déduit

$$\deg A = 2m - d \leq 2(d-1) - d = d - 2$$

qui est impair. A admet donc un facteur irréductible B de degré impair, et en appelant β une racine de ce facteur dans un bon corps de décomposition (B est donc le polynôme minimal de β dans l'extension $k \hookrightarrow k(\beta)$), on voit que $\sum \lambda_i P_i(\beta)^2 = 0$, d'où un vecteur isotrope dans l'extension $k(\beta)$ qui est de degré $\deg B$ impair $< d$, d'où par récurrence $P_i(\beta) = 0$ pour tout i, ce qui contredit la primalité relative des P_i dans $k(\beta)$ (le pgcd est inchangé par extension de corps...).

Remarque. On vient de montre la conservation de l'anisotropie par extension impaire ou transcendante pure des scalaires. Le résultat tombe complètement en défaut pour les extension de degré pair. En effet, si le nombre de variables est au moins 2, on peut contruire une extension de k de degré pair où apparait un vecteur isotrope. Soit $q(x_1, ..., x_n) = \lambda x_1^2 + \mu x_2^2 + ...$ dans une base de diagonalisation, avec $\lambda \neq 0$ (sinon $q = 0$ est isotrope). Le polynôme $\lambda X^2 + \mu$ est irréductible sur k car sans racine par anisotropie de q, donc admet une racine α dans l'extension quadratique $k \hookrightarrow {}^{k[X]}\!/_{\lambda X^2+\mu}$. Le vecteur $(\alpha, 1, 0, ..., 0)$ est alors isotrope dans l'extension.

11 Connexité de $\mathcal{O}_n, \mathcal{U}_n, \mathcal{SO}_n, \mathcal{SU}_n$

Étudier la connexité des groupes $\mathcal{O}_n$, $\mathcal{U}_n$, $\mathcal{SO}_n$ et $\mathcal{SU}_n$.

Solution proposée.

On va essayer de passer des formes réduites à l'identité.

Pour $\mathcal{O}_n$, la forme réduite est une diagonale de blocs de rotation du type $\begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}$ suivi d'une diagonale de 1 puis éventuellement d'un -1 si l'isométrie est indirecte. Les rotations glissent vers l'identité en étouffant a vers 0, ce qui montre que $\mathcal{O}_n$ a au plus deux composantes connexes (selon le déterminant). S'il

n'en avait qu'une, son image $\{-1,1\}$ par det serait un connexe de $\mathbb{R}$, ce qui n'est pas. Finalement, $\mathcal{O}_n$ a deux composantes connexes (par arcs) : $\mathcal{SO}_n$ et $\mathcal{O}_n^-$.

Concernant $\mathcal{U}_n$, on diagonalise avec des valeurs propres unitaires. En faisant glisser chacune de ces valeurs le long du cercle unité vers 1, on relie tout le monde à l'identité, d'où la connexité de $\mathcal{U}_n$.

Pour $\mathcal{SU}_n$, les valeurs propres sont de produit 1. Pour conserver cette condition, lorsque qu'on fait glisser une valeur le long du cercle unité vers 1, on fait bouger en même temps la valeur suivante dans le sens inverse, ce qui ne change pas le déterminant. On envoie ainsi les $n-1$ premières valeurs propres sur 1. La valeur propre restante est alors le déterminant de notre matrice de départ, *i.e.* 1. Nous venons de montrer que $\mathcal{SU}_n$ est connexe par arcs.

12 Points extrémaux de la boule unité dans $M_n(\mathbb{C})$

Soit C un convexe d'un evn. On appelle *point extrémal* de C tout point $e \in C$ qui n'est dans l'intérieur d'aucun segement de C :

$$e \in [ab] \subset C \implies e = a \text{ ou } b.$$

On notera $\operatorname{Extr} C$ l'ensemble des points extrémaux de C. On pourra admettre (voire redémontrer...) qu'un point $c \in C$ est extrémal ssi

$$\forall a, b \in C,\ c = \frac{a+b}{2} \implies a = b = c.$$

On se place dans $M_n(\mathbb{C})$ vu comme $\mathcal{L}(\mathbb{C}^n)$ muni de la norme subordonnée à la norme hermitienne $\|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}$. On rappelle au besoin que la norme subordonnée d'une matrice $A \in M_n(\mathbb{C})$ est donnée par

$$\|A\| = \sqrt{\rho(AA^*)}$$

où $\rho(A) = \max_{\lambda \in \operatorname{Sp} A} |\lambda|$ désigne le rayon spectral de A.

On s'intéresse à la boule unité fermée $\mathcal{B}_n$ de $M_n(\mathbb{C})$. Montrer que ses points extrémaux sont exactement les matrices unitaires :

$$\operatorname{Extr} \mathcal{B}_n = \mathcal{U}_n.$$

Solution proposée.

Soit U unitaire et x un vecteur non nul. Puisque U conserve la norme, $\frac{\|Ux\|}{\|x\|}$ vaut 1, donc U est sur la sphère unité, *a fortiori* dans $\mathcal{B}_n$.

Montrons alors que U est extrémale. Si on peut écrire $U = \frac{A+B}{2}$ où A et B sont dans $\mathcal{B}_n$, pour $x \neq 0$ on a alors

$$\|x\| = \|Ux\| = \left\| \frac{A+B}{2} x \right\| \leq \frac{\|Ax\| + \|Bx\|}{2} \leq \frac{\|A\|\,\|x\| + \|B\|\,\|x\|}{2} \leq \|x\|.$$

Le troisième cas d'égalité donne $\|A\| = \|B\| = 1$, le second cas d'égalité impose $\begin{cases} \|Ax\| = \|A\|\,\|x\| = \|x\| \\ \|Bx\| = \|B\|\,\|x\| = \|x\| \end{cases}$, et le premier cas d'égalité nous donne un réel $\lambda > 0$ tel que $Ax = \lambda Bx$; en prenant la norme il vient $\lambda = 1$. Finalement, $Ax = Bx$ pour tout x non nul, *i.e.* $A = B$, ce qui montre que U est extrémale.

Soit maintenant E extrémale dans $\mathcal{B}_n$. On veut E unitaire. Remarquons déjà que E est à la frontière de $\mathcal{B}_n$: E serait sinon le centre d'une boule ouverte incluse dans $\mathcal{B}_n$, et en prenant un diamètre $[ab]$ de cette boule E serait dans $]ab[$, contredisant son extrémalité. On a ainsi $\|E\| = 1$.

En introduisant la décomposition polaire $E = HU$ où H est hermitienne positive et U unitaire, il s'agit de montrer que $H = I_n$. Traduisons l'hypothèse $\|E\| = 1$:

$$1 = \|E\| = \sqrt{\rho(EE^*)} = \sqrt{\rho(HUU^*H)} = \sqrt{\rho(H^2)} = \rho(H).$$

Ainsi, si H n'est pas l'identité, une de ses valeurs propres est < 1, mettons $0 \leq \lambda < 1$. En diagonalisant $H = V \begin{pmatrix} \lambda & \\ & D \end{pmatrix} V^*$ où $\rho(D) \leq \rho(H) = 1$, et en écrivant $\lambda = \frac{1}{2} + \frac{2\lambda - 1}{2}$ comme un barycentre de réels

distincts dans $[-1,1]$, les matrices $\begin{pmatrix} 1 & \\ & D \end{pmatrix}$ et $\begin{pmatrix} 2\lambda-1 & \\ & D \end{pmatrix}$ sont dans $\mathcal{B}_n$; en multipliant par des matrices unitaires (U et V, au hasard), on reste dans $\mathcal{B}_n$, donc l'écriture

$$E = \frac{1}{2}\left[V\begin{pmatrix} 1 & \\ & D \end{pmatrix}V^*U\right] + \frac{1}{2}\left[V\begin{pmatrix} 2\lambda-1 & \\ & D \end{pmatrix}V^*U\right],$$

impose $1 = 2\lambda - 1$ par extrémalité de E, d'où contradiction.

13 Sur la densité de $\mathcal{O}_n(\mathbb{Q})$ dans $\mathcal{O}_n(\mathbb{R})$

- *En notant $\mathbb{S}_n(K)$ la sphère euclidienne de K^n, montrer que $\mathbb{S}_n(\mathbb{Q})$ est dense dans $\mathbb{S}_n(\mathbb{R})$.*
- *En déduire que $\mathcal{O}_n(\mathbb{Q})$ est dense dans $\mathcal{O}_n(\mathbb{R})$.*

Solution proposée.

• On raisonne par récurrence.

Pour $n = 1$, les deux sphères sont réduites à $\{\pm 1\}$.

Pour $n = 2$, on obtient le cercle unité, qui est paramétré par $(\cos\theta, \sin\theta)$. Le premier réflexe qui vient est d'approcher θ par une suite de rationnels, mais cela ne donne rien vu que les fonctions cos et sin envoient les rationnels un peu n'importe où (et n'ont aucune raison de stabiliser $\mathbb{Q}$).On pense alors à un autre paramétrage des matrices de rotations, le paramétrage rationnel (le nom est quand même bien choisi !) en fonction de l'arc moitié $\left(\frac{1-t^2}{1+t^2}, \frac{-2t}{1+t^2}\right)$.Sous cette forme, le problème est résolu en approchant t par des rationnels.

Supposons le résultat montré pour un $n \geq 2$ et soit $x \in \mathbb{S}_{n+1}(\mathbb{R})$. L'idée (faire $n = 3$ et dessiner une sphère peut aider à voir...) est de laisser tomber une coordonnée, d'approcher ce qui reste par récurrence, mettons $\frac{(x_1,\ldots,x_n)}{\sqrt{x_1^2+\ldots+x_n^2}} \simeq q \in \mathbb{S}_n(\mathbb{Q})$, puis de rajouter la coordonnée manquante x_0 en la remplaçant par une bonne approximation $q_0 \in \mathbb{Q}$. En observant que

$$(x_1, ..., x_n) \simeq \sqrt{x_1^2 + ... + x_n^2}\, q = \sqrt{1 - x_0^2}\, q,$$

la coordonnée supplémentaire doit vérifier

$$\begin{aligned} x &= (x_0, 0, ..., 0) + (0, x_1, ..., x_n) \\ &\simeq (q_0, 0, ..., 0) + \sqrt{1 - q_0^2}\,(0, q)\ ; \end{aligned}$$

il est alors judicieux d'approcher $\left(x_0, \sqrt{1-x_0^2}\right)$ par un rationnel $\left(q_0, \sqrt{1-q_0^2}\right)$, ce qui est toujours possible d'après le cas $n = 2$.

• Soit une matrice orthogonale réelle que l'on cherche à approcher par des matrices de $\mathcal{O}_n(\mathbb{Q})$. Une idée est que toute isométrie se décompose en produit de réflexions orthogonales, chacune étant entièrement déterminée selon la formule

$$r_a = \mathrm{Id} - 2\langle \cdot, a\rangle a$$

par l'un des deux vecteurs a unitaires situés sur son axe. En approximant $a \in \mathbb{S}_n(\mathbb{R})$ par un vecteur unitaire a' de $\mathbb{S}_n(\mathbb{Q})$ (cf. question précédente), on voit (à l'aide de la formule ci-dessus) que la nouvelle matrice de $r_{a'}$ dans la base canonique devient à coefficients rationnels et est encore une réflexion orthogonale, ce qui, en faisant le produit de ces approximations, donne une isométrie qui approche bien notre isométrie de départ.

Si l'on souhaite être précis, on évaluera la différence

$$\|r_a - r_b\| = 2\,\|\langle \cdot, a\rangle a - \langle \cdot, b\rangle b\| \leq 2\,\|\langle \cdot, a - b\rangle a\| + 2\,\|\langle \cdot, b\rangle (b - a)\|,$$

on remarquera que la norme triple de $\varphi(\cdot)\, a$ où a est un vecteur et φ une forme linéaire est donnée par le produit

$$\|\varphi(\cdot)\, a\| = \|\varphi\|\,\|a\|,$$

on observera que pour le cas d'un produit scalaire la norme de φ est donnée grâce à Cauchy-Schwarz par

$$\|\langle \cdot, a\rangle\| = \|a\|,$$

d'où, en approchant a par une suite (a_p) de vecteurs rationnels, l'évaluation

$$\left\|r_a - r_{a_p}\right\| \leq 2\,\|a - a_p\|\,\|a\| + 2\,\|a_p\|\,\|a - a_p\| \longrightarrow 0.$$

Enfin, puisque le produit est continu pour la norme triple, le produit des approximations ci-dessus fournit bien une approximation du produit de réflexions de départ qu'était notre isométrie.

14 Conjugaison et simplicité dans $\mathcal{SO}_3$

- *Qu'obtient-on en conjuguant une rotation r de $\mathbb{R}^3$ par une isométrie ω (*i.e. *que vaut $\omega r\omega^{-1}$) ?*
- *Soit G un sous-groupe non trivial de $\mathcal{SO}_3$ stable par conjugaison,* i.e. *tel que*

$$\forall g \in G,\ \forall r \in \mathcal{SO}_3,\ rgr^{-1} \in G.$$

Montrer que G vaut $\mathcal{SO}_3$ tout entier.

Solution proposée.

• Notons u un vecteur dirigeant l'axe de r et θ l'angle correspondant. $\mathcal{SO}_3$ étant un groupe, $\omega r\omega^{-1}$ est une rotation. Pour trouver son axe, on regarde ses points fixes :

$$\begin{aligned} x \in \operatorname{Fix}\left(\omega r\omega^{-1}\right) &\iff \omega r\omega^{-1}(x) = x \iff r\omega^{-1}(x) = \omega^{-1}(x) \\ &\iff \omega^{-1}(x) \in \operatorname{Fix} r \iff x \in \omega(\operatorname{Fix} r) = \mathbb{R}\omega(u). \end{aligned}$$

L'angle s'obtient au signe près en prenant la trace, laquelle est inchangée par conjugaison ; il reste donc à trouver le signe. Pour cela, on prend un vecteur y arbitraire (non lié à l'axe $\mathbb{R}\omega(u)$) et on regarde le signe du produit mixte

$$\begin{aligned} \left[y, \omega r\omega^{-1}(y), \omega(u)\right] &= \operatorname{Det}\left(y, \omega r\omega^{-1}(y), \omega(u)\right) \\ &= \det\omega \cdot \operatorname{Det}\left(\omega^{-1}(y), r\omega^{-1}(y), u\right) \\ &= 1 \cdot \left[\omega^{-1}(y), r\omega^{-1}(y), u\right] \\ &= [x, rx, u] \text{ avec } x = \omega^{-1}(y) \notin \mathbb{R}u \end{aligned}$$

qui est du signe de θ.

Finalement, conjuguer par ω revient juste à faire un changement de point de vue : l'action de la rotation est la même que celle de r, on tourne juste autour d'un autre axe.

• Les éléments de $\mathcal{SO}_3$ étant tous conjugués dans $\mathcal{SO}_3$ à une matrice de rotation standard $\begin{pmatrix} R_\theta & \\ & 1 \end{pmatrix}$, on peut supposer que G contient un élément g de la forme $\begin{pmatrix} R_\theta & \\ & 1 \end{pmatrix}$ avec θ non trivial. Il s'agit de récupérer n'importe quelle rotation d'angle φ pour φ arbitraire : en effet, pour obtenir une rotation autour d'un vecteur v à partir d'une rotation de même angle autour d'un vecteur u, il suffit (d'après le premier point) de conjuguer par une isométrie qui envoie u sur v et G est justement stable par conjugaison. Mieux : l'angle d'une rotation étant défini au signe près (l'axe a deux directions), il suffit de récupérer dans G une rotation d'angle φ pour chaque $\cos\varphi$ de $[-1, 1]$.

Pour obtenir, à partir de R_θ, une rotation d'angle φ quelconque, il ne suffit pas de conjuguer puisque cela ne change pas l'angle de la rotation. On va donc faire le produit de g par un de ses conjugués, puis récupérer l'angle φ en prenant la trace (laquelle vaut $1 + 2\cos\varphi$). Géométriquement, conjuguer g par r revient à modifier son axe Δ, donc prendre le produit reviendra à faire la composée de deux rotations d'axes distincts dont on cherche l'angle ; on prie pour que ce dernier décrive $[0, \pi]$ pour une modification de l'axe bien choisie.

Pour que les calculs se passent bien, on pense d'abord à conjuguer g par une matrice de même forme, *i.e.* du type $\begin{pmatrix} R_\nu & \\ & 1 \end{pmatrix}$: mais alors tout commute et le calcul se passe "trop" bien (la conjugaison n'a aucun effet). On pense alors à "mélanger" les deux blocs $R_?$ en conjugant par $r = \begin{pmatrix} 1 & \\ & R_v \end{pmatrix}$; géométriquement, cela revient à incliner l'axe (Oz) de g d'un angle ν dans le plan (Oyz). Pour mener le calcul, on explicite $R_\theta = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ et $R_\nu = \begin{pmatrix} \lambda & -\mu \\ \mu & \lambda \end{pmatrix}$ puis on y va la tête haute, en se souvenant que c'est la trace qui nous intéresse :

$$\begin{aligned} g\ rgr^{-1} &= \begin{pmatrix} a & -b & \\ b & a & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & \lambda & -\mu \\ & \mu & \lambda \end{pmatrix} \begin{pmatrix} a & -b & \\ b & a & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & \lambda & \mu \\ & -\mu & \lambda \end{pmatrix} \\ &= \begin{pmatrix} a & -b\lambda & b\mu \\ b & a\lambda & -a\mu \\ 0 & \mu & \lambda \end{pmatrix} \begin{pmatrix} a & -b\lambda & -b\mu \\ b & a\lambda & a\mu \\ 0 & -\mu & \lambda \end{pmatrix} \\ &= \begin{pmatrix} a^2 - b^2\lambda & ? & ? \\ ? & -b^2\lambda + a^2\lambda^2 + a\mu^2 & ? \\ ? & ? & a\mu^2 + \lambda^2 \end{pmatrix}. \end{aligned}$$

On récupère la trace que l'on met sous la forme d'un trinôme en λ, en utilisant les relations $a^2+b^2=1=\lambda^2+\mu^2$, puis on en déduit l'angle ν :

$$\begin{aligned}\operatorname{tr}\left(g\ rgr^{-1}\right) &= a^2-2\left(1-a^2\right)\lambda+a^2\lambda^2+2a\left(1-\lambda^2\right)+\lambda^2\\ &= (1-a)^2\lambda^2-2\left(1-a^2\right)\lambda+(1+a)^2-1\\ 1+2\cos\nu &= [(1-a)\lambda-(1+a)]^2-1\end{aligned}$$

(le premier qui a calculé un discriminant pour factoriser le trinôme a le droit de copier 100 fois "je regarde s'il n'y a pas de factorisation évidente avant de faire des calculs inutiles" !)

$$\cos\nu=\frac{[(1-a)\lambda-(1+a)]^2}{2}-1.$$

Puisque $a\leq 1$, le truc sous le carré croît en λ : ce dernier étant arbitraire dans $[-1,1]$, le truc sous le carré varie de -2 à $-2a$, donc parcourt $[-2,0]$ pour $a\leq 0$, donc $\cos\nu$ varie dans tout $[-1,1]$, *CQFD*. La condition $a\leq 0$ n'est pas bien dure à obtenir : si elle n'est pas vérifiée, cela signie que l'angle θ est dans la partie droite du cercle trigonométrique, et en itérant suffisamment (θ est non trivial !) on le fait sortir là où on veut.

Remarque. La première question illustre un principe très général en algèbre : conjuguer revient à changer de point de vue et ne change rien à l'"action" de l'objet étudié.

Le lecteur connait déjà la conjugaison dans $M_n(K)$ qui correspond à un changement de base : l'action de l'endomorphisme n'en est en rien modifiée.

Un autre exemple très simple est celui des permutations des ensembles finis : un tel ensemble E est en bijection avec $\{1,...,n\}$ via une bijection φ. On dispose alors d'une bijection $\left\{\begin{array}{ccc}\mathfrak{S}_E & \longrightarrow & \mathfrak{S}_n\\ f & \longmapsto & \varphi f\varphi^{-1}\end{array}\right.$ ramenant l'étude de $\mathfrak{S}_E$ à celle de $\mathfrak{S}_n$: ici, conjuguer par φ revient juste à mettre des étiquettes numérotées de 1 à n sur les objets de E, ce qui permet, au lieu d'étudier les permutations des objets de E, d'étudier les permutations de leurs étiquettes. C'est un changement de point de vue :-).

15 Ellipsoïde de John-Lowner

On appelle *ellipsoïde* de $\mathbb{R}^n$ toute boule unité pour la norme issue d'une forme quadratique définie positive. Si S est une matrice symtrique, on notera q_S la forme quadratique associée (dans la base canonique de $\mathbb{R}^n$) et E_S l'ellipsoïde associé à q_S.

Rappelons qu'un corps convexe est un compact convexe non vide.

On cherche à montrer qu'un corps convexe est inclus dans un unique ellipsoïde de volume minimal.

- *Montrer que le volume de l'ellipsoïde E_S est donné par*

$$\operatorname{vol}E_S=\frac{1}{\sqrt{\det S}}\operatorname{vol}\mathbb{B}$$

où $\mathbb{B}$ désigne la vraie boule unité euclidienne (correspondant à la forme quadratique standard).

- *Montrer la convexité de l'application* $v:\left\{\begin{array}{ccc}\mathcal{S}_n^{++} & \longrightarrow & \mathbb{R}\\ S & \longmapsto & \frac{1}{\sqrt{\det S}}\end{array}\right.$
- *En déduire le résultat.*

Solution proposée.

- Intuitivement, l'ellipsoïde E_S est une déformation $f(\mathbb{B})$ de la boule euclidienne, d'où

$$\operatorname{vol}E_S=\int_{E_S}1=\int_{f(\mathbb{B})}1=\int_{\mathbb{B}}|\det f|\ \text{ (en supposant } f \text{ de classe } C^1).$$

Précisons cela, en se rappelant l'existence de la racine carrée dans S_n^{++} :

$$\begin{aligned}x\in E_S &\iff q_S(x)\leq 1 \iff x^*Sx\leq 1 \iff x^*\sqrt{S}\sqrt{S}x\leq 1\\ &\iff \left(\sqrt{S}x\right)^*\sqrt{S}x\leq 1 \iff \sqrt{S}x\in\mathbb{B} \iff x\in\frac{1}{\sqrt{S}}\mathbb{B}.\end{aligned}$$

On peut donc prendre $f = \frac{1}{\sqrt{S}}$, d'où

$$\operatorname{vol} E_q = \int_{\mathbb{B}} \left|\det \frac{1}{\sqrt{S}}\right| = \frac{1}{\sqrt{\det S}} \int_{\mathbb{B}} 1 = \frac{1}{\sqrt{\det S}} \operatorname{vol} \mathbb{B},\ \textit{CQFD.}$$

• v fournit le volume de l'ellipsoïde associé à une forme quadratique donnée ; montrer sa convexité est donc intéréssant pour chercher ses minima (cf. point suivant).

Étant donnés deux matrices S et T dans $\mathcal{S}_n^{++}$, il suffit de vérifier que $v\left(\frac{S+T}{2}\right) \leq \frac{v(S)+v(T)}{2}$ puisque v est continue. Il convient déjà de remarquer que le milieu $\frac{S+T}{2}$ reste dans $\mathcal{S}_n^{++}$: la matrice $\frac{S+T}{2}$ est clairement symétrique, et définie positive comme somme de truc définis positifs. On veut donc

$$v\left(\frac{S+T}{2}\right) \overset{?}{\leq} \frac{v(S)+v(T)}{2},\ \textit{i.e.}\ \frac{1}{\sqrt{\det \frac{S+T}{2}}} \overset{?}{\leq} \frac{\frac{1}{2}}{\sqrt{\det S}} + \frac{\frac{1}{2}}{\sqrt{\det T}}.$$

L'inégalité à montrer étant invariante par conjugaison commune de S et T, on pense à écrire $S = P^*P$ et $T = PDP^*$ où P est inversible et D diagonale. On veut alors

$$\frac{1}{\sqrt{\det \frac{S+T}{2}}} \overset{?}{\leq} \frac{\frac{1}{2}}{\sqrt{\det I_n}} + \frac{\frac{1}{2}}{\sqrt{\det D}},\ \textit{i.e.}\ \frac{1}{\sqrt{\prod \frac{1+d_i}{2}}} \overset{?}{\leq} \frac{1}{2} + \frac{\frac{1}{2}}{\sqrt{\prod d_i}}.$$

C'est joli, mais pas autant que ça : trop d'interaction entre sommes et racines qui sont tout sauf compatibles. On va donc faire sauter la somme en essayant de montrer la log-convexité de v. Le résultat à montrer s'énonce alors

$$v\left(\frac{S+T}{2}\right) \overset{?}{\leq} \sqrt{v(S)\,v(T)} \iff \frac{1}{\sqrt{\prod \frac{1+d_i}{2}}} \overset{?}{\leq} \frac{1}{\sqrt{\sqrt{\prod d_i}}} \iff \prod \sqrt{d_i} \overset{?}{\leq} \prod \frac{1+d_i}{2}.$$

Il suffirait de montrer que $\sqrt{d_i} \leq \frac{1+d_i}{2}$ pour tout i. Or, si v est effectivement log-convexe, il doit l'être pour $n = 1$, donc il nous *faut* montrer $\sqrt{d_i} \leq \frac{1+d_i}{2}$, ce qui devient trivial une fois réécrit sous la forme

$$\left(1 - \sqrt{d_i}\right)^2 \geq 0.$$

• Soit K notre corps convexe. On cherche les minima de v sur les matrices S telles que $K \subset E_S$. Afin d'avoir l'existence d'un minimum, on va imposer une condition de plus pour avoir un domaine d'étude $\mathcal{D} \subset \mathcal{S}_n^{++}$ qui soit compact.

K étant bornée car compact, il y a une grosse boule qui l'englobe, mettons de volume V. Puisque l'on veut un truc qui englobe K et qui soit de volume minimal, ce volume minimal doit être $\leq V$. Cela nous assure que l'ellipsoïde cherché se trouve dans le domaine $\mathcal{D}$ défini par les conditions

$$K \subset E_S \text{ et } v(S) \leq V.$$

La première condition se réécrit $\forall k \in K,\ k^*Sk \leq 1$ et donc est clairement fermée. La seconde l'est également puisque v est continue. $\mathcal{D}$ étant par construction borné, il est compact comme souhaité et v y atteint son minimum.

Pour montrer qu'il est unique, il suffit d'invoquer la stricte convexité de v (qui est log-convexe). Mais encore faut-il que notre domaine $\mathcal{D}$ soit convexe ! Montrons ce dernier point, ce qui concluera.

$\mathcal{D}$ étant fermé, il suffit de montrer que $\mathcal{D}$ est stable par passage au milieu. Soit donc S et T dans $\mathcal{D}$: on veut que $\frac{S+T}{2} \in \mathcal{D}$. Pour vérifier la condition d'inclusion $K \subset E_{\frac{S+T}{2}}$, on pioche un élément $k \in K$: il est dans $E_S \cap E_T$, donc vérifie $k^*Sk \leq 1$ et $k^*Tk \leq 1$, d'où

$$k^* \frac{S+T}{2} k = \frac{k^*Sk + k^*Tk}{2} \leq \frac{1+1}{2} = 1,\ \textit{i.e.}\ k \in E_{\frac{S+T}{2}},\ \textit{CQFD.}$$

La condition de bornitude est par ailleurs triviale à vérifier par convexité de v :

$$v\left(\frac{S+T}{2}\right) \leq \frac{v(S)+v(T)}{2} \leq \frac{V+V}{2} = V.$$

Remarque. On pourrait se poser la même question en intervertissant les rôles contenant/contenu du corps convexe et de l'ellipsoïde : la réponse est alors négative. Il suffit de considérer un rectangle (non carré) et deux cercles inscrits maximaux à intérieur, chacun collant un petit côté du rectangle.