# Méthodes – Espaces préhilbertiens et euclidiens

## 1 ✓ Produit scalaire, et norme euclidienne

### 1.1 ✓ Montrer qu'une application est un produit scalaire

Nous nous attardons **uniquement** sur le caractère défini, mais les autres propriétés sont aussi à savoir démontrer.

#### 1.1.1 ✓ Produits scalaires sur $\mathbb{R}^n$ : somme de carrés et produits

Ils sont souvent définis concrètement :

$$\varphi_1 : \begin{cases} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((x_1, y_1), (x_2, x_2)) & \mapsto & x_1x_2 + 2y_1y_2 - x_1y_2 - x_2y_1 \end{cases},$$

$$\varphi_2 : \begin{cases} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((x_1, y_1), (x_2, x_2)) & \mapsto & 12x_1x_2 + x_2y_1 + x_1y_2 + 9y_1y_2 \end{cases}.$$

S'ils sont définis abstraitement, alors il y a fort à parier qu'ils sont définis à partir d'un endomorphisme autoadjoint, et dans ce cas ses valeurs propres sont nécessaires pour montrer qu'il s'agit d'un produit scalaire : nous n'en parlerons pas ici, et vous renvoyons à la section 4.3 (plus précisément l'exercice 24) → page 29 pour les détails.

Montrer la bilinéarité est une affaire de routine. Pour la symétrie, faites bien attention : il faut que le coefficient en facteur de $x_1y_2$ (par exemple) soit le même que celui en facteur de $x_2y_1$, et de même s'il y a d'autres termes « croisés » (ce qui n'est pas le cas sur $\mathbb{R}^2$). Par exemple, l'application :

$$\psi : \begin{cases} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((x_1, y_1), (x_2, x_2)) & \mapsto & 3x_1x_2 - 2y_1y_2 + 2x_1y_2 - x_2y_1 \end{cases}$$

n'est pas bilinéaire, puisque par exemple : $\psi((1{,}0), (0{,}1)) = 2$, alors que $\psi((0{,}1), (1{,}0)) = -1$, si bien que $\psi((1{,}0), (0{,}1)) \neq \psi((0{,}1), (1{,}0))$.

Mais la vraie difficulté est pour la positivité et le caractère défini, qui se traitent à peu près en même temps. Pour cela, vous utilisez le résultat suivant, **à citer tel quel quand vous en avez besoin** :

| Une somme de réels positifs est nulle si et seulement si chaque terme de la somme est nul. |
|---|

Cela tombe bien : quand vous évaluez ces formes bilinéaires en deux fois le même vecteur, il apparaît une somme de carrés, qui sont donc positifs. Mais le problème est qu'il y a d'autres termes dont on ne connaît pas le signe ! Par exemple, pour $\varphi_1$ définie ci-dessus :

$$\forall (x, y) \in \mathbb{R}^2, \quad \varphi_1((x, y), (x, y)) = x^2 + 2y^2 - 2xy.$$

On ne connaît pas le signe de $x$ et de $y$, donc on ne connaît pas le signe de $-2xy$. On ne peut pas en déduire que chaque terme est nul.

**La méthode pour y remédier est : ajouter et enlever un carré convenable pour pouvoir « éliminer » le terme problématique en reconnaissant une identité remarquable** :

$$ax^2 + bxy = a\left(x^2 + \frac{b}{a}xy\right) = a\left(x^2 + 2 \cdot \frac{b}{2a}xy + \left(\frac{by}{2a}\right)^2 - \left(\frac{by}{2a}\right)^2\right) = a\left(x + \frac{b}{2a}y\right)^2 - \frac{b^2y^2}{4a}.$$

Il n'y a plus le terme problématique $bxy$ : il n'y a plus que des carrés, dont on connaît le signe. On recommence au besoin (si on n'est pas sur $\mathbb{R}^2$) jusqu'à n'avoir plus que des carrés.

**Exemple 1.** Montrons que $\varphi_1$, définie ci-dessus, est définie positive. Appliquer cette méthode donne :

$$\forall (x,y) \in \mathbb{R}^2, \quad \varphi_1((x,y),(x,y)) = x^2 - 2xy + 2y^2 = x^2 - 2xy + y^2 + y^2 = (x+y)^2 + y^2 \geqslant 0,$$

donc $\varphi_1$ est positive. De plus, une somme de termes **positifs** est nulle si et seulement si chaque terme est nul, donc l'égalité $\varphi_1((x,y),(x,y)) = 0$ implique $(x+y)^2 = 0$ et $y^2 = 0$. De là on déduit immédiatement $x = y = 0$, donc $\varphi_1((x,y),(x,y)) = 0$ si et seulement si $(x,y) = (0,0)$. L'application $\varphi_1$ est définie positive, en plus d'être une forme bilinéaire symétrique : c'est donc un produit scalaire sur $\mathbb{R}^2$.

**Exercice 1.** Montrer de même que $\varphi_2$ est un produit scalaire sur $\mathbb{R}^2$.

### 1.1.2 ✓ Produits scalaires sur $\mathrm{M}_{n,1}(\mathbb{R})$

Les produits scalaires sur $\mathrm{M}_{n,1}(\mathbb{R})$ sont toujours, implicitement ou non, de la forme :

$$(X,Y) \mapsto X^\top AY$$

avec $A$ une matrice symétrique. Voir exercices 21 et 24 de ce document pour en savoir plus. → page 26

Si vous devez montrer que c'est un produit scalaire **avec $n$ petit et $A$ explicite**, alors la méthode est la même que pour les produits scalaires sur $\mathbb{R}^n$, car tout produit scalaire sur $\mathbb{R}^n$ peut être écrit sous la forme $X^\top AY$ quitte à identifier $\mathbb{R}^n$ et $\mathrm{M}_{n,1}(\mathbb{R})$. Par exemple, si $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, alors pour tous $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ :

$$X^\top AY = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} ay_1 + by_2 \\ by_1 + cy_2 \end{pmatrix} = ax_1y_1 + b\,(x_1y_2 + x_2y_1) + cx_2y_2.$$

**Faites le calcul vous-mêmes pour vous en convaincre**. Comme pour les produits scalaires sur $\mathbb{R}^n$, on démontre le caractère défini en faisant apparaître une somme de carrés.

En revanche, si ce n'est pas explicite, il y a deux complications :

— **la symétrie** : elle n'est vraie que si $A$ est symétrique, et pour la justifier sans calcul vous suivez cette démarche : comme $X^\top AY$ est une matrice d'ordre 1 (identifiée à un réel), elle est nécessairement symétrique (on a $(a)^\top = (a)$ pour tout $(a) \in \mathrm{M}_1(\mathbb{R})$), donc :

$$X^\top AY = \left(X^\top AY\right)^\top = Y^\top A^\top \left(X^\top\right)^\top \overset{[A \in \mathrm{S}_n(\mathbb{R})]}{=} Y^\top AX,$$

donc $(X,Y) \mapsto X^\top AY$ est symétrique ;

— **la positivité et le caractère défini** : en général, c'est lié à un problème de valeurs propres, comme nous le disons ci-dessous ; mais sinon, ne perdez pas de vue que $X^\top AX$ *est le produit scalaire usuel de $X$ par $AX$* : selon les hypothèses, vous pourrez peut-être le simplifier ou voir à quelle condition sur $X$ il est nul.

Nous parlons plus en détails des produits scalaires sur $\mathrm{M}_{n,1}(\mathbb{R})$ dans l'exercice 24 de ce document : → page 29 $(X,Y) \mapsto X^\top AY$ est un produit scalaire si et seulement si $A$ est symétrique et définie positive. **C'est un exercice très important**. C'est votre principal recours si $A$ est définie abstraitement.

### 1.1.3 ✓ Cas particulier fréquent dans les espaces de fonctions

Les produits scalaires sur les espaces de fonctions sont souvent définis avec une intégrale :

$$(f,g) \mapsto \int_I fg.$$

Lorsque les fonctions ont une régularité (classe $C^1$), voire lorsqu'elles vérifient une équation différentielle, on peut privilégier des intégrales où apparaissent leurs dérivées :

$$\varphi_1 : \begin{cases} C^1([0,1]) \times C^1([0,1]) & \to \mathbb{R} \\ (f,g) & \mapsto \displaystyle\int_0^1 (fg + f'g') \end{cases}, \quad \varphi_2 : \begin{cases} C^1([0,1]) \times C^1([0,1]) & \to \mathbb{R} \\ (f,g) & \mapsto f(0)g(0) + \displaystyle\int_0^1 f'g' \end{cases}.$$

Le caractère défini se démontre semblablement dans tous les cas : on se retrouve avec une fonction (ou une somme de fonctions) à valeurs réelles et au carré. Elle est donc positive. Si une fonction **positive** et **continue** est **d'intégrale nulle**, alors elle est elle-même nulle par séparation de l'intégrale.

Citez très soigneusement ces hypothèses, on vous note PRÉCISÉMENT là-dessus !

Si ce n'est pas $f$ qu'on intègre mais $f'$, $f''$, etc., alors la continuité de ces dérivées découle de la classe de $f$ : la vérifier soigneusement. Si vous en déduisez que $f' = 0$ ou $f'' = 0$ par exemple, ce n'est pas encore suffisant pour conclure : intégrez ces relations pour en arriver à $f$, *sans oublier les constantes d'intégration*, et utilisez les autres quantités du produit scalaire pour en déduire que $f = 0$.

**Exercice 2.** Montrer que les deux applications ci-dessus sont des produits scalaires.

Si ces produits scalaires sont définis sur $\mathbb{R}[X]$ ou $\mathbb{R}_n[X]$, **on n'oublie pas de passer des applications polynomiales aux polynômes** : le fait que $P(x) = 0$ pour tout $x \in I$ signifie que $x$ est une racine de $P$ pour tout $x \in I$. Cela fournit une infinité de racines de $P$, donc $P = 0_{\mathbb{R}[X]}$. **Cette étape n'est pas facultative**.

### 1.1.4 ✓ Produits scalaires dans $\mathbb{R}_n[X]$ et $\mathbb{R}[X]$ : raisonnements sur les racines

Outre le cas des produits scalaires avec intégrales (voir ci-dessus), vous aurez souvent à étudier des produits scalaires sur $\mathbb{R}[X]$ ou $\mathbb{R}_n[X]$ qui s'obtiennent en sommant différentes évaluations des polynômes, ou de leurs dérivées successives. Exemples :

$$\varphi_1 : \begin{cases} \mathbb{R}_n[X] \times \mathbb{R}_n[X] & \to \mathbb{R} \\ (P,Q) & \mapsto \sum\limits_{i=0}^{n} P(i)Q(i) \end{cases}, \quad \varphi_2 : \begin{cases} \mathbb{R}[X] \times \mathbb{R}[X] & \to \mathbb{R} \\ (P,Q) & \mapsto \sum\limits_{i=0}^{+\infty} P^{(i)}(2)Q^{(i)}(2) \end{cases},$$

$$\varphi_3 : \begin{cases} \mathbb{R}_2[X] \times \mathbb{R}_2[X] & \to \mathbb{R} \\ (P,Q) & \mapsto P(0)Q(0) + P'(0)Q'(0) + P(1)Q(1) \end{cases}.$$

Quand les produits scalaires sont définis à l'aide des dérivées successives, il n'est pas rare d'avoir une somme infinie : il n'y a pourtant aucun problème de convergence, car les dérivées sont nulles à partir d'un certain rang et la somme est en vérité finie.

Pour démontrer le caractère défini, on utilise d'abord le fait qu'une somme de termes positifs soit nulle si et seulement si chaque terme est nul, nous ramenant à plusieurs égalités de la forme : $P^{(k)}(a) = 0$. Ces égalités permettent d'invoquer trois résultats essentiels sur les racines :

- $P(a) = 0$ si et seulement si $a$ est racine de $P$ ;
- plus généralement, $P(a) = P'(a) = \cdots = P^{(k)}(a) = 0$ si et seulement si $a$ est racine de $P$ d'ordre de multiplicité au moins $k+1$ (et non $k$, attention : regardez pour $k = 1$ ce que cela donnerait) ;
- seul le polynôme nul a strictement plus de racines (comptées avec multiplicités) que son degré ;

La connaissance du degré des polynômes est essentielle. Si l'on est sur $\mathbb{R}[X]$ au lieu de $\mathbb{R}_n[X]$, on n'a pas de borne sur le degré. Néanmoins, dans ces cas-là, nous arrivons souvent à démontrer qu'il existe une infinité de racines (ou une racine d'ordre arbitrairement élevé), ce qui permet à coup sûr de dépasser le degré.

**Exercice 3.** Montrer que les trois applications proposées ci-dessus sont des produits scalaires. Raisonnez sur les racines y compris pour $\varphi_3$, ou vous ne tirerez aucun enseignement de l'exercice.

En revanche, vous OUBLIEZ cette méthode lorsque les évaluations des dérivées de $P$ n'ont pas de rapport avec les évaluations de $P$. Par exemple, pour vérifier le caractère défini du produit scalaire défini sur $\mathbb{R}_1[X]$ par $(P,Q) \mapsto P(0)Q(0) + P'(1)Q'(1)$, le fait $P'(1) = 0$ n'apporterait AUCUNE information sur les racines de $P$, et certainement pas le fait que 1 soit racine de $P$ (certes, cela nous dit que 1 est racine de $P'$, mais cela n'implique pas que c'est une racine de $P$). Plus généralement, pour que l'information $P^{(k)}(a) = 0$ ait une influence sur les racines, il faut AUSSI $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$ !
Dans ce cas-là, vous devez utiliser d'autres choses sur $P^{(k)}$ : son degré, son coefficient dominant ou constant, etc.

**Exercice 4.** Les méthodes de la section ne s'appliquent pas à l'application $(P,Q) \mapsto \sum\limits_{k=0}^{+\infty} P^{(k)}(k)Q^{(k)}(k)$ (chaque dérivée successive est évaluée en un autre réel). Démontrer malgré tout le caractère défini en explicitant le terme de la somme correspondant à $k = \deg(P)$ (si $P$ est un polynôme non nul).

## 1.2 ✓ Calculs de produits scalaires et (surtout) de normes euclidiennes

Dans cette section je ne parle pas de calculs « savants » avec les normes euclidiennes :

— pour les inégalités entre normes et produits scalaires, voir section 1.3 ; → page 6

— pour le calcul des normes impliquées dans les calculs de distance, voir section 3. → page 20

Pour les inégalités sur les normes utilisant des endomorphismes autoadjoints ou antisymétriques, vous tirerez plus d'enseignements des sections 2.3 et 4. Ici je ne parle que : → page 26

— des conseils généraux sur les calculs ;

— des calculs de normes dans $\mathrm{M}_{n,1}(\mathbb{R})$ (avec $X^\top Y$) ou $\mathrm{M}_n(\mathbb{R})$ (avec $\mathrm{tr}(A^\top B)$) ;

— des inégalités basiques sur les normes (valant pour tous vecteurs, sans trop d'hypothèses).

Lorsqu'on nous demande de calculer une norme, nous vous recommandons de plutôt calculer **la norme au carré**. C'est plus pratique afin d'utiliser :

— les identités remarquables suivantes :

$$\|\vec{x} \pm \vec{y}\|^2 = \|\vec{x}\|^2 \pm 2\langle \vec{x}, \vec{y}\rangle + \|\vec{y}\|^2, \quad \|\alpha\vec{x} + \beta\vec{y}\|^2 = \alpha^2\|\vec{x}\|^2 + 2\alpha\beta\langle \vec{x}, \vec{y}\rangle + \beta^2\|\vec{y}\|^2;$$

— le théorème de Pythagore s'il y a des relations d'orthogonalité : **parfois, un dessin vous fera réaliser des relations d'orthogonalité que vous n'aviez pas aperçues**.

Le seul cas où je vous *déconseille* cette élévation au carré est lorsque *vous avez à majorer la norme d'une somme de nombreux vecteurs*. Dans ce cas, même avec l'inégalité de Cauchy-Schwarz, majorer les produits scalaires apparaissant dans le développement est pénible, alors qu'avec l'inégalité triangulaire c'est immédiat.

**Exercice 5.** Montrer :

$$\forall(\vec{a},\vec{b},\vec{c},\vec{d}) \in E^4, \quad \|\vec{b}-\vec{a}\|^2 + \|\vec{c}-\vec{b}\|^2 + \|\vec{d}-\vec{c}\|^2 + \|\vec{a}-\vec{d}\|^2 = \|\vec{c}-\vec{a}\|^2 + \|\vec{d}-\vec{b}\|^2 + \|\vec{a}-\vec{b}+\vec{c}-\vec{d}\|^2.$$

### 1.2.1 ✓ Produit scalaire et norme euclidienne sur $\mathrm{M}_{n,1}(\mathbb{R})$ ou $\mathrm{M}_n(\mathbb{R})$

Pour la norme euclidienne sur $\mathrm{M}_{n,1}(\mathbb{R})$, je veux simplement attirer l'attention sur le fait que... Souvent vous ne la reconnaissez pas. Et c'est ainsi que vous bloquez bêtement dès que vous êtes en présence d'un produit matriciel tel que $X^\top Y$ ou $X^\top X$.

> On a tout simplement $X^\top Y = \sum\limits_{i=1}^{n} x_i y_i$ et $X^\top X = \sum\limits_{i=1}^{n} x_i^2$ : c'est le même produit scalaire que sur $\mathbb{R}^n$ !

Souvent dans ces cas-là, après avoir simplifié du mieux possible en développant : *il n'y a pas d'autre choix que d'écrire explicitement les coordonnées pour poursuivre.*

Les calculs de normes sur $\mathrm{M}_n(\mathbb{R})$ sont plus subtils. Pour rappel, il y a deux façons d'écrire le produit scalaire et la norme euclidienne usuels :

| | Version abstraite | Version concrète |
|---|---|---|
| $\langle A, B\rangle$ | $\operatorname{tr}(A^\top B)$ | $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}b_{i,j}$ |
| $\lVert A\rVert^2$ | $\operatorname{tr}(A^\top A)$ | $\sum_{i=1}^{n}\sum_{j=1}^{n} (a_{i,j})^2$ |

Comme on le voit, ils ont une expression explicite très proche du produit scalaire usuel sur $\mathbb{R}^n$ : c'est simplement la somme des produits des coefficients !

Vous allez privilégier la « version concrète » pour le calcul **explicite** (matrices dont on connaît les coefficients), et la « version abstraite » pour des matrices **dont la transposée vérifie des propriétés spéciales**. En particulier :

— si $A$ est symétrique ou antisymétrique (vu que $A^\top = \pm A$ dans ce cas) ;

— si $A$ est orthogonale (vu que $A^\top = A^{-1}$ dans ce cas).

Plus particulièrement dans le cas **symétrique**, vous pouvez de plus simplifier le calcul **en vous souvenant que** $A = PDP^{-1}$ **avec** $D$ **diagonale et** $P$ **orthogonale**. C'est le théorème spectral.

**Exemple 2.** Si $A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \\ 7 & -8 & 9 \end{pmatrix}$, alors :

$$\lVert A\rVert^2 = 1^2+(-2)^2+3^2+(-4)^2+5^2+(-6)^2+7^2+(-8)^2+9^2 = \sum_{k=1}^{9} k^2 = \frac{9\cdot 10\cdot 19}{6} = 3\cdot 5\cdot 19 = 15\cdot 19 = 285.$$

Cela se fait de tête : $15\cdot 19 = 15\cdot 20 - 15$. Vous auriez bien sûr songé à simplifier la fraction avant de multiplier comme des cochons, n'est-ce pas ? Remarquez que je ne m'embête pas à calculer $A^\top A$ puis sa trace.

**Mise en garde 1.** Lorsque vous faites du calcul de produits scalaires ou de normes avec des matrices explicites :

| NE CALCULEZ PAS $A^\top B$ OU $A^\top A$ POUR EN PRENDRE ENSUITE LA TRACE ! VOUS PERDEZ DU TEMPS ! |
|---|

Perte de temps d'autant plus bête que vous calculez $n^2$ coefficients pour finalement n'en retenir que $n$.

### 1.2.2 ✓ Inégalités basiques sur les normes

Lorsque vous devez démontrer une inégalité avec des normes, souvent :

— soit vous avez **un produit de normes dans le majorant** : songez à l'inégalité de Cauchy-Schwarz ;

— soit vous avez **une somme de normes, ou une norme de sommes** : songez à l'inégalité triangulaire.

Si vous n'êtes pas dans ces cas-là, essayez de développer une norme au carré pour faire apparaître des simplifications. Les produits scalaires qui apparaissent dans le développement s'encadrent avec l'inégalité de Cauchy-Schwarz (**pas systématiquement** : analysez prudemment, au brouillon, par quoi vous devez majorer ces produits scalaires, pour voir s'il n'y a pas une autre approche plus maligne).

S'il apparaît également un endomorphisme dans l'inégalité (isométrie, projecteur orthogonal, etc.), les conseils de la page 9 (*Quand penser à l'inégalité de Cauchy-Schwarz : cas abstrait*) restent valables ici. → page 9

Si l'on vous demande de comparer des normes de différences ou sommes, par exemple :

$$\forall(\vec{x},\vec{y},\vec{z}) \in E^3, \quad \|\vec{x}-\vec{z}\| \leqslant \|\vec{x}-\vec{y}\| + \|\vec{y}-\vec{z}\|.$$

Vous aurez souvent besoin de « l'astuce » d'ajouter et soustraire le vecteur qui vous manque dans le membre de gauche, afin de faire apparaître celui de droite. Puis vous utilisez l'inégalité triangulaire. Ainsi :

$$\forall(\vec{x},\vec{y},\vec{z}) \in E^3, \quad \|\vec{x}-\vec{z}\| = \|(\vec{x}-\vec{y}) + (\vec{y}-\vec{z})\| \leqslant \|\vec{x}-\vec{y}\| + \|\vec{y}-\vec{z}\|.$$

**Exemple 3.** Soit $(\vec{x},\vec{y},\vec{z}) \in E^3$. Nous allons combiner plusieurs des principes plus haut pour démontrer :

$$\|\vec{x}-\vec{z}\|^2 \leqslant 2\left(\|\vec{x}-\vec{y}\|^2 + \|\vec{y}-\vec{z}\|^2\right).$$

Comme plus haut nous allons faire apparaître $\vec{x}-\vec{y}$ et $\vec{y}-\vec{z}$ dans le membre de gauche, puis nous allons développer le carré de la norme à l'aide d'une « identité remarquable ». Notons qu'on veut avoir $\|\vec{x}-\vec{y}\|^2$ et $\|\vec{y}-\vec{z}\|^2$ dans le membre de droite : ce serait donc contre-productif de développer ces deux normes au carré, et nous les laisserons ainsi. Ceci étant dit :

$$\begin{aligned}
\|\vec{x}-\vec{z}\|^2 = \|(\vec{x}-\vec{y}) + (\vec{y}-\vec{z})\|^2 &= \|\vec{x}-\vec{y}\|^2 + 2\langle \vec{x}-\vec{y}, \vec{y}-\vec{z}\rangle + \|\vec{y}-\vec{z}\|^2 \\
&\leqslant \|\vec{x}-\vec{y}\|^2 + 2\|\vec{x}-\vec{y}\|\|\vec{y}-\vec{z}\| + \|\vec{y}-\vec{z}\|^2 && \text{(inégalité de Cauchy-Schwarz)} \\
&\leqslant \|\vec{x}-\vec{y}\|^2 + \left(\|\vec{x}-\vec{y}\|^2 + \|\vec{y}-\vec{z}\|^2\right) + \|\vec{y}-\vec{z}\|^2. && \text{(inégalité } 2ab \leqslant a^2+b^2\text{)}
\end{aligned}$$

L'inégalité $2ab \leqslant a^2+b^2$ se démontre à partir de : $(a-b)^2 \geqslant 0$, et en développant. Il reste à regrouper les termes. Notons qu'on peut aussi démontrer cette inégalité avec l'identité du parallélogramme : comment ?

**Exercice 6.** Démontrer de même : $\forall(\vec{x},\vec{y}) \in E^2$, $2 + \|\vec{x}+\vec{y}\|^2 \leqslant 2\left(1+\|\vec{x}\|^2\right)\left(1+\|\vec{y}\|^2\right)$.

## 1.3 ✓ Reconnaître et utiliser l'inégalité de Cauchy-Schwarz

Rappelons l'inégalité de Cauchy-Schwarz : $\forall(\vec{x},\vec{y}) \in E^2$, $|\langle\vec{x},\vec{y}\rangle| \leqslant \|\vec{x}\| \cdot \|\vec{y}\|$. Dans le cas de $\mathbb{R}^n$ et des espaces de fonctions, cela donne :

$$\forall(a_1,\dots,a_n) \in \mathbb{R}^n,\ \forall(b_1,\dots,b_n) \in \mathbb{R}^n, \quad \left|\sum_{i=1}^n a_i b_i\right| \leqslant \sqrt{\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2}, \tag{1}$$

$$\forall f,g \in \mathrm{L}^2(I,\mathbb{R}) \cap \mathrm{C}^0(I,\mathbb{R}), \quad \left|\int_I fg\right| \leqslant \sqrt{\int_I f^2 \int_I g^2}. \tag{2}$$

On note que dans chacun des cas, il apparaît un carré dans la définition des normes, et donc dans le majorant de l'inégalité de Cauchy-Schwarz. Cela va nous aider à savoir quand la reconnaître dans un exercice, et comment déterminer les deux vecteurs à utiliser.

### 1.3.1 Quand penser à l'inégalité de Cauchy-Schwarz : cas concret

S'il n'apparaît pas explicitement des normes ou produits scalaires, on y pense **lorsqu'il apparaît une inégalité entre des sommes ou des intégrales, avec la présence de carrés ou de racines carrées** (à l'intérieur ou à l'extérieur des sommes ou intégrales, peu importe). Dans le cas des matrices, on remplace les sommes ou intégrales par des traces. Alors, pour trouver *comment* l'appliquer :

1. Identifiez le produit scalaire usuel ressemblant le plus aux forces en présence.
2. Dans votre analyse *au brouillon* : comparez ce que vous voulez et ce que vous savez, et « identifiez ».
3. Dans votre synthèse au propre : écrivez l'inégalité de Cauchy-Schwarz avec les vecteurs trouvés au brouillon.

**Un conseil très important pour « l'identification » : ne raisonnez pas sur le produit scalaire (minorant), mais les normes (majorant).** Pour une raison très simple : trouver les bons réels $x$, $x'$ vérifiant $x^2 = y$ et $x'^2 = y'$ est immédiat : on a $x = \pm\sqrt{y}$ et $x' = \pm\sqrt{y'}$. Trouver des réels $x$, $x'$ vérifiant $xx' = z$ est voué à l'échec : il y a trop de solutions. Or c'est le genre d'identification que vous serez amenés à faire, selon que vous regardiez le minorant ou le majorant.

**Exemple 4.** On veut montrer que pour tous réels $x_1, \dots, x_n$, on a :

$$\left|\sum_{k=1}^{n} k^2 x_k^3\right| \leqslant \sqrt{\left(\sum_{k=1}^{n} k^2 x_k^2\right)\left(\sum_{k=1}^{n} k^2 x_k^4\right)}.$$

Au vu de la majoration à démontrer, il ne fait aucun doute qu'il s'agit d'utiliser l'inégalité de Cauchy-Schwarz avec des vecteurs $\vec{a} = (a_1, \dots, a_n)$ et $\vec{b} = (b_1, \dots, b_n)$ de $\mathbb{R}^n$, mais lesquels ? Il faudrait $a_k b_k = k^2 x_k^3$, mais il y a plein de façons de définir $a_k$ et $b_k$ pour que cette égalité soit vérifiée : prend-on $(a_k, b_k) = (k^2 x_k, x_k^2)$ ? ou $(a_k, b_k) = (kx_k, kx_k)$ ? ou $(a_k, b_k) = (x_k^2, k^2 x_k)$ ? Et je suis loin d'embrasser toutes les possibilités.

En revanche, si l'on regarde les normes, on voit qu'il n'y a pas ambiguïté : on voudrait (1) avec $a_k^2 = k^2 x_k^2$ et $b_k^2 = k^2 x_k^4$. Cela incite à prendre $a_k = kx_k$ et $b_k = kx_k^2$ pour tout $k \in [\![1, n]\!]$. Vérifions ce que donne l'inégalité de Cauchy-Schwarz dans ce cas :

$$\left|\sum_{k=1}^{n} a_k b_k\right| \leqslant \sqrt{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2} \iff \left|\sum_{k=1}^{n} kx_k \cdot kx_k^2\right| \leqslant \sqrt{\sum_{k=1}^{n} k^2 x_k^2 \sum_{k=1}^{n} k^2 x_k^4} \iff \left|\sum_{k=1}^{n} k^2 x_k^3\right| \leqslant \sqrt{\left(\sum_{k=1}^{n} k^2 x_k^2\right)\left(\sum_{k=1}^{n} k^2 x_k^4\right)},$$

d'où le résultat. Voyez comme il est plus facile d'identifier avec le majorant.

**Exemple 5.** On veut montrer que pour tous réels $x_1, \dots, x_n$ strictement positifs, on a :

$$n^2 \leqslant \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \frac{1}{x_i}.$$

On fait la même « identification » que ci-dessus : si cette inégalité est celle de Cauchy-Schwarz avec des vecteurs convenables de $\mathbb{R}^n$, il faudrait $a_i^2 = x_i$ et $b_i^2 = \frac{1}{x_i}$. Ceci conduit au choix naturel $a_i = \sqrt{x_i}$ et $b_i = \sqrt{\frac{1}{x_i}}$ (c'est ici que la positivité intervient). Vérifions :

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2 \iff \left(\sum_{i=1}^{n} \sqrt{\frac{x_i}{x_i}}\right)^2 \leqslant \sum_{i=1}^{n} (\sqrt{x_i})^2 \sum_{i=1}^{n} \left(\sqrt{\frac{1}{x_i}}\right)^2 \iff \left(\sum_{i=1}^{n} 1\right)^2 \leqslant \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \frac{1}{x_i},$$

et comme $\displaystyle\sum_{i=1}^{n} 1 = n$ on a le résultat attendu.

**Exemple 6.** On veut montrer que pour toute fonction $f : [0,1] \to \mathbb{R}$ continue et positive, on a :

$$\int_0^1 (f(x))^2 \mathrm{d}x \leqslant \sqrt{\int_0^1 f(x)e^x \mathrm{d}x \cdot \int_0^1 f(x)e^{-x} \mathrm{d}x}.$$

Cette fois-ci, la présence de la racine carrée sur un produit d'intégrales fait penser à l'inégalité de Cauchy-Schwarz sur les espaces de fonctions (inégalité (2)). Mais lesquelles ? Cherchons quelles fonctions $\varphi$ et $\psi$ (je change les noms pour ne pas avoir de doublon avec $f$) donneraient : $\displaystyle\int_0^1 f(x)e^x \mathrm{d}x = \int_0^1 (\varphi(x))^2 \mathrm{d}x$, et : $\displaystyle\int_0^1 f(x)e^{-x} \mathrm{d}x = \int_0^1 (\psi(x))^2 \mathrm{d}x$. Il semble que des choix naturels seraient $\varphi(x)^2 = f(x)e^x$ et $\psi(x)^2 = f(x)e^{-x}$. Posons donc :

$$\forall x \in [0,1], \quad \varphi(x) = \sqrt{f(x)}e^{x/2}, \quad \psi(x) = \sqrt{f(x)}e^{-x/2}, \quad \text{(c'est ici qu'apparaît la positivité de } f\text{)}$$

Ce sont bien des fonctions continues sur [0,1]. L'inégalité de Cauchy-Schwarz donne :

$$\left|\int_0^1 \varphi(x)\psi(x)\mathrm{d}x\right| \leqslant \sqrt{\int_0^1 (\varphi(x))^2\mathrm{d}x \cdot \int_0^1 (\psi(x))^2\mathrm{d}x} \iff \int_0^1 f(x)\cancel{e^{x/2}}f(x)\cancel{e^{-x/2}}\mathrm{d}x \leqslant \sqrt{\int_0^1 f(x)e^{x}\mathrm{d}x \cdot \int_0^1 f(x)e^{-x}\mathrm{d}x},$$

d'où le résultat.

**Mise en garde 2.** Ces identifications ne sont pas rigoureuses DU TOUT et fausses en général. C'est pour trouver des idées AU BROUILLON que vous les écrivez, mais elles ne figurent pas dans une copie.

**Et s'il n'apparaît qu'une seule norme ?** Souvent il n'apparaît qu'une seule norme dans le majorant, parce que l'autre norme a déjà été simplifiée dans l'énoncé. À moins que vous n'ayez du flair, voici comment procéder :

1. En inspectant la norme dans le majorant, vous trouvez l'un des deux vecteurs à utiliser (disons $\vec{x}$).
2. Pour trouver l'autre vecteur, demandez-vous : quel vecteur donne un produit scalaire avec $\vec{x}$ égal au minorant ?

Nous donnons deux exemples.

**Exemple 7.** On veut montrer que pour tous réels $x_1, \ldots, x_n$, on a :

$$\left(\sum_{k=1}^{n} kx_k\right)^2 \leqslant \frac{n(n+1)(2n+1)}{6}\sum_{k=1}^{n} x_k^2.$$

On peut effectivement penser à l'inégalité de Cauchy-Schwarz : c'est une inégalité entre sommes avec présence de carrés. Mais il n'apparaît *a priori* qu'une seule somme, donc une seule norme, dans le majorant. Pas grave : notons $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, et voyons quel vecteur $\vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ donnerait $\langle \vec{x}, \vec{y}\rangle = \sum\limits_{k=1}^{n} kx_k$. Comme $\langle \vec{x}, \vec{y}\rangle = \sum\limits_{k=1}^{n} x_k y_k$, un choix naturel serait : $\forall k \in [\![1, n]\!]$, $y_k = k$. Vérifions que ce choix de $\vec{y}$ fonctionne :

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leqslant \sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} y_k^2 \iff \left(\sum_{k=1}^{n} kx_k\right)^2 \leqslant \sum_{k=1}^{n} x_k^2 \cdot \sum_{k=1}^{n} k^2,$$

et on sait que : $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$, d'où le résultat.

**Exemple 8.** On veut montrer que pour toute fonction $f : [0,1] \to \mathbb{R}$ continue, on a :

$$\left(\int_0^1 f(x)\mathrm{d}x\right)^2 \leqslant \int_0^1 (f(x))^2\mathrm{d}x.$$

On a une inégalité entre intégrales et la présence de carrés. Mieux, on note que pour le produit scalaire $(f, g) \mapsto \displaystyle\int_0^1 f(x)g(x)\mathrm{d}x$, le membre de droite est égal à $\|f\|^2$ ; le majorant fait apparaître une norme, et le minorant une quantité ressemblant au produit scalaire, au carré : on peut donc espérer qu'il s'agit de l'inégalité de Cauchy-Schwarz appliquée à de bonnes fonctions.

On va l'appliquer avec $f$ et une autre fonction à déterminer ; pour déterminer laquelle, je me demande : comment trouver $g$ de sorte que $\langle f, g\rangle = \displaystyle\int_0^1 f(x)\mathrm{d}x$ ? Vu que $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)\mathrm{d}x$, un choix naturel serait pour $g$ la fonction constante égale à 1.

Vérifions : si l'on applique l'inégalité de Cauchy-Schwarz avec $f$ et $g = 1$, on obtient : $(\langle f,1\rangle)^2 \leqslant \|1\|^2\|f\|^2$. Or : $\|1\|^2 = \displaystyle\int_0^1 \mathrm{d}x = 1$, et comme prévu : $\langle f,1\rangle = \displaystyle\int_0^1 f(x)\mathrm{d}x$. On obtient donc bien $\left(\displaystyle\int_0^1 f(x)\mathrm{d}x\right)^2 \leqslant \displaystyle\int_0^1 (f(x))^2\mathrm{d}x$ et c'est ce qu'on voulait démontrer.

Même si ce n'est pas demandé : le cas d'égalité est obtenu si $f$ et 1 sont linéairement dépendantes, c'est-à-dire si et seulement si $f$ est constante.

Dans le cas particulier des fonctions, s'il apparaît un intégrande dans le membre de droite qui n'a RIEN À VOIR avec le membre de gauche, alors il faut probablement faire une opération préalable sur l'intégrale, pour faire apparaître l'intégrande du membre de droite, avant de lui appliquer l'inégalité de Cauchy-Schwarz : intégration par parties ou formule du changement de variable.

**Exercice 7.** Soit $f$ une application de classe $\mathrm{C}^1$ sur $[0,1]$ et à valeurs réelles. On suppose : $f(1) = 0$. Montrer :

$$\left(\int_0^1 f(t)\mathrm{d}t\right)^2 \leqslant \frac{1}{3}\int_0^1 \left(f'(t)\right)^2 \mathrm{d}t.$$

On est bien dans le cadre d'application de l'inégalité de Cauchy-Schwarz, avec $\|f'\|$ qui apparaît dans le membre de droite, le produit scalaire $\langle f,1\rangle$ dans le membre de gauche... Mais comme on le voit, appliquer cette inégalité, avec $f$ ou $f'$, ne donne pas du tout le résultat attendu. Demandez-vous donc : comment « passer » de $f$ à $f'$ ?

### 1.3.2 Quand penser à l'inégalité de Cauchy-Schwarz : cas abstrait

Une situation où vous pouvez y songer est **lorsqu'on vous demande de majorer un produit scalaire de deux vecteurs dont on connaît les normes**, grâce aux données du problème (vecteurs unitaires, isométries, projecteurs orthogonaux...).

À cet égard, souvenez-vous de l'effet des endomorphismes remarquables du chapitre sur les normes :

— les projections orthogonales diminuent la norme : $\|p(\vec{x})\| \leqslant \|\vec{x}\|$ (inégalité de Bessel) ;

— les symétries orthogonales et les isométries plus généralement conservent la norme : $\|f(\vec{x})\| = \|\vec{x}\|$ ;

— les normes avec les endomorphismes autoadjoints s'écrivent à l'aide de leurs valeurs propres grâce au théorème spectral (section 4.2). → page 28

Ce dernier point n'est intéressant que si l'on a des informations sur le spectre, ou si l'on veut faire apparaître une quantité dépendant des valeurs propres (trace, déterminant, autre produit scalaire avec l'endomorphisme...).

On pourrait aussi recycler le conseil du cas « concret » : songer à l'utiliser lorsqu'on voit apparaître un produit de normes comme majorant. Mais souvent l'énoncé de l'exercice a déjà simplifié une ou deux de ces normes grâce aux hypothèses (vecteurs unitaires en général). Vous ne les reconnaîtrez donc pas clairement.

**Exemple 9.** Soit $p$ un projecteur orthogonal d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$. Nous voulons montrer que pour tout vecteur $\vec{x} \in E$ **unitaire**, on a :

$$|\langle\vec{x}, p(\vec{x})\rangle| \leqslant 1.$$

**Puisqu'il s'agit de majorer le produit scalaire de vecteurs dont on connaît les normes** ($\vec{x}$ est unitaire, et l'inégalité de Bessel permet de majorer $\|p(\vec{x})\|$), **on est dans le cadre d'application de l'inégalité de Cauchy-Schwarz**. En le faisant, on obtient :

$$|\,\langle\vec{x}, p(\vec{x})\rangle\,| \leqslant \|\vec{x}\| \cdot \|p(\vec{x})\|.$$

On a par hypothèse $\|\vec{x}\| = 1$, et d'après l'inégalité de Bessel on a $\|p(\vec{x})\| \leqslant \|\vec{x}\| = 1$, d'où le résultat.

On sait aussi à quel condition il y a égalité : d'abord, il faut que l'inégalité de Bessel soit dans le cas d'égalité, et c'est vrai uniquement si $\vec{x} \in \mathrm{im}(p)$. Dans ce cas-là, on a $p(\vec{x}) = \vec{x}$ et donc $\langle\vec{x}, p(\vec{x})\rangle = \|\vec{x}\|^2 = 1$.

Parfois, le produit scalaire est « déguisé ». C'est le cas en particulier lorsqu'il apparaît un endomorphisme (anti)symétrique : le fait que le $f$ dans le produit scalaire puisse « changer de place » permet alors de transformer un produit scalaire en norme (ou inversement). C'est le cas dans l'exercice ci-dessous.

**Exercice 8.** Soit $f$ un endomorphisme autoadjoint d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$.

1. Montrer que pour tout vecteur $\vec{x} \in E$ **unitaire**, on a : $\|f(\vec{x})\|^2 \leqslant \|f^2(\vec{x})\|$.
2. Montrer que l'égalité est vérifiée si et seulement si $\vec{x}$ est un vecteur propre de $f$ (*subtil*).

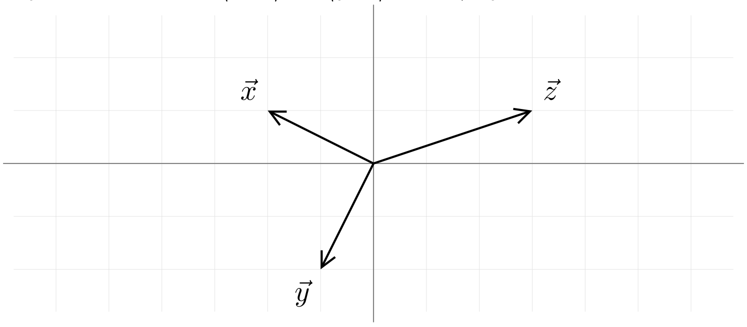
### 1.4 Comment passer de $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ à $\vec{x} = \vec{y}$

**Mise en garde 3.** Ne me faites pas dire ce que je n'ai pas dit. Il est important d'insister là-dessus :

En toute généralité, l'implication $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle \Rightarrow \vec{x} = \vec{y}$ est FAUSSE et ABOMINABLE.

On ne peut pas « identifier » dans un produit scalaire ! C'est évident : si $\vec{x}$, $\vec{y}$ et $\vec{z}$ sont pris de même norme (pour simplifier), alors l'égalité $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ signifie simplement que l'angle entre $\vec{x}$ et $\vec{z}$ est le même que l'angle entre $\vec{y}$ et $\vec{z}$. Situation extrêmement banale, qu'on rencontre aisément avec $\vec{x}$ et $\vec{y}$ distincts.

FIGURE 1 – Vecteurs $\vec{x}$, $\vec{y}$ et $\vec{z}$ tels que $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ et $\vec{x} \neq \vec{y}$. Faites le calcul pour vous en convaincre.




L'objectif de cette section est de discuter des situations où on *peut* tout de même déduire une égalité entre vecteurs à partir d'une égalité entre produits scalaires, et comment l'obtenir RIGOUREUSEMENT :

Comment déduire des égalités entre VECTEURS en partant d'égalités entre PRODUITS SCALAIRES ?

Pour cela, il faut que $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ soit vraie **pour suffisamment de vecteurs** $\vec{z}$. Mettons-nous dans le cas favorable, et le plus fréquent, où l'égalité $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ est vraie pour TOUT $\vec{z} \in E$. Alors :

1. Pour tout $\vec{z} \in E$, on a : $\langle \vec{x} - \vec{y}, \vec{z} \rangle = 0$.
2. C'est en particulier le cas pour $\vec{z} = \vec{x} - \vec{y}$, donc $\|\vec{x} - \vec{y}\|^2 = 0$.
3. Par propriété de séparation de la norme, on a donc $\vec{x} - \vec{y} = \vec{0}$, puis $\vec{x} = \vec{y}$.

Ou, plus rapidement : si $\langle \vec{x} - \vec{y}, \vec{z} \rangle = 0$ pour tout $\vec{z} \in E$, alors $\vec{x} - \vec{y} \in E^{\perp} = \{\vec{0}\}$, donc $\vec{x} - \vec{y} = \vec{0}$ puis $\vec{x} = \vec{y}$.

À noter que si vous ne savez pas si cette égalité est vraie pour tout $\vec{z} \in E$, c'est malgré tout suffisant dans les deux cas suivants :

— $\langle \vec{x}, \vec{e_i} \rangle = \langle \vec{y}, \vec{e_i} \rangle$ pour tout vecteur $\vec{e_i}$ d'une base de $E$ (et même d'une famille génératrice) : on sait en effet que $\vec{x} - \vec{y}$ est dans l'orthogonal d'un espace si et seulement s'il est orthogonal à tout vecteur d'une de ses bases, ce qui permet de reprendre le raisonnement ci-dessus ;

— si l'on sait que $\vec{x}$ et $\vec{y}$ sont dans un même sous-espace bien connu $F$, et que $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ pour tout $\vec{z} \in F$, alors c'est suffisant : on peut toujours poser $\vec{z} = \vec{x} - \vec{y} \in F$ pour conclure.

Le deuxième cas de figure est extrêmement rare.

#### 1.4.1 Cas particulier fréquent : montrer qu'une application est linéaire, est nulle, est une homothétie, etc.

Imaginons que nous ayons une application définie de $E$ dans $E$, et dont nous ne connaissons pas explicitement sa correspondance, **mais vérifiant une certaine propriété impliquant un produit scalaire**. Nous avons deux exemples de telles applications dans le cours : les endomorphismes autoadjoints et les isométries (je mentionne aussi les endomorphismes antisymétriques dans la section 4.6). → page 36

On peut nous demander alors de démontrer que $f$ est linéaire (ce n'est qu'un exemple parmi d'autres), si ce n'est pas déjà établi, c'est-à-dire :

$$\forall(\vec{x},\vec{y}) \in E^2,\ \forall\lambda\in\mathbb{R},\quad f(\vec{x}+\lambda\vec{y}) = f(\vec{x})+\lambda f(\vec{y}).$$

On adapte alors la stratégie développée plus haut dans cette section, en montrant que pour tout $(\vec{x},\vec{y}) \in E^2$ et tout $\lambda\in\mathbb{R}$, on a :

$$\forall\vec{z}\in E,\quad \left\langle f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y}))\,,\vec{z}\right\rangle = 0,$$

dont en déduit $f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y})) \in E^\perp$, puis $f(\vec{x}+\lambda\vec{y}) = f(\vec{x})+\lambda f(\vec{y})$ pour tout $(\vec{x},\vec{y}) \in E^2$ et tout $\lambda\in\mathbb{R}$.

La stratégie se transpose de manière générale à la démonstration d'une identité de la forme $f(\vec{x}) = g(\vec{x})$ ; il suffit alors d'étudier $\langle f(\vec{x}) - g(\vec{x}), \vec{y}\rangle$ pour démontrer que $f(\vec{x}) - g(\vec{x}) \in E^\perp = \{\vec{0}\}$.

On peut se demander comment on « peut penser » à cette idée saugrenue de démontrer l'égalité $\left\langle f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y}))\,,\vec{z}\right\rangle = 0$, plutôt que de démontrer directement $f(\vec{x}+\lambda\vec{y}) = f(\vec{x})+\lambda f(\vec{y})$, comme on le fait dans l'immense majorité des cas. L'idée, dans ces situations, est que même si l'application n'est pas linéaire *a priori* (on veut le démontrer), le produit scalaire *l'est, lui, par rapport à chaque variable.* Ainsi l'application « hérite » de la bilinéarité du produit scalaire grâce à la propriété qu'elle vérifie à son égard.

**Exemple 10.** Soit $(E,\langle\cdot,\cdot\rangle)$ un espace euclidien. Nous avons défini un endomorphisme autoadjoint de $E$ comme étant un endomorphisme de $E$ vérifiant :

$$\forall(\vec{x},\vec{y}) \in E^2,\quad \langle f(\vec{x}),\vec{y}\rangle = \langle \vec{x}, f(\vec{y})\rangle\,. \tag{3}$$

Nous allons montrer qu'en vérité, l'hypothèse que c'est un endomorphisme est superflu : nous allons démontrer que si $f : E\to E$ est une application QUELCONQUE vérifiant (3), alors elle est nécessairement linéaire, c'est-à-dire :

$$\forall(\vec{x},\vec{y}) \in E^2,\ \forall\lambda\in\mathbb{R},\quad f(\vec{x}+\lambda\vec{y}) = f(\vec{x})+\lambda f(\vec{y}).$$

**Nous sommes bien dans le contexte de cette section**, car nous avons des égalités entre des PRODUITS SCALAIRES et nous voulons en déduire des égalités entre VECTEURS. Cela nous incite à démontrer que pour tout $(\vec{x},\vec{y}) \in E^2$ et tout $\lambda\in\mathbb{R}$, le vecteur :

$$f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y}))$$

est nul, en démontrant qu'il est orthogonal à tout vecteur de l'espace. Soit $(\vec{x},\vec{y}) \in E^2$, soit $\lambda\in\mathbb{R}$. Pour tout $\vec{z}\in E$, on a :

$$\langle f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y})),\vec{z}\rangle = \langle f(\vec{x}+\lambda\vec{y}),\vec{z}\rangle - \langle f(\vec{x}),\vec{z}\rangle - \lambda\,\langle f(\vec{y}),\vec{z}\rangle\,,$$

et en utilisant la propriété (3) on a, pour tout $\vec{z}\in E$,

$$\langle f(\vec{x}+\lambda\vec{y}),\vec{z}\rangle \overset{(3)}{=} \langle \vec{x}+\lambda y, f(\vec{z})\rangle = \langle\vec{x}, f(\vec{z})\rangle + \lambda\,\langle\vec{y}, f(\vec{z})\rangle \overset{(3)}{=} \langle f(\vec{x}),\vec{z}\rangle + \lambda\,\langle f(\vec{y}),\vec{z}\rangle\,.$$

On en déduit :

$$\forall\vec{z}\in E,\quad \langle f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y})),\vec{z}\rangle = \langle f(\vec{x}),\vec{z}\rangle + \lambda\,\langle f(\vec{y}),\vec{z}\rangle - \langle f(\vec{x}),\vec{z}\rangle - \lambda\,\langle f(\vec{y}),\vec{z}\rangle = 0,$$

donc : $f(\vec{x}+\lambda\vec{y}) - (f(\vec{x})+\lambda f(\vec{y})) \in E^\perp = \{\vec{0}\}$, et on déduit : $f(\vec{x}+\lambda\vec{y}) = f(\vec{x})+\lambda f(\vec{y})$. Ceci étant vrai pour tout $(\vec{x},\vec{y}) \in E^2$ et tout $\lambda\in\mathbb{R}$, l'application $f$ est bien linéaire, donc un endomorphisme de $E$, autoadjoint par définition du fait de vérifier (3).

**Exercice 9.** Soit $(E,\langle\cdot,\cdot\rangle)$ un espace euclidien, et soit $f : E\to E$ une application vérifiant :

$$\forall(\vec{x},\vec{y}) \in E^2,\quad \langle f(\vec{x}), f(\vec{y})\rangle = \langle\vec{x},\vec{y}\rangle\,.$$

Montrer que $f$ est linéaire. C'est donc une isométrie. *Ici, il vaut mieux travailler directement avec une norme au carré plutôt qu'un produit scalaire.*

### 1.4.2 Et si l'on n'a pas des produits scalaires, mais des normes ?

Dans ce cas, il y a des situations plus ou moins favorables. Soit, comme dans l'exercice 9 ci-dessus, développer l'expression :

$$\|f(\vec{x}+\lambda\vec{y}) - (f(\vec{x}) + \lambda f(\vec{y}))\,\|^2$$

ne fait apparaître que des normes et produits scalaires qu'on sait simplifier (cas ultra-favorable), soit on passe des normes aux produits scalaires afin de raisonner comme ci-dessus, grâce à une **identité de polarisation** :

$$\forall(\vec{x},\vec{y})\in E^2,\quad \langle\vec{x},\vec{y}\rangle = \frac{1}{4}\left(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\right) = \frac{1}{2}\left(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2\right).$$

Hormis dans de très rares cas que je ne mentionnerai pas, le choix de l'identité de polarisation importe peu. Alors, les hypothèses sur les normes permettent de simplifier le membre de droite, et d'en déduire une expression du produit scalaire qu'on veut simplifier.

Pour savoir à quels vecteurs appliquer l'identité de polarisation, à vrai dire, cela dépend de ce qui nous arrange. En général ce sera soit à $\langle f(\vec{x}),\vec{y}\rangle$, soit à $\langle f(\vec{x}), f(\vec{y})\rangle$, pour tout $(\vec{x},\vec{y})\in E^2$. Vous en jugerez selon que les normes soient simplifiables ou non.

**Exercice 10.** Soit $(E,\langle\cdot,\cdot\rangle)$ un espace euclidien, dont on note $\|\cdot\|$ la norme euclidienne. Soit $f : E\to E$ une application vérifiant $f(\vec{0})=\vec{0}$, et :

$$\forall(\vec{x},\vec{y})\in E^2,\quad \|f(\vec{x})-f(\vec{y})\| = \|\vec{x}-\vec{y}\|.$$

Montrer que $f$ est linéaire, et que c'est une isométrie. Indication : se ramener à la situation de l'exercice 9 ci-dessus.

### 1.4.3 Et si l'on n'a pas des normes, mais des produits scalaires... de la forme $\langle f(\vec{x}),\vec{x}\rangle$ ?

Dans tous ces raisonnements, on veut impérativement manipuler le produit scalaire $\langle f(\vec{x}),\vec{y}\rangle$, afin d'avoir toute latitude sur le choix de $\vec{y}$, on l'a vu ; mais il arrive malheureusement que nos hypothèses ne soient pas sur $\langle f(\vec{x}),\vec{y}\rangle$ (et ne puissent pas s'y ramener trivialement), mais sur $\langle f(\vec{x}),\vec{x}\rangle$. Comment faire, dans ce cas ?

On peut se ramener malgré tout aux cas précédents si l'on a de bonnes hypothèses : si $f$ est un endomorphisme autoadjoint, par exemple, alors $(\vec{x},\vec{y})\mapsto\langle f(\vec{x}),\vec{y}\rangle$ est une forme bilinéaire symétrique, et cela implique en particulier une identité de polarisation :

$$\forall(\vec{x},\vec{y})\in E^2,\quad \langle f(\vec{x}),\vec{y}\rangle = \frac{1}{4}\left(\langle f(\vec{x}+\vec{y}),\vec{x}+\vec{y}\rangle - \langle f(\vec{x}-\vec{y}),\vec{x}-\vec{y}\rangle\right) \tag{4}$$

Nous l'admettons ici et en parlons plus amplement dans la section 4.1 (exercice 22). L'intérêt est qu'alors, si l'on connaît $\langle f(\vec{x}),\vec{x}\rangle$ pour tout $\vec{x}$, alors nous connaissons aussi $\langle f(\vec{x}\pm\vec{y}),\vec{x}\pm\vec{y}\rangle$, et donc nous connaissons aussi $\langle f(\vec{x}),\vec{y}\rangle$ : on se ramène à la situation précédente ! → page 26

**Exemple 11.** Soit $f$ un endomorphisme autoadjoint. On suppose que pour tous $\vec{x}$, $\vec{y}\in E$ **non nuls** :

$$\frac{\langle f(\vec{x}),\vec{x}\rangle}{\|\vec{x}\|^2} = \frac{\langle f(\vec{y}),\vec{y}\rangle}{\|\vec{y}\|^2} \tag{5}$$

(on pourrait trouver tordue une telle propriété, mais ce n'est que l'identité raisonnable $\langle f(\vec{x}),\vec{x}\rangle = \langle f(\vec{y}),\vec{y}\rangle$ supposée pour tous vecteurs unitaires, et qu'on étend à tous vecteurs en les divisant par leurs normes).

On veut en déduire que $f$ est une homothétie, c'est-à-dire : il existe $\lambda\in\mathbb{R}$ tel que :

$$\forall\vec{x}\in E,\quad f(\vec{x}) = \lambda\vec{x}.$$

Encore faudrait-il avoir idée de la valeur de ce $\lambda$. Je vous laisserai vous en convaincre par une analyse et synthèse, mais je vais démontrer qu'un choix de $\lambda$ qui convient est la quantité :

$$\lambda = \frac{\langle f(\vec{x}),\vec{x}\rangle}{\|\vec{x}\|^2}, \tag{6}$$

l'égalité (5) démontrant que c'est bien une quantité constante. Notons qu'on en déduit une égalité qui vaut pour tout $\vec{x} \in E$ (y compris nul) : $\langle f(\vec{x}), \vec{x}\rangle = \lambda \|\vec{x}\|^2$.

**Nous sommes pile dans la configuration de cette section**. En effet, nous avons des égalités entre des PRODUITS SCALAIRES (identité (5)), et nous voulons en déduire des égalités entre VECTEURS ($f(\vec{x}) = \lambda\vec{x}$). Conformément aux conseils de cette section, nous allons démontrer que si $\vec{x} \in E$, alors le vecteur $f(\vec{x}) - \lambda\vec{x}$ est nul, en montrant qu'il est orthogonal à tout vecteur $\vec{y}$ de l'espace.

Soient $\vec{x} \in E$ et $\vec{y} \in E$. Alors :

$$\langle f(\vec{x}) - \lambda\vec{x}, y\rangle = \langle f(\vec{x}), \vec{y}\rangle - \lambda \langle \vec{x}, \vec{y}\rangle .$$

L'identité (5) ne permet pas de simplifier $\langle f(\vec{x}), \vec{y}\rangle$ ; c'est là que je vais utiliser le conseil ci-dessus de passer par une identité de polarisation pour me ramener aux produits scalaires $\langle f(\vec{x} \pm \vec{y}), \vec{x} \pm \vec{y}\rangle$ que je saurai simplifier : c'est possible parce que $f$ est un endomorphisme autoadjoint. On a :

$$\begin{aligned}\langle f(\vec{x}), \vec{y}\rangle = \frac{1}{4}\left(\langle f(\vec{x}+\vec{y}), \vec{x}+\vec{y}\rangle - \langle f(\vec{x}-\vec{y}), \vec{x}-\vec{y}\rangle\right) &\overset{(6)}{=} \frac{1}{4}\left(\lambda\|\vec{x}+\vec{y}\|^2 - \lambda\|\vec{x}-\vec{y}\|^2\right)\\ &= \lambda \cdot \frac{1}{4}\left(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\right) = \lambda\langle \vec{x}, \vec{y}\rangle ,\end{aligned}$$

la dernière égalité découlant encore d'une identité de polarisation. On trouve bien :

$$\langle f(\vec{x}) - \lambda\vec{x}, \vec{y}\rangle = \langle f(\vec{x}), \vec{y}\rangle - \lambda \langle \vec{x}, \vec{y}\rangle = 0,$$

et ceci vaut pour tout $\vec{y} \in E$ donc : $f(\vec{x}) - \lambda\vec{x} \in E^{\perp} = \{\vec{0}\}$. On en déduit $f(\vec{x}) - \lambda\vec{x} = \vec{0}$, donc $f(\vec{x}) = \lambda\vec{x}$. Ceci vaut pour tout $\vec{x} \in E$, donc $f$ est une homothétie : comme promis.

**Exercice 11.** Redémontrer ce résultat en une ligne avec la version matricielle du théorème spectral.

Cet exercice est une illustration parmi d'autres de la section 2.3 consacrée à l'intérêt des bases orthonormées, en particulier dans leur utilisation pour expliciter la forme d'une matrice.

**Exercice 12.**

1. Soient $f$ et $g$ deux endomorphismes autoadjoints. Montrer que si : $\forall \vec{x} \in E$, $\langle \vec{x}, f(\vec{x})\rangle = \langle \vec{x}, g(\vec{x})\rangle$, alors : $f = g$.
2. Soient $A$ et $B$ deux matrices symétriques. Montrer que si : $\forall X \in \mathrm{M}_{n,1}(\mathbb{R})$, $X^{\top}AX = X^{\top}BX$, alors : $A = B$.

Faire le lien avec la section 2.3.

# 2 ✓ L'importance des bases orthonormées

## 2.1 ✓ Pour le calcul de produits scalaires et normes

Le principal usage des bases orthonormées est, de très loin, pour le calcul de produits scalaires et de normes. Si $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base orthonormée de $(E, \langle\cdot,\cdot\rangle)$, et si $\vec{x}$ et $\vec{y}$ sont de coordonnées respectives $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans cette base, alors :

$$\|\vec{x}\|^2 = \sum_{k=1}^{n} x_k^2, \quad \langle \vec{x}, \vec{y}\rangle = \sum_{k=1}^{n} x_k y_k.$$

Ce sont les mêmes expressions que le produit scalaire et la norme usuels : elles sont simples et c'est ce qui en fait l'intérêt. Vous avez bien dû remarquer la difficulté que l'on rencontre dans le calcul de certaines normes définies sur des espaces de fonctions ou de polynômes, qui se ramènent à du calcul intégral long et pénible. Nous voulons donc nous en passer. À noter que nous savons très simplement expliciter $x_k$ et $y_k$ si nous en avons besoin (et là encore c'est spécifique aux bases orthonormées), puisqu'on a $x_k = \langle \vec{x}, \vec{e}_k\rangle$ et $y_k = \langle \vec{y}, \vec{e}_k\rangle$.

En particulier, lorsque nous calculons des projections orthogonales sur un sous-espace $F$ en utilisant une base orthonormée (obtenue en général à partir de l'algorithme de Gram-Schmidt) :

$$p(\vec{x}) = \langle \vec{x}, \vec{v}_1 \rangle \vec{v}_1 + \langle \vec{x}, \vec{v}_2 \rangle \vec{v}_2 = \alpha \vec{v}_1 + \beta \vec{v}_2,$$

VOUS NE TOUCHEZ PAS À $\vec{v}_1$ ET $\vec{v}_2$ ! En utilisant le fait que $(\vec{v}_1, \vec{v}_2)$ soit orthonormée, vous avez directement $\|p(\vec{x})\|^2 = \alpha^2 + \beta^2$. Si vous remplacez $\vec{v}_1$ et $\vec{v}_2$ par leurs expressions respectives (dans la base canonique s'il y en a une), vous n'aurez en général plus une base orthonormée et le calcul de norme deviendra infernal. Nous reparlons plus amplement des projections orthogonales dans la section 3.1 (*Calculs avec les projecteurs orthogonaux, les symétries orthogonales*). → page 22

Tout ce qui est prodigué ici est TOTALEMENT FAUX si la base n'est pas orthonormée et on prendra garde à ne pas appliquer ces formules hors contexte.

## 2.2 Un cas particulier TRÈS fréquent : les polynômes orthogonaux

Dans toute cette partie, $(P_k)_{k \in [\![0,n]\!]}$ est une base de polynômes orthogonaux obtenue en appliquant l'algorithme d'orthonormalisation de Gram-Schmidt à la base canonique $(X^k)_{k \in [\![0,n]\!]}$, avec un produit scalaire défini par :

$$\forall (P,Q) \in \mathbb{R}_n[X], \quad \langle P, Q \rangle = \int_a^b P(t)Q(t)w(t)\mathrm{d}t,$$

où $w : [a,b] \to \mathbb{R}$ une fonction continue, positive et non nulle partout (en exercice nous avons été plus restrictif et exigé la stricte positivité).

Tous les problèmes sur les familles de polynômes orthogonaux commenceront invariablement par la démonstration de ces propriétés très classiques, **qu'il convient donc de savoir traiter** :

1. Pour tout $k \in [\![0,n]\!]$, on a : $\deg(P_k) = k$ (*conséquence de l'égalité entre « Vect » dans l'algorithme*).
2. Pour tout $k \in [\![1,n]\!]$, on a : $P_k \in \mathbb{R}_{k-1}[X]^\perp$ (*conséquence de l'orthogonalité à une base de* $\mathbb{R}_{k-1}[X]$).

D'autres propriétés classiques valables pour tout $k \in [\![0,n]\!]$ :

1. Le coefficient dominant de $P_k$ est positif (*conséquence de* $\left\langle P_k, X^k \right\rangle > 0$ *d'après l'algorithme de Gram-Schmidt*).
2. Le polynôme $P_k$ est scindé et à racines simples, qui sont toutes dans $[a,b]$ (*difficile*).

Et enfin, nous avons un résultat d'unicité très utilisé dans les sujets de concours :

> Pour tout $k \in [\![0,n]\!]$, $P_k$ est l'unique polynôme de degré $k$, à coefficient dominant strictement positif, vérifiant : $\displaystyle\int_a^b (P_k(t))^2 w(t)\mathrm{d}t = 1$, et : $\forall Q \in \mathbb{R}_{k-1}[X]$, $\displaystyle\int_a^b P_k(t)Q(t)w(t)\mathrm{d}t = 0$.

Je ne démontre rien de tout cela, c'est un exercice.

### 2.2.1 Comment utiliser ces polynômes orthogonaux ?

Je ne m'étendrai pas sur leurs nombreuses utilisations (dans le calcul approché d'intégrales et l'étude de certaines équations différentielles linéaires par exemple). **Je ne parlerai pas non plus de leur utilisation dans le calcul de distances**, puisqu'en cela ils ne montrent rien de spécifique aux polynômes. Je veux seulement que vous sachiez comment mobiliser leurs propriétés. La première propriété majeure utilisée est l'orthogonalité aux polynômes de degré inférieur :

$$\boxed{\text{Si } \deg(Q) < k \text{ alors } \langle P_k, Q \rangle = \int_a^b P_k(t)Q(t)w(t)\mathrm{d}t = 0. \ (*)}$$

Soyez bien attentifs à cela. Souvent de l'orthonormalité vous ne retenez que l'aspect unitaire, alors que c'est (presque) inutile. La seconde propriété majeure utilisée est l'**unicité**, que je réécris un peu différemment :

> Si $P$ est orthogonal à tout polynôme de degré strictement inférieur (c'est-à-dire :
> $$\forall Q \in \mathbb{R}[X],\ \deg(Q) < \deg(P) \Rightarrow \int_a^b P(t)Q(t)w(t)\mathrm{d}t = 0),$$
> alors il existe $\lambda \in \mathbb{R}$ tel que : $P = \lambda P_k$, où $k = \deg(P)$. (†)

Réfléchissez à comment le déduire du résultat d'unicité (remarquez les hypothèses devenues moins contraignantes). On explicite alors $\lambda$ en comparant les coefficients dominants, ou en calculant la norme de $P$ si on y parvient.

Enfin, la troisième propriété majeure est... Que c'est une base orthonormée ! En particulier, on a les différentes formules rappelées dans la section : $P = \sum\limits_{k=0}^{n} \langle P, P_k\rangle P_k$, $\|P\|^2 = \sum\limits_{k=0}^{n} (\langle P, P_k\rangle)^2$, etc.

Cela vous donne trois stratégies pour démontrer une formule invoquant des polynômes orthogonaux. Si vous devez démontrer une égalité :

— entre **intégrales** : faites des produits scalaires avec les $P_k$ et utilisez $(*)$ pour simplifier ;

— de la forme « un polynôme $= \bigstar \cdot P_k$ » : montrez que le membre de gauche vérifie la propriété (†) (on vous donne en général des indications pour montrer que les intégrales sont nulles : intégrer par parties, etc.) ;

— de la forme « un polynôme = une somme avec plusieurs $P_k$ » : utilisez les expressions explicites des vecteurs, normes, etc., dans une base orthonormée.

Ce n'est pas exhaustif mais cela vous permettra déjà d'y voir plus clair.

**Exemple 12. (illustration du premier cas)** Soit $n \in \mathbb{N} \setminus \{0\}$. Que vaut $\int_a^b P_n(t)w(t)\mathrm{d}t$ ? J'affirme que même sans expliciter $P_n$, nous pouvons calculer cette intégrale. En effet, ce n'est rien d'autre que $\langle P_n,1\rangle$, or $0 < n$ par hypothèse et donc d'après $(*)$ : $\langle P_n,1\rangle = \int_a^b P_n(t)w(t)\mathrm{d}t = 0$.

**Exercice 13. (illustration du second cas)** Soit $n \in \mathbb{N}$. On prend ici $w = 1$, $[a,b] = [-1,1]$, de sorte que $(P_k)_{k\in[\![0,n]\!]}$ soit orthonormée pour le produit scalaire $(P,Q) \mapsto \int_{-1}^1 P(t)Q(t)\mathrm{d}t$. On pose : $L_n = \frac{1}{2^n n!}\left((X^2-1)^n\right)^{(n)}$.

1. Pour tout $P \in \mathbb{R}[X]$ et tout $k \in [\![0,n]\!]$, montrer en intégrant par parties :
$$\int_{-1}^1 L_n(t)P(t)\mathrm{d}t = \frac{(-1)^k}{2^n n!}\int_{-1}^1 \left((t^2-1)^n\right)^{(n-k)} P^{(k)}(t)\mathrm{d}t.$$
On déplorera bien sûr l'abus de notation « à la physicienne ».
2. En déduire que $L_n$ vérifie (†), puis : $L_n = \sqrt{\frac{2}{n+1}}P_n$ (il faudra calculer $\|L_n\|^2$). Cette identité permet d'en déduire $P_n$ sans même avoir à appliquer l'algorithme d'orthonormalisation, ce qui est plus commode.

**Exemple 13. (illustration du troisième cas)** Soit $n \in \mathbb{N} \setminus \{0\}$. Démontrons qu'il existe $(\alpha,\beta,\gamma) \in \mathbb{R}^3$ tel que : $XP_n = \alpha P_{n+1} + \beta P_n + \gamma P_{n-1}$. Le fait qu'il apparaisse une somme de $P_k$ dans le membre de droite fait penser au troisième cas de figure : voyons en quoi le déduire de l'expression explicite de $XP_n$ dans la base orthonormée $(P_k)_{k\in[\![0,n+1]\!]}$ (attention au fait que $\deg(XP_n) = n+1$). On a :
$$XP_n = \sum_{k=0}^{n+1} \langle XP_n, P_k\rangle P_k.$$

Si on isole les termes pour $k \in \{n-1, n, n+1\}$, on obtient une égalité proche du résultat voulu :
$$XP_n = \underbrace{\langle XP_n, P_{n+1}\rangle}_{=\alpha} P_{n+1} + \underbrace{\langle XP_n, P_n\rangle}_{=\beta} P_n + \underbrace{\langle XP_n, P_{n-1}\rangle}_{=\gamma} P_{n-1} + \sum_{k=0}^{n-2} \langle XP_n, P_k\rangle P_k.$$

Il reste à justifier que $\langle XP_n, P_k\rangle = 0$ pour tout $k \leqslant n-2$. Pour cela, on note que :

$$\forall k \in [\![0, n-2]\!], \quad \langle XP_n, P_k\rangle = \int_a^b tP_n(t)P_k(t)w(t)\mathrm{d}t = \int_a^b P_n(t)(tP_k(t))w(t)\mathrm{d}t = \langle P_n, XP_k\rangle \overset{(*)}{=} 0$$

car $\deg(XP_k) = k+1 \leqslant n-1 < n$, d'où le résultat : $XP_n = \alpha P_{n+1} + \beta P_n + \gamma P_{n-1}$.

## 2.3 ✓ Pour d'autres calculs : trace, produit matriciel, etc.

Retenez UNE quantité qui revient tout le temps dans un contexte géométrique :

$$\langle \vec{e}_i, f(\vec{e}_j)\rangle, \text{ et son équivalent matriciel : } E_i^\top A E_j,$$

où $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base orthonormée de $E$ dans le premier cas, et $(E_1, \ldots, E_n)$ la base canonique de $\mathrm{M}_{n,1}(\mathbb{R})$ (qui est aussi une base orthonormée, pour le produit scalaire usuel $(X,Y) \mapsto X^\top Y$). Connaître ces deux quantités permet en principe de TOUT savoir sur $f$ ou sur $A$ : comme on l'a vu, elles donnent leurs coefficients (ceux de la matrice de $f$ dans $(\vec{e}_1, \ldots, \vec{e}_n)$ dans le cas de l'endomorphisme). Notez par ailleurs :

$$\forall (i,j) \in [\![1, n]\!],\ a_{i,j} = E_i^\top A E_j.$$

Faites le calcul pour vous en convaincre, ou utilisez l'expression des coordonnées d'un vecteur dans une base orthonormée du cours (puis appliquez ce résultat à $f(\vec{e}_j)$).

Puisque ces quantités donnent des coefficients matriciels, elles donnent aussi la trace :

$$\mathrm{tr}(f) = \sum_{k=1}^n \langle f(\vec{e}_k), \vec{e}_k\rangle, \quad \text{et :} \quad \mathrm{tr}(A) = \sum_{k=1}^n E_k^\top A E_k.$$

Guettez donc toute propriété que l'énoncé donne sur ces produits scalaires. Mais notez aussi (on y pense moins) qu'on peut en déduire des produits matriciels : il y a une ressemblance entre les formules :

$$c_{i,j} = \sum_{k=1}^n a_{i,k}b_{k,j}, \text{ et : } \langle \vec{x}, \vec{y}\rangle = \sum_{k=1}^n \langle \vec{x}, \vec{e}_k\rangle \langle \vec{e}_k, \vec{y}\rangle.$$

Si les vecteurs $\vec{x}$ et $\vec{y}$ sont deux vecteurs $\vec{x} = \vec{u}_i$ et $\vec{y} = \vec{u}_j$ d'une famille de vecteurs, alors le parallèle devient absolu. Vous pouvez par conséquent simplifier des produits entre matrices de familles de vecteurs (exprimés dans une base ORTHONORMÉE) dès que vous avez des relations d'orthogonalité.

**Exercice 14.** Soit $\mathcal{F} = (\vec{u}_1, \ldots, \vec{u}_n)$ une famille de vecteurs de $E$, et soit $G_\mathcal{F}$ la matrice de $\mathcal{F}$ dans une base orthonormée $(\vec{e}_1, \ldots, \vec{e}_n)$. Montrer que $\mathcal{F}$ est une famille de vecteurs orthogonaux si et seulement si $G_\mathcal{F}^\top G_\mathcal{F}$ est diagonale, et est une base orthonormée si et seulement si $G_\mathcal{F}^\top G_\mathcal{F} = \mathrm{I}_n$.

Songez enfin que pour les endomorphismes **autoadjoints**, vous avez toujours une base qui recoupe les deux meilleures propriétés possibles d'une base : une base ORTHONORMÉE constituée de VECTEURS PROPRES. C'est le théorème spectral. Commencez **toujours** une réflexion sur ces endomorphismes en invoquant une telle base (plus de détails dans la section 4.2).

### 2.3.1 ☢ Démonstration de nouvelles identités très générales (et magnifiques)

La formule :

$$\forall \vec{x} \in E, \quad \vec{x} = \sum_{i=1}^n \langle x, \vec{e}_i\rangle \vec{e}_i$$

donne une écriture EXPLICITE de TOUT vecteur $\vec{x}$ en fonction des $\vec{e}_i$. L'explicitation est extrêmement utile.

On peut utiliser cette identité en des circonstances qui ne sont *a priori* pas géométriques. Si l'on vous demande de démontrer une identité de la forme :

$$\forall \vec{x} \in E, \quad \vec{x} = \sum (\text{réel dépendant de } \vec{x}) \times (\text{vecteur indépendant de } \vec{x}),$$

avec $E$ un espace préhilbertien ($\mathbb{R}^n$, $\mathrm{M}_n(\mathbb{R})$, les espaces de fonctions, de polynômes, etc.), alors il y a fort à parier que ces vecteurs indépendants de $\vec{x}$ soient une base orthonormée pour un produit scalaire que vous aurez à déterminer, et dans ce cas les réels en facteur ne sont rien d'autre que les produits scalaires de $\vec{x}$ avec les vecteurs de cette base orthonormée : en les calculant, vous aurez l'identité désirée !

Il manque un aspect crucial de la méthode pour qu'elle soit payante : comment reconnaître le produit scalaire à utiliser ? et la base orthonormée ? Pour cela, suivez ces conseils, applicables en quasiment toutes circonstances :

— vous prenez pour produit scalaire :

$$\forall(\vec{x},\vec{y}) \in E^2, \quad \langle \vec{x},\vec{y}\rangle = \sum (\text{le réel dépendant de } \vec{x}) \times (\text{le réel dépendant de } \vec{y})\,;$$

pour comprendre pourquoi ce choix est pertinent, comparons ce qu'on sait être vrai en toute généralité dans une base orthonormée, et ce que l'on a ici :

| Ce qu'on sait | Ce qu'on veut |
|---|---|
| $\langle \vec{x},\vec{y}\rangle = \sum_{i=1}^{n} \langle \vec{x},\vec{e}_i\rangle \langle \vec{y},\vec{e}_i\rangle$ | ? |
| $\vec{x} = \sum_{i=1}^{n} \langle x,\vec{e}_i\rangle \vec{e}_i$ | $\vec{x} = \sum (\text{réel dépendant de } \vec{x}) \times (\text{vecteur indépendant de } \vec{x})$ |

Ce parallèle est cohérent si l'on pose, par identification :

$$\langle \vec{x},\vec{e}_i\rangle = (\text{le réel dépendant de } \vec{x})\,, \qquad (\text{et : } \vec{e}_i = (\text{vecteur indépendant de } \vec{x})\,)$$

et alors, toujours en suivant ce parallèle, on peut compléter la case en haut à droite du tableau :

| Ce qu'on sait | Ce qu'on veut |
|---|---|
| $\langle \vec{x},\vec{y}\rangle = \sum_{i=1}^{n} \langle \vec{x},\vec{e}_i\rangle \langle \vec{y},\vec{e}_i\rangle$ | $\langle \vec{x},\vec{y}\rangle = \sum (\text{le réel dépendant de } \vec{x}) \times (\text{le réel dépendant de } \vec{y})$ |
| $\vec{x} = \sum_{i=1}^{n} \langle x,\vec{e}_i\rangle \vec{e}_i$ | $\vec{x} = \sum (\text{réel dépendant de } \vec{x}) \times (\text{vecteur indépendant de } \vec{x})$ |

— vous vérifiez que la famille constituée des vecteurs indépendants de $\vec{x}$, dans la somme de l'identité à prouver, est orthonormée pour ce produit scalaire, et si le cardinal est suffisant alors c'est une base (à noter que les espaces de fonctions et polynômes sont en général de dimension infinie : vous aurez éventuellement besoin de restreindre le degré, par exemple, pour vous ramener à la dimension finie).

La stratégie sera plus parlante avec l'exemple ci-dessous.

**En général, ce type de raisonnement est fructueux dans des espaces vectoriels de fonctions ou de polynômes, parce que la variété de produits scalaires est plus riche**. Dans $\mathbb{R}^n$, $\mathrm{M}_{n,1}(\mathbb{R})$ ou $\mathrm{M}_n(\mathbb{R})$, tous les produits scalaires se ressemblent trop pour proposer une grande variété de nouvelles identités issues de cette formule. Vous aurez donc, en général, à penser à ce laïus pour démontrer une identité de la forme :

$$\forall P \in \mathbb{R}_n[X], \quad P = \sum (\text{réel dépendant de } P) \times (\text{polynôme sans rapport avec } P)\,,$$

et en général le réel dépendant de $P$ est proportionnel à $P(a)$, $P^{(k)}(a)$, $\int_I P(t)f(t)\mathrm{d}t$, etc. (je donne les exemples les plus courants issus de produits scalaires). De même si je remplace les polynômes par des fonctions.

**Exemple 14.** Démonstration géométrique de la formule de Taylor pour les polynômes. Soit $a \in \mathbb{R}$. Montrons :

$$\forall P \in \mathbb{R}_n[X], \quad P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k.$$

Pour voir en quoi cela découlerait des considérations géométriques développées ci-dessus, il faut identifier : quel produit scalaire utiliser ? quelle base orthonormée ? Suivant la discussion ci-dessus :

— le « réel dépendant de $P$ » est $P^{(k)}(a)$ ;

— le « polynôme sans rapport avec $P$ » est $\frac{(X-a)^k}{k!}$.

Nous introduisons donc le produit scalaire défini par :

$$\forall (P,Q) \in (\mathbb{R}_n[X])^2, \quad \langle P,Q\rangle = \sum_{\ell=0}^{n} P^{(\ell)}(a)Q^{(\ell)}(a).$$

Nous vous laissons vérifier qu'il s'agit effectivement d'un produit scalaire. La seule subtilité est pour la propriété de séparation : pour cela, vous aurez besoin de remarquer que si $\langle P,P\rangle = 0$, alors $a$ est une racine de $P$ d'ordre de multiplicité au moins $n+1$, ce qui est impossible si $P$ est non nul pour des raisons de degré.

Pour montrer l'identité voulue, regardons d'abord si les polynômes $\frac{(X-a)^k}{k!}$, pour $k \in [\![0,n]\!]$, forment une base orthonormée de $\mathbb{R}_n[X]$ : nous les diviserons au besoin par leurs normes s'ils ne sont qu'orthogonaux. Pour cela, il nous est nécessaire de calculer les dérivées successives de $(X-a)^k$ pour $k \in [\![0,n]\!]$ : on a facilement :

$$\forall \ell \in [\![0,k-1]\!], \quad \left((X-a)^k\right)^{(\ell)} = k(k-1)\cdots(k-\ell+1)(X-a)^{k-\ell} = \frac{k!}{(k-\ell)!}(X-a)^{k-\ell}, \quad \text{et :}$$

$$\left((X-a)^k\right)^{(k)} = k!, \quad \text{et :} \quad \forall \ell \in [\![k+1,n]\!], \quad \left((X-a)^k\right)^{(\ell)} = 0.$$

On en déduit alors que TOUTES les dérivées successives de $(X-a)^k$ s'annulent en $a$, sauf la $k^{\mathrm{e}}$ (qui est un polynôme constant égal à $k!$). Ainsi, dès qu'on fait un produit scalaire impliquant $(X-a)^k$ (pour le produit scalaire défini plus haut), tous les termes de la somme sont nuls sauf celui correspondant à l'indice $k$. On a donc facilement :

$$\forall (k,k') \in [\![0,n]\!]^2, \quad \left\langle (X-a)^k, (X-a)^{k'} \right\rangle = \begin{cases} \displaystyle\sum_{\substack{\ell=0 \\ \ell\neq k}}^{n} 0\times 0 + (k!)^2 = (k!)^2 & \text{si } k=k', \\ \displaystyle\sum_{\substack{\ell=0 \\ \ell\neq k \\ \ell\neq k'}}^{n} 0\times 0 + k!\times 0 + 0\times (k')! = 0 & \text{si } k\neq k', \end{cases}$$

ce dont on déduit que la famille $\left(\frac{(X-a)^k}{k!}\right)_{k\in[\![0,n]\!]}$ est orthonormale et de cardinal $n+1$, donc est une base orthonormale de $\mathbb{R}_n[X]$. On en déduit, en adaptant ici la formule $\vec{x} = \sum\limits_{i=1}^{n} \langle \vec{x}, \vec{e_i}\rangle\, \vec{e_i}$ :

$$\forall P \in \mathbb{R}_n[X], \quad P = \sum_{k=0}^{n} \left\langle P, \frac{(X-a)^k}{k!} \right\rangle \frac{(X-a)^k}{k!},$$

et comme, pour tout $P \in \mathbb{R}_n[X]$ et tout $k \in [\![0,n]\!]$, on a :

$$\left\langle P, \frac{(X-a)^k}{k!} \right\rangle = \frac{1}{k!}\sum_{\ell=0}^{n} P^{(\ell)}(a) \underbrace{\left(\left((X-a)^k\right)^{(\ell)}\right)(a)}_{=0 \text{ si } \ell\neq k} = \frac{1}{k!}P^{(k)}(a)\cdot k! = P^{(k)}(a),$$

on en déduit :

$$\forall P \in \mathbb{R}_n[X], \quad P = \sum_{k=0}^{n} P^{(k)}(a)\frac{(X-a)^k}{k!},$$

ce qu'on voulait démontrer.

Nous vous laissons vérifier l'efficacité de cette approche avec la classique *interpolation de Lagrange*, dont nous redémontrons la formule la plus importante. En vérité, ce sera plus facile dans cet exercice que dans l'exemple ci-dessus.

**Exercice 15. (interpolation de Lagrange)** Soient $n \in \mathbb{N}$ et $a_0, \ldots, a_n$ des réels distincts. On pose :

$$\forall (P,Q) \in (\mathbb{R}_n[X])^2, \quad \langle P, Q\rangle = \sum_{k=0}^{n} P(a_k)Q(a_k).$$

1. Montrer que l'on définit ainsi un produit scalaire sur $\mathbb{R}_n[X]$.
2. Soit $(L_0, \ldots, L_n)$ la famille des polynômes interpolateurs de Lagrange associée à $(a_0, \ldots, a_n)$. Montrer que c'est une base orthonormée de $\mathbb{R}_n[X]$.
3. Conclure que, pour tout $P \in \mathbb{R}_n[X]$, on a : $P = \sum\limits_{i=0}^{n} P(a_i)L_i$.

Cette approche n'est pas seulement fructueuse avec l'identité $\vec{x} = \sum\limits_{i=0}^{n} \langle x, \vec{e}_i\rangle\, \vec{e}_i$. Tout le propos ci-dessus se généralise, par exemple, à l'identité : $\|\vec{x}\|^2 = \sum\limits_{i=0}^{n} (\langle x, \vec{e}_i\rangle)^2$. Voici, en exercice, un exemple célèbre que vous avez dû rencontrer en Physique.

**Exercice 16. (inégalité de Bessel et séries de Fourier)** On définit un produit scalaire sur $E = \mathrm{C}^0([-\pi,\pi],\mathbb{R})$ ainsi :

$$\forall (f,g) \in E^2, \quad \langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\mathrm{d}t.$$

1. On pose :

$$\forall n \in \mathbb{N}\setminus\{0\},\ \forall t \in [-\pi,\pi], \quad c_n(t) = \cos(nt),\ s_n(t) = \sin(nt), \quad \text{et : } c_0 = \frac{1}{\sqrt{2}}.$$

   et : $\forall N \in \mathbb{N}\setminus\{0\}$, $F_N = \mathrm{Vect}_{\mathbb{R}}\left((c_0, c_1, \ldots, c_N, s_1, \ldots, s_N)\right)$. Montrer que pour tout $N \in \mathbb{N}\setminus\{0\}$, la famille $(c_0, c_1, \ldots, c_N, s_1, \ldots, s_N)$ est une base orthonormale de $F_N$ pour ce produit scalaire.
2. On note $p_N$ la projection orthogonale de $E$ sur $F_N$ (qui existe car $F_N$ est de dimension finie). On pose aussi :

$$\forall f \in E,\ \forall n \in \mathbb{N}\setminus\{0\},\ a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)\mathrm{d}t,\ b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)\mathrm{d}t, \quad a_0 = \frac{1}{\sqrt{2}\pi}\int_{-\pi}^{\pi} f(t)\mathrm{d}t.$$

   Montrer que pour toute fonction $f \in E$ et tout entier $N \geqslant 1$, on a :

$$\|f\|^2 \geqslant \|p_N(f)\|^2 = a_0^2 + \sum_{n=1}^{N}\left(a_n^2 + b_n^2\right).$$

3. Soit $f \in E$. Montrer que la série $\sum\limits_{n\geqslant 1}\left(a_n^2 + b_n^2\right)$ converge, et qu'on a :

$$a_0^2 + \sum_{n=1}^{+\infty}\left(a_n^2 + b_n^2\right) \leqslant \|f\|^2.$$

   En vérité, nous pouvons démontrer qu'il y a égalité avec ces hypothèses : c'est l'identité de Parseval.
4. **Application.** Soit $f \in E$. Montrer le lemme de Riemann-Lebesgue :

$$\lim_{n\to+\infty}\int_{-\pi}^{\pi} f(t)\cos(nt)\mathrm{d}t = 0, \qquad \lim_{n\to+\infty}\int_{-\pi}^{\pi} f(t)\sin(nt)\mathrm{d}t = 0.$$

   Contrairement aux situations où nous l'avions démontré en devoir surveillé ou travaux dirigés, ICI NOUS N'AVONS PAS BESOIN DE LA CLASSE $\mathrm{C}^1$ DE $f$ ! Au vu des complications que nécessitait cette hypothèse dès que nous en avions besoin (calcul de limite avec un développement limité lourd, et théorème de la limite de la dérivée), c'est un IMMENSE apport.

Cette méthode est un des nombreux succès géniaux de la géométrie euclidienne généralisée à des espaces vectoriels abstraits (nous n'avons pas besoin de vous convaincre que la théorie des séries de Fourier a des applications hors des mathématiques), et elle donne tout son crédit à ce chapitre.

# 3 ✓ Minimisation de distance

Dans cette section je traite du problème de *reconnaître* une distance à calculer. Pour le calcul effectif, c'est dans la section suivante (*Calculs avec les projecteurs orthogonaux, les symétries orthogonales*). → page 22

La question se pose dès qu'on nous demande de déterminer une borne inférieure de la forme :

$$\inf_{(a,b,\dots)\in\mathbb{R}^\star} \cdots (\heartsuit - a\cdot\spadesuit - b\cdot\clubsuit - \cdots)^2 \cdots$$

où les points de suspension cachent en général une somme ou une intégrale, et les symboles $\heartsuit$, $\spadesuit$, $\clubsuit$ peuvent au choix être des réels, des fonctions, des polynômes, etc. Seulement, pour y reconnaître une distance à un sous-espace vectoriel, c'est-à-dire une quantité de la forme $\mathrm{d}(\vec{x},F) = \inf_{\vec{y}\in F} \|\vec{x}-\vec{y}\|$ avec $\vec{x}\in E$, il faut reconnaître :

1. L'espace vectoriel $E$.
2. Le produit scalaire $\langle\cdot,\cdot\rangle$.
3. Le vecteur $\vec{x}\in E$.
4. Le sous-espace vectoriel $F$ dont on mesure la distance à $\vec{x}$.

Pour cela :

1. **Vous déterminez le produit scalaire**. Il est en général usuel. S'il ne l'est pas : **vous regardez la quantité au carré** $(\heartsuit - a\cdot\spadesuit - b\cdot\clubsuit - \cdots)^2$**, et vous la remplacez par un produit de deux éléments quelconques mais de « même nature »** (deux fonctions, deux coordonnées de vecteurs, deux polynômes, etc.), et le résultat obtenu est le produit scalaire qu'il vous faut (voir exemple 17 plus bas).

   On détermine le produit scalaire avant l'espace vectoriel $E$, ce qui peut paraître un peu étonnant. Ce serait uniquement dans votre démarche au brouillon : parce que c'est le plus facile à reconnaître et c'est ce qui guide le reste de votre réflexion.
2. **Le vecteur $\vec{x}\in E$ est donné par $\heartsuit$** : la seule quantité n'ayant pas en facteur un des paramètres $a$, $b$, etc.
3. Les deux étapes précédentes devraient vous permettre d'écrire la quantité à minimiser sous la forme : $\langle \vec{x} - a\cdot\spadesuit' - b\cdot\clubsuit' - \cdots, \vec{x} - a\cdot\spadesuit' - b\cdot\clubsuit' - \cdots\rangle$. **Vous prenez alors** $F = \left\{a\spadesuit' + b\clubsuit' + \cdots \mid (a,b,\dots)\in\mathbb{R}^\star\right\} = \mathrm{Vect}_{\mathbb{R}}\left((\spadesuit',\clubsuit',\dots)\right)$ C'est le sous-espace engendré par $\spadesuit'$, $\clubsuit'$, etc., c'est-à-dire le « Vect » de tous les vecteurs en facteur de $a$, $b$, etc.

   C'est parfois un sous-espace vectoriel connu, par exemple vous pouvez trouver $F = \mathrm{Vect}_{\mathbb{R}}(1,X) = \mathbb{R}_1[X]$, mais c'est inutile de le remarquer pour traiter convenablement le problème.

   Il arrive parfois que la borne inférieure ne soit pas prise sur tous les réels $a$, $b$, etc., mais seulement ceux vérifiant une certaine équation (je vais prendre $a+b=0$ pour l'exemple). Dans ce cas-là, ce n'est pas très grave. On prend à la place : $F = \left\{a\spadesuit' + b\clubsuit' + \cdots \mid (a,b,\dots)\in\mathbb{R}^\star ; a+b=0\right\}$. L'égalité nous permettra dans ce cas, au choix, de remarquer que $F$ est un hyperplan (un des cas particuliers fréquents dans $\mathrm{M}_n(\mathbb{R})$, voir section 3.2), ou d'exprimer un paramètre en fonction des autres (de sorte à se ramener à une description plus classique). L'exemple 16 l'illustre. → page 24
4. **Vous définissez $E$ comme un espace vectoriel usuel quelconque, tant qu'il contient tous les vecteurs en présence (fonctions, polynômes, $n$-uplets, etc.)**. Prendre l'espace vectoriel engendré par $\heartsuit$, $\spadesuit$ et $\clubsuit$ suffirait, mais il vaut mieux prendre un espace préhilbertien usuel pour invoquer sans scrupule les résultats qu'on y connaît.

   C'est important de faire attention s'il faut la convergence d'une intégrale ou d'une série : c'est ce qui pourra justifier de prendre $E = \mathrm{L}^2(I,\mathbb{R})\cap\mathrm{C}^0(I,\mathbb{R})$ plutôt que $E = \mathrm{C}^0(I,\mathbb{R})$ par exemple.

Après toutes ces transformations, on a effectivement :

$$\inf_{(a,b,\dots)\in\mathbb{R}^\star} \cdots (\heartsuit - a\cdot\spadesuit - b\cdot\clubsuit - \cdots)^2 \cdots = \inf_{\vec{y}\in F} \|\vec{x}-\vec{y}\|^2 = \mathrm{d}(\vec{x},F)^2,$$

et il reste à calculer cette distance grâce aux projections orthogonales : voir section 3.1. → page 22

**Exemple 15.** On cherche à calculer : $\inf_{(a,b)\in\mathbb{R}^2} \int_0^\pi (\ln(x) - a\cos(x) - b\sin(x))^2 \,\mathrm{d}x$ :

$E$ doit contenir ces trois fonctions, et l'intégrale doit converger : $E = \mathrm{L}^2(]0,\pi]) \cap \mathrm{C}^0(]0,\pi],\mathbb{R})$ (à cause du logarithme en 0)

$$\inf_{(a,b)\in\mathbb{R}^2} \int_0^\pi (\ln(x) - \underbrace{a\cos(x) - b\sin(x)}_{})^2 \,\mathrm{d}x.$$

produit scalaire usuel : $\langle f,g\rangle = \int_0^\pi f(x)g(x)\mathrm{d}x$

le seul vecteur présent pas dans $F$

$F$ contient ces fonctions : $F = \mathrm{Vect}_{\mathbb{R}}\,(\cos,\sin)$

Munissons $E = \mathrm{L}^2(]0,\pi]) \cap \mathrm{C}^0(]0,\pi],\mathbb{R})$ du produit scalaire défini par : $\forall (f,g) \in E^2,\ \langle f,g\rangle = \int_0^\pi f(x)g(x)\mathrm{d}x$, et notons $\|\cdot\|$ la norme euclidienne associée. Posons aussi $F = \mathrm{Vect}_{\mathbb{R}}\,((\cos,\sin))$. Alors :

$$\begin{aligned}\inf_{(a,b)\in\mathbb{R}^2} \int_0^\pi (\ln(x) - a\cos(x) - b\sin(x))^2 \,\mathrm{d}x &= \inf_{(a,b)\in\mathbb{R}^2} \langle \ln - a\cos - b\sin, \ln - a\cos - b\sin\rangle \\ &= \inf_{(a,b)\in\mathbb{R}^2} \|\ln - a\cos - b\sin\|^2 = \inf_{f\in F} \|\ln - f\|^2 = \mathrm{d}(\ln, F)^2.\end{aligned}$$

Nous avons ramené le problème à un calcul de distance à un sous-espace vectoriel. Notez bien ce qui a motivé le choix de $E$.

**Exemple 16.** On cherche à déterminer la borne inférieure $\inf_{\substack{(a_1\dots,a_n)\in\mathbb{R}^n\\ a_1+\dots+a_n=0}} \sum_{k=1}^n (k - a_k)^2$ :

$E$ doit contenir des $n$-uplets : $E = \mathbb{R}^n$ suffit.

ces coordonnées ne dépendent pas de $F$ : ce sont donc celles du vecteur dont on mesure la distance à $F$ : $\vec{b} = (b_k)_{k\in[\![1,n]\!]} = (k)_{k\in[\![1,n]\!]}$

$$\inf_{\substack{(a_1\dots,a_n)\in\mathbb{R}^n\\ a_1+\dots+a_n=0}} \sum_{k=1}^n (k - a_k)^2$$

on reconnaît facilement la norme usuelle sur $\mathbb{R}^n$ : $\|\vec{x}\|^2 = \sum_{k=1}^n x_i^2$, avec ici $\vec{x} = (k - a_k)_{k\in[\![1,n]\!]}$

$F$ contient les $n$-uplets $(a_1,\dots,a_n)$ vérifiant la condition : $\sum_{k=1}^n a_k = 0$

On munit $E = \mathbb{R}^n$ de la norme euclidienne usuelle $\|\cdot\|$. Posons aussi $F = \left\{(a_1,\dots,a_n) \in \mathbb{R}^n \mid \sum_{k=1}^n a_k = 0\right\}$, et $\vec{b} = (k)_{k\in[\![1,n]\!]} = (1,2,\dots,n) \in \mathbb{R}^n$. Alors :

$$\inf_{\substack{(a_1\dots,a_n)\in\mathbb{R}^n\\ a_1+\dots+a_n=0}} \sum_{k=1}^n (k - a_k)^2 = \inf_{\substack{(a_1\dots,a_n)\in\mathbb{R}^n\\ a_1+\dots+a_n=0}} \|\vec{b} - (a_1,\dots,a_n)\|^2 = \inf_{\vec{a}\in F} \|\vec{b} - \vec{a}\|^2 = \mathrm{d}(\vec{b}, F)^2.$$

Nous avons ramené le problème à un calcul de distance à un sous-espace vectoriel. Remarquez la différence avec les autres exemples, dans la définition de $F$. Notez aussi que c'est un hyperplan de $\mathbb{R}^n$ dont un vecteur normal est $(1,\dots,1)$, ce qui pourrait faciliter un calcul ultérieur de projection orthogonale.

**Exemple 17.** On cherche à déterminer la borne inférieure $\inf_{(a,b)\in\mathbb{R}^2} \int_{-\pi}^\pi \sin(x)\left(e^x - ax - bx^3\right)^2 \mathrm{d}x$ :

$E$ doit contenir ces trois fonctions : $E = \mathcal{C}^0([-\pi,\pi],\mathbb{R})$ (on n'a pas besoin que le sinus soit dans $E$ vu qu'il n'apparaîtra pas dans les normes, mais peu importe : on prend $E$ assez gros pour contenir tout le monde de toute façon)

le carré n'est pas sur le sinus, donc ce n'est pas la norme euclidienne de $x \mapsto \sin(x)\left(e^x - ax - bx^3\right)$ pour le produit scalaire usuel. Définissons plutôt : $\langle f,g\rangle = \int_{-\pi}^{\pi} \sin\cdot fg$. Là, le produit scalaire de $x \mapsto e^x - ax - bx^3$ par lui-même donne ce qu'on veut.

$$\left\{ \inf_{(a,b)\in\mathbb{R}^2} \int_{-\pi}^{\pi} \sin(x)\left(e^x - \underbrace{ax - bx^3}\right)^2 \mathrm{d}x. \right.$$

$F$ contient ces fonctions : $F = \mathrm{Vect}_{\mathbb{R}}\left(x\mapsto x, x\mapsto x^3\right)$

le seul vecteur présent pas dans $F$

Munissons $E = \mathcal{C}^0([-\pi,\pi],\mathbb{R})$ du produit scalaire défini par : $\forall (f,g)\in E^2$, $\langle f,g\rangle = \int_{-\pi}^{\pi} \sin(x) f(x) g(x)\mathrm{d}x$ (vérifiez que c'est effectivement un produit scalaire, avec une petite subtilité aux extrémités du segment pour le caractère défini), et notons $\|\cdot\|$ la norme euclidienne associée. Posons aussi $g : x \mapsto x^3$, et $F = \mathrm{Vect}_{\mathbb{R}}\left(\left(\mathrm{Id}_{[-\pi,\pi]}, g\right)\right)$. Alors :

$$\begin{aligned}
\inf_{(a,b)\in\mathbb{R}^2} \int_{-\pi}^{\pi} \sin(x)\left(e^x - ax - bx^3\right)^2 \mathrm{d}x &= \inf_{(a,b)\in\mathbb{R}^2} \left\langle \exp - a\mathrm{Id}_{[-\pi,\pi]} - bg, \exp - a\mathrm{Id}_{[-\pi,\pi]} - bg \right\rangle \\
&= \inf_{(a,b)\in\mathbb{R}^2} \left\| \exp - a\mathrm{Id}_{[-\pi,\pi]} - bg \right\|^2 \\
&= \inf_{f\in F} \|\exp - f\|^2 = \mathrm{d}(\exp, F)^2.
\end{aligned}$$

Nous avons ramené le problème à un calcul de distance à un sous-espace vectoriel. Notez bien qu'ici, nous n'avions pas un produit scalaire usuel : nous avons donc utilisé la méthode de « duplication » conseillée dans cette section, pour trouver le bon produit scalaire. Remarquez en particulier que le sinus est dans la définition du produit scalaire, et *uniquement* là. Pas dans $\left\langle \exp - a\mathrm{Id}_{[-\pi,\pi]} - bg, \exp - a\mathrm{Id}_{[-\pi,\pi]} - bg \right\rangle$ ni après.

**Mise en garde 4.** Vous remarquerez que ce qu'on nous demande n'est pas une distance en général, mais une distance au CARRÉ. Ne prenez donc pas la racine carrée à la fin, ou vous aurez très bêtement un résultat faux.

**Exercice 17.** Reconnaître $\inf_{(a,b)\in\mathbb{R}^2} \sum_{n=0}^{+\infty} \left(\frac{1}{4^n} - \frac{a}{2^n} - \frac{b(-1)^n}{2^n}\right)^2$ comme une distance au carré dans un espace vectoriel convenable (*attention, il n'est pas usuel, donc vous devrez bien réfléchir à chacune des étapes*). Vous pouvez essayer de la calculer si vous le désirez, mais ce n'est pas la compétence évaluée ici.

Nous donnons des exemples matriciels dans la section 3.2. → page 24

## 3.1 ✓ Calculs avec les projecteurs orthogonaux, les symétries orthogonales

Nous sommes en général amenés à expliciter une projection orthogonale pour un calcul de distance, en vertu de l'égalité $\mathrm{d}(\vec{x}, F) = \|\vec{x} - p(\vec{x})\|$.

Rappelons les quatre moyens proposés dans le cours pour calculer des projections orthogonales :

| Calcul de la projection orthogonale... | | | |
|---|---|---|---|
| Sur $D = \mathrm{Vect}_{\mathbb{R}}(\vec{n})$ | Sur $H = \vec{n}^{\perp}$ | On a une b.o.n $(\vec{v}_i)_i$ | On a une base « tout court » $(\vec{e}_i)_i$ |
| $p(\vec{x}) = \frac{\langle \vec{x}, \vec{n}\rangle}{\|\vec{n}\|^2}\vec{n}$ | $p(\vec{x}) = \vec{x} - \frac{\langle \vec{x}, \vec{n}\rangle}{\|\vec{n}\|^2}\vec{n}$ | $p(\vec{x}) = \sum_{i=1}^{k} \langle \vec{x}, \vec{v}_i\rangle \vec{v}_i$ | $\begin{cases} \forall i \in [\![1,k]\!], \quad \langle \vec{x} - p(\vec{x}), \vec{e}_i\rangle = 0, \\ p(\vec{x}) = \sum_{i=1}^{k} \alpha_i \vec{e}_i. \end{cases}$ |

Pensez à l'homogénéité des dimensions, comme en Physique, pour ne pas oublier de mettre au carré la norme au dénominateur ! Si on l'oublie, alors nous avons une longueur au cube divisé par une longueur :

cela donne une longueur au carré, ce qui est incompatible avec le fait que $p(\vec{x})$ soit un vecteur (et donc ait une longueur « tout court »).

Privilégiez TOUJOURS les deux premières expressions si la situation se présente. **Cela nécessite de regarder la dimension de l'espace sur lequel on projette avant toute chose**. De plus, dans le cas de la **distance** à un hyperplan, **ne calculez pas la différence $\vec{x} - \frac{\langle \vec{x}, \vec{n} \rangle}{\|\vec{n}\|^2}\vec{n}$, c'est inutile et une perte de temps**. En effet, dans le calcul de distance, on obtient ensuite :

$$\mathrm{d}(\vec{x}, H) = \|\vec{x} - p(\vec{x})\| = \left\| \cancel{\vec{x}} - \left( \cancel{\vec{x}} - \frac{\langle \vec{x}, \vec{n} \rangle}{\|\vec{n}\|^2}\vec{n} \right) \right\| = \frac{|\langle \vec{x}, \vec{n} \rangle|}{\|\vec{n}\|^2} \times \|\vec{n}\| = \frac{|\langle \vec{x}, \vec{n} \rangle|}{\|\vec{n}\|}.$$

Vous auriez perdu la simplification des $\vec{x}$ si vous aviez simplifié la différence préalablement.

Si nous ne sommes pas dans le cas particulier d'une projection sur une droite ou un hyperplan, nous revenons aux deux méthodes données dans le cours (page 22), qui correspondent aux deux dernières colonnes du tableau ci-dessus. Avec une petite variante dans le calcul de distance qui suit :

1. Si vous avez exprimé $p(\vec{x})$ grâce à une base orthonormée, disons sous la forme $p(\vec{x}) = \langle \vec{x}, \vec{v}_1 \rangle \vec{v}_1 + \langle \vec{x}, \vec{v}_2 \rangle \vec{v}_2$, vous calculez ces produits scalaires et vous avez des coordonnées explicitées $p(\vec{x}) = \alpha\vec{v}_1 + \beta\vec{v}_2$. VOUS NE REMPLACEZ SURTOUT PAS $\vec{v}_1$ ET $\vec{v}_2$ PAR LEURS EXPRESSIONS EXPLICITES ET VOUS LES LAISSEZ ! L'intérêt est que $(\vec{v}_1, \vec{v}_2)$ étant orthonormée, on a : $\|p(\vec{x})\|^2 = \alpha^2 + \beta^2$. Et donc :

$$\mathrm{d}(\vec{x}, F)^2 = \|\vec{x}\|^2 - \|p(\vec{x})\|^2 = \|\vec{x}\|^2 - \alpha^2 - \beta^2.$$

   C'est immédiat dans une base orthonormée et on doit en tirer profit. Il ne reste plus que $\|\vec{x}\|^2$ à calculer.

2. Si vous avez explicité $p(\vec{x})$ grâce à une base « tout court » et la résolution d'un système linéaire, alors la formule $\mathrm{d}(\vec{x}, F) = \|\vec{x}\|^2 - \|p(\vec{x})\|^2$ n'est pas toujours la plus pratique (si la norme revient à calculer une intégrale ou une somme compliquée). Une « astuce » est d'utiliser le fait que $p(\vec{x}) \in F$ et $\vec{x} - p(\vec{x}) \in F^\perp$ pour en déduire $\langle p(\vec{x}), \vec{x} - p(\vec{x}) \rangle = 0$, et donc, par linéarité du produit scalaire par rapport à sa première variable :

$$\mathrm{d}(\vec{x}, F)^2 = \|\vec{x} - p(\vec{x})\|^2 = \langle \vec{x} - p(\vec{x}), \vec{x} - p(\vec{x}) \rangle = \langle \vec{x}, \vec{x} - p(\vec{x}) \rangle - \underbrace{\langle p(\vec{x}), \vec{x} - p(\vec{x}) \rangle}_{=0} = \langle \vec{x}, \vec{x} - p(\vec{x}) \rangle.$$

   Calculer ce dernier produit scalaire est (légèrement) moins lourd.

| | Avantages | Défauts |
|---|---|---|
| 1re méthode | L'expression de la projection orthogonale est simple et totalement explicite, valable pour *tout* vecteur, et très pratique pour le calcul de normes (c'est « comme la norme usuelle » dans une base orthonormée), en particulier pour les calculs de distances. | Si on ne connaît pas de base orthonormée, il faut en produire une avec l'algorithme de Gram-Schmidt (si le calcul est faisable). C'est long et calculatoire si $\dim(F) \geqslant 3$. |
| 2e méthode | On évite le fastidieux algorithme de Gram-Schmidt, si on ne connaît pas de base orthonormée. Résoudre un système linéaire est facile et rapide. | Le calcul de norme ou de produit scalaire, pour calculer la distance d'un vecteur à $F$, s'avère plus technique.<br>De plus, si doit faire le calcul de projection pour plusieurs vecteurs, on doit résoudre autant de systèmes que de vecteurs : peu rentable à la longue. |

**Exercice 18.** Calculer les distances des exemples 16 et 17, éventuellement en variant les méthodes pour se convaincre des avantages et défauts présentés.

Il ne coûte pas cher de donner les formules analogues pour les symétries orthogonales. On sait en effet que la symétrie orthogonale par rapport à $F$ (notée $s$) et la projection orthogonale sur $F$ (notée $p$) sont reliées par la formule $s = 2p - \mathrm{Id}_E$ (*faire un dessin*). En partant du tableau ci-dessus, on en déduit ces expressions :

| Calcul de la symétrie orthogonale par rapport... | | | |
|---|---|---|---|
| à $D = \mathrm{Vect}_{\mathbb{R}}(\vec{n})$ | à $H = \vec{n}^{\perp}$ | On a une b.o.n $(\vec{v}_i)_i$ | On a une base « tout court » $(\vec{e}_i)_i$ |
| $s(\vec{x}) = 2\dfrac{\langle \vec{x}, \vec{n}\rangle}{\lVert\vec{n}\rVert^2}\vec{n} - \vec{x}$ | $s(\vec{x}) = \vec{x} - 2\dfrac{\langle \vec{x}, \vec{n}\rangle}{\lVert\vec{n}\rVert^2}\vec{n}$ | $s(\vec{x}) = 2\sum\limits_{i=1}^{k} \langle \vec{x}, \vec{v}_i\rangle \vec{v}_i - \vec{x}$ | $\begin{cases} \forall i \in [\![1,k]\!], \quad \langle \vec{x} - p(\vec{x}), \vec{e}_i\rangle = 0, \\ p(\vec{x}) = \sum\limits_{i=1}^{k} \alpha_i \vec{e}_i, \\ s(\vec{x}) = 2p(\vec{x}) - \vec{x}. \end{cases}$ |

Pour conclure, nous donnons deux cas, le premier étant très rarement rencontré dans ce contexte, où nous connaissons *déjà* (au sens où c'est du cours, même de première année) la décomposition EXPLICITE de tout vecteur dans la somme $E = F \oplus F^{\perp}$. On en déduit alors immédiatement les projections orthogonales sur chaque espace :

| Décomposition | Toute fonction est somme d'une fonction paire et d'une fonction impaire | Toute matrice est somme d'une matrice symétrique et d'une matrice antisymétrique. |
|---|---|---|
| Formellement | $E = \text{Paires} \oplus \text{Impaires}$ | $\mathrm{M}_n(\mathbb{R}) = \mathrm{S}_n(\mathbb{R}) \oplus \mathrm{A}_n(\mathbb{R})$ |
| Décomposition explicite | $\forall x \in I,\, f(x) = \dfrac{f(x)+f(-x)}{2} + \dfrac{f(x)-f(-x)}{2}$ | $A = \dfrac{1}{2}\left(A + A^{\top}\right) + \dfrac{1}{2}\left(A - A^{\top}\right)$ |

**Exemple 18.** En citant sans les redémontrer les propriétés de ce tableau : si l'on munit $\mathbb{R}_n[X]$ du produit scalaire $(P,Q) \mapsto \displaystyle\int_{-1}^{1} P(t)Q(t)\mathrm{d}t$, et si $F$ est l'ensemble des polynômes $P \in \mathbb{R}_n[X]$ tels que $P(-X) = P(X)$, alors $F^{\perp}$ est l'ensemble des polynômes $P \in \mathbb{R}_n[X]$ tels que $P(-X) = -P(X)$, par une démonstration analogue à celle vue en cours, et de plus la décomposition est explicite :

$$\forall P \in \mathbb{R}_n[X], \quad P = \underbrace{\frac{1}{2}\left(P(X) + P(-X)\right)}_{\in F} + \underbrace{\frac{1}{2}\left(P(X) - P(-X)\right)}_{\in F^{\perp}}.$$

On en déduit que la projection orthogonale de tout polynôme $P \in \mathbb{R}_n[X]$ sur $F$ est : $p(P) = \dfrac{1}{2}\left(P(X) + P(-X)\right)$. La distance de $P \in \mathbb{R}_n[X]$ à $F$ est alors :

$$\mathrm{d}(P,F) = \lVert P - p(P)\rVert = \frac{1}{2}\lVert P(X) - P(-X)\rVert = \frac{1}{2}\sqrt{\int_{-1}^{1} (P(t) - P(-t))^2 \mathrm{d}t}.$$

**Exercice 19.** Déterminer la dimension de $F$ et de $F^{\perp}$. Remarquer qu'on sait facilement écrire la matrice de $p$ dans la base canonique, et donc en déduire $\dim(F) = \mathrm{rang}(p) = \mathrm{tr}(p)$ (sachez justifier ces deux égalités).

Nous donnons aussi l'exemple beaucoup plus fréquent de $\mathrm{S}_n(\mathbb{R})$ et $\mathrm{A}_n(\mathbb{R})$ dans la prochaine section.

### 3.2 Cas particulier fréquent : distance dans $\mathrm{M}_n(\mathbb{R})$

Lorsqu'on vous donne à calculer une distance à un sous-espace vectoriel $F$ de $\mathrm{M}_n(\mathbb{R})$ (pour le produit scalaire et la norme euclidienne usuels), on vous place en général dans l'un de ces trois cas :

1. Avec $n$ « **petit** » et $F$ défini explicitement (de sorte qu'on pourrait en donner une base facilement).
2. Avec $F$ défini par une *équation* linéaire, où apparaît soit la **trace**, soit une **somme sur des coefficients**.
3. Avec $F = \mathrm{S}_n(\mathbb{R})$ ou $F = \mathrm{A}_n(\mathbb{R})$.

Calculer la distance de $M \in \mathrm{M}_n(\mathbb{R})$ à $F$ revient à calculer la projection orthogonale $p_F(M)$. Selon le cas :

1. **On écrit $F$ comme un « Vect », de sorte à pouvoir exprimer $p_F(M)$ dans cette base avec la méthode connue.**
   Il est même souvent facile d'obtenir une base **orthonormée** de $F$, ce qui est très pratique pour calculer $p_F(M)$. Pour vérifier qu'elle l'est, n'oubliez que le produit scalaire sur $\mathrm{M}_n(\mathbb{R})$ est en vérité très simple : c'est la somme des produits des coefficients !
2. **Dans ce cas $F$ est un hyperplan, et il suffit d'en expliciter un vecteur normal pour calculer facilement $p_F(M)$ et $\mathrm{d}(M, F)$.**
   Si l'équation est de la forme $\mathrm{tr}(AM) = 0$, alors un vecteur normal est $N = A^\top$ (vu que $\left\langle A^\top, M \right\rangle = \mathrm{tr}((A^\top)^\top M) = \mathrm{tr}(AM)$). Si l'équation est de la forme $\sum\limits_{i,j} \alpha_{i,j} m_{i,j} = 0$, où les $\alpha_{i,j}$ sont des réels ne dépendant pas de $M$, alors un vecteur normal est $N = ((\alpha_{i,j}))_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{R})$ (on a bien $\langle N, M \rangle = \sum\limits_{i,j} \alpha_{i,j} m_{i,j}$ : c'est la somme des produits des coefficients).
3. **Si $F = \mathrm{S}_n(\mathbb{R})$ ou $F = \mathrm{A}_n(\mathbb{R})$, alors on se souvient que la décomposition de toute matrice dans la somme directe $\mathrm{M}_n(\mathbb{R}) = F \oplus F^\perp$ est explicite** et issue de l'égalité $M = \dfrac{1}{2}\left(M + M^\top\right) + \dfrac{1}{2}\left(M - M^\top\right)$, nous l'avons rappelé dans la section précédente. Ainsi on a simplement, si $F = \mathrm{A}_n(\mathbb{R})$ : $p_F(M) = \dfrac{1}{2}\left(M - M^\top\right)$, et : $\mathrm{d}(M, F) = \|M - p_F(M)\| = \frac{1}{2}\|M + M^\top\|$. Expression analogue si $F = \mathrm{S}_n(\mathbb{R})$.

**Exemple 19. (premier cas)** Calculons la distance de $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$ au sous-espace vectoriel :

$$F = \left\{ \begin{pmatrix} a & b \\ a & -b \end{pmatrix} \mid (a, b) \in \mathbb{R}^2 \right\}.$$

On a facilement :

$$F = \left\{ \begin{pmatrix} a & b \\ a & -b \end{pmatrix} \mid (a, b) \in \mathbb{R}^2 \right\} = \left\{ a \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \mid (a, b) \in \mathbb{R}^2 \right\} = \mathrm{Vect}_{\mathbb{R}} \left( \left( \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right) \right).$$

Ainsi $\left( \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right)$ est une base de $F$, même une base orthogonale (c'est immédiat), et il suffit de rendre les matrices unitaires pour avoir une base orthonormée. Ainsi $(V_1, V_2) = \left( \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \dfrac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right)$ est une base orthonormée de $F$. On en déduit que la projection orthogonale de $A$ sur $F$ est :

$$p(A) = \langle A, V_1 \rangle V_1 + \langle A, V_2 \rangle V_2 = \frac{3}{\sqrt{2}} V_1 - \frac{2}{\sqrt{2}} V_2 = \frac{3}{\sqrt{2}} V_1 - \sqrt{2} V_2.$$

On en déduit : $\mathrm{d}(A, F)^2 = \|A\|^2 - \|p(A)\|^2 = (1^2 + 2^2 + 2^2 + 4^2) - \left(\dfrac{3}{\sqrt{2}}\right)^2 - \left(\sqrt{2}\right)^2 = \dfrac{37}{2}$, puis : $\mathrm{d}(A, F) = \sqrt{\frac{37}{2}}$.

**Exemple 20. (deuxième cas)** Calculons la distance de $A = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in \mathrm{M}_n(\mathbb{R})$ au sous-espace vectoriel $F = \{M \in \mathrm{M}_n(\mathbb{R}) \mid \mathrm{tr}(M) = 0\}$. On note que $\mathrm{I}_n$ est un vecteur normal de $F$, puisque : $\forall M \in \mathrm{M}_n(\mathbb{R})$, $\langle \mathrm{I}_n, M \rangle = \mathrm{tr}(\mathrm{I}_n^\top M) = \mathrm{tr}(M)$, donc : $F = \{M \in \mathrm{M}_n(\mathbb{R}) \mid \langle \mathrm{I}_n, M \rangle = 0\} = \mathrm{Vect}_{\mathbb{R}}(\mathrm{I}_n)^\perp$.

On en déduit que la projection orthogonale de $A$ sur $F$ est :

$$p(A) = A - \frac{\langle A, \mathrm{I}_n\rangle}{\|\mathrm{I}_n\|^2}\mathrm{I}_n = A - \frac{n}{n}\mathrm{I}_n = A - \mathrm{I}_n,$$

puis : $\mathrm{d}(A,F) = \|A - p(A)\| = \|\mathrm{I}_n\| = \sqrt{n}$. Plus généralement : $\forall M \in \mathrm{M}_n(\mathbb{R})$, $\mathrm{d}(M,F) = \frac{|\langle M, \mathrm{I}_n\rangle|}{\|\mathrm{I}_n\|} = \frac{|\mathrm{tr}(M)|}{\sqrt{n}}$.

**Exercice 20.** Procéder de même pour obtenir la distance de $A$ à :

$$F = \left\{ M = ((m_{i,j}))_{1\leqslant i,j\leqslant n} \mid \sum_{1\leqslant i<j\leqslant n} m_{i,j} = 0 \right\}$$

(c'est l'ensemble des matrices dont la somme des coefficients strictement au-dessus de la diagonale est nulle).

# 4 Utilisation des endomorphismes autoadjoints

Les endomorphismes autoadjoints sont des objets extrêmement riches, bien plus que ne le suggère le temps qu'on y consacre en cours ; ils peuvent donc occuper, dans un sujet de concours, une place non proportionnelle à celle dans le programme de PSI.

Les trois prochaines sections (de *Généraliser des produits scalaires* à *Matrices positives, définies positives*) sont très liées entre elles. Vous êtes très fortement encouragés à les lire à la suite, si vous voulez du recul et une vue d'ensemble parfaite sur ce qui motive *réellement* les endomorphismes autoadjoints.

Le résultat suivant est à savoir démontrer IMPÉRATIVEMENT avant cette lecture, ou elle est inutile :

> La décomposition $\langle f(\vec{x}), \vec{y}\rangle = \sum_{i=1}^{n} \lambda_i x_i y_i$ ou $X^\top AY = \sum_{i=1}^{n} \lambda_i x_i y_i$ (attention à ne pas retenir bêtement et de travers ce que sont les $x_i$ et $y_i$ ici : CE NE SONT PAS LES COORDONNÉES DE $X$ ET $Y$ DANS LA BASE CANONIQUE).

## 4.1 Généraliser des produits scalaires

Si $f$ est un endomorphisme *autoadjoint* d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$, et $A$ une matrice *symétrique*, alors :

$$(\vec{x}, \vec{y}) \mapsto \langle \vec{x}, f(\vec{y})\rangle, \quad \text{et } (X,Y) \mapsto X^\top AY$$

sont des formes bilinéaires symétriques sur $E$ et $\mathrm{M}_{n,1}(\mathbb{R})$ respectivement : pour le cas de l'endomorphisme autoadjoint, c'est essentiellement la définition (qui assure la... symétrie). Pour la matrice symétrique, c'est un exercice.

D'ailleurs, TOUTES les formes bilinéaires symétriques sur $\mathrm{M}_{n,1}(\mathbb{R})$ sont de cette forme :

**Exercice 21.** Soit $\varphi$ une forme bilinéaire symétrique sur $\mathrm{M}_{2,1}(\mathbb{R})$. On pose $E_1 = \begin{pmatrix}1\\0\end{pmatrix}$ et $E_2 = \begin{pmatrix}0\\1\end{pmatrix}$.

1. Vérifier qu'on a, pour tout $(\lambda_1, \lambda_2, \mu_1, \mu_2) \in \mathbb{R}^4$ :

$$\varphi(\lambda_1 E_1 + \lambda_2 E_2, \mu_1 E_1 + \mu_2 E_2) = \lambda_1\mu_1\varphi(E_1,E_1) + \lambda_1\mu_2\varphi(E_1,E_2) + \lambda_2\mu_1\varphi(E_2,E_1) + \lambda_2\mu_2\varphi(E_2,E_2).$$

2. On pose : $A = \begin{pmatrix}\varphi(E_1,E_1) & \varphi(E_1,E_2)\\ \varphi(E_2,E_1) & \varphi(E_2,E_2)\end{pmatrix}$. Montrer que $A$ est symétrique, et qu'on a :

$$\forall (X,Y) \in \mathrm{M}_{2,1}(\mathbb{R}), \quad \varphi(X,Y) = X^\top AY.$$

3. Généraliser ce qui précède pour démontrer que toute forme bilinéaire symétrique sur $\mathrm{M}_{n,1}(\mathbb{R})$ est de la forme $(X,Y) \mapsto X^\top AY$ où $A$ est une matrice symétrique réelle d'ordre $n$.

Pour voir ce qu'il manque à une telle forme bilinéaire symétrique pour être un produit scalaire, voir la section 4.3 (*Matrices positives, définies positives)*, et plus particulièrement l'exercice 24. → page 28

Une démonstration plus élégante du résultat de cet exercice découle du théorème de représentation de Riesz, appliqué pour tout $\vec{y} \in E$ (fixé) à la forme linéaire $\vec{x} \mapsto \varphi(\vec{x}, \vec{y})$. Je n'en parlerai pas, car là n'est pas ma préoccupation dans cette section.

Une forme bilinéaire symétrique n'est pas tout à fait aussi intéressante qu'un produit scalaire, mais on peut déjà en tirer quelques formules. En effet, les identités *calculatoires* vérifiées par un produit scalaire n'utilisent pas le caractère défini ni positif. On s'en convainc en faisant l'exercice suivant :

**Exercice 22.** Soit $f$ un endomorphisme autoadjoint de $(E, \langle\cdot,\cdot\rangle)$. Montrer les propriétés suivantes :

1. L'identité du parallélogramme :
$$\forall(\vec{x},\vec{y}) \in E^2, \quad \langle f(\vec{x}+\vec{y}), \vec{x}+\vec{y}\rangle + \langle f(\vec{x}-\vec{y}), \vec{x}-\vec{y}\rangle = 2\left(\langle f(\vec{x}), \vec{x}\rangle + \langle f(\vec{y}), \vec{y}\rangle\right).$$
2. Les identités de polarisation ; pour tout $(\vec{x},\vec{y}) \in E^2$, on a :
$$\langle f(\vec{x}), \vec{y}\rangle = \frac{1}{4}\left(\langle f(\vec{x}+\vec{y}), \vec{x}+\vec{y}\rangle - \langle f(\vec{x}-\vec{y}), \vec{x}-\vec{y}\rangle\right) = \frac{1}{2}\left(\langle f(\vec{x}+\vec{y}), \vec{x}+\vec{y}\rangle - \langle f(\vec{x}), \vec{x}\rangle - \langle f(\vec{y}), \vec{y}\rangle\right).$$
3. Refaire l'exercice en remplaçant l'étude de $(\vec{x},\vec{y}) \mapsto \langle f(\vec{x}), \vec{y}\rangle$ par celle de $(X,Y) \mapsto X^\top AY$ sur $\mathrm{M}_{n,1}(\mathbb{R})$, où $A$ est une matrice symétrique réelle d'ordre $n$.

Prenez bien soin d'identifier où sert la propriété d'être symétrique, et si vous avez du temps : déterminer ce que l'on aurait à la place en enlevant cette hypothèse.

Nous utilisons abondamment les identités de polarisation, parce que nous aurons souvent des informations sur les produits scalaires de la forme $\langle f(\vec{x}), \vec{x}\rangle$, au lieu de $\langle f(\vec{x}), \vec{y}\rangle$, et nous voudrons nous y ramener (ce sont en effet ces derniers produits scalaires qui nous permettent de tirer profit de l'orthogonalité, ou d'en déduire la matrice de $f$ dans une base orthonormée) : voir la section 1.4 (*Comment passer de* $\langle\vec{x},\vec{z}\rangle = \langle\vec{y},\vec{z}\rangle$ *à* $\vec{x} = \vec{y}$). ← page 10

**Culture scientifique.** On peut se demander : à quoi bon étudier ces formes bilinéaires symétriques, abstraites et n'ayant même pas l'avantage de l'interprétation géométrique, à l'instar d'un produit scalaire ? En vérité, elles n'ont rien d'abstrait : **dès qu'il apparaît une combinaison linéaire de carrés et de produits de coordonnées, il apparaît une quantité de la forme $\langle\vec{x}, f(\vec{y})\rangle$ ou $X^\top AY$.**

Par exemple, si j'écris l'équation d'une ellipse sous la forme :
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$
je peux la réécrire sous la forme :
$$X^\top AX = 1,\ \text{où : } X = \begin{pmatrix} x \\ y \end{pmatrix},\ \text{et : } A = \begin{pmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{pmatrix}.$$
On peut de même transformer l'équation d'un certain cône d'axe de révolution $(Oz)$ dans $\mathbb{R}^3$ :
$$x^2+y^2 = z^2 \iff x^2+y^2-z^2 = 0 \iff X^\top AX = 0,\ \text{où : } X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ \text{et : } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
Ce sont donc des sujets d'étude CENTRAUX de la géométrie. D'ailleurs, le théorème spectral, conjointement à la section suivante, permettent de démontrer que toute équation de degré 2 en $x$ et $y$ décrit soit une ellipse, soit une parabole, soit une hyperbole (sauf cas pathologiques : on parle alors de conique *dégénérée*), avec des résultats analogues en dimension 3.

Pour vraiment avoir un produit scalaire (et donc profiter de la positivité, du caractère défini, de l'inégalité de Cauchy-Schwarz, etc.), il nous faut un endomorphisme ou une matrice avec des valeurs propres *strictement positives* : voir la section 4.3 (*Matrices positives, définies positives*), et plus particulièrement l'exercice 24. **Mais même si ce n'est pas un produit scalaire, le théorème spectral nous permet d'étudier** $(\vec{x}, \vec{y}) \mapsto \langle \vec{x}, f(\vec{y}) \rangle$ **ou** $(X, Y) \mapsto X^\top AY$ **comme si c'en était « presque un »**. C'est l'objet de la section suivante. → page 28

## 4.2 ✓ Théorème spectral et produits scalaires

À l'aide du théorème spectral, nous avons démontré dans le cours que si $f$ est un endomorphisme autoadjoint d'un espace euclidien $(E, \langle \cdot, \cdot \rangle)$, dont $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base orthonormée constituée de vecteurs propres de $f$ (respectivement associés à des valeurs propres $\lambda_1, \ldots, \lambda_n$), alors pour tous vecteurs $\vec{x}$ et $\vec{y}$ de $E$ de coordonnées $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans $(\vec{e}_1, \ldots, \vec{e}_n)$, on a :

$$\langle f(\vec{x}), \vec{y} \rangle = \sum_{i=1}^{n} \lambda_i x_i y_i.$$

C'EST À SAVOIR REDÉMONTRER À CHAQUE FOIS ! En particulier, si $\vec{x} = \vec{y}$, alors : $\langle f(\vec{x}), \vec{x} \rangle = \sum\limits_{i=1}^{n} \lambda_i x_i^2$. C'est « presque » comme le produit scalaire usuel, et très confortable pour le calcul.

Nous avons une formule analogue pour les matrices, qu'on peut déduire directement de l'expression d'un produit scalaire dans une base orthonormée pour écrire $X^\top AY = \langle f(\vec{x}), \vec{y} \rangle = \sum\limits_{i=1}^{n} \lambda_i x_i y_i$, mais il est bon de savoir la démontrer directement au lieu de passer systématiquement par un endomorphisme associé. C'EST À SAVOIR REDÉMONTRER À CHAQUE FOIS ! C'est le point de départ de nombreux problèmes sur les endomorphismes et matrices symétriques, il faut y penser et même l'écrire « dans le doute » (au brouillon) quand on ne sait pas répondre à une question.

Le cadre d'application de ces identités est : **dès que vous devez déduire des propriétés des éléments propres de $A$ à partir de produits scalaires, et inversement**. Il s'agit en général d'inégalités. Cela permet aussi de produire des encadrements très efficaces de produits scalaires, à condition de connaître un minimum de choses sur les valeurs propres.

## 4.3 Matrices positives, définies positives

Le cours définit la notion de matrice positive, et prouve l'équivalence suivante :

$$\boxed{\forall X \in \mathrm{M}_{n,1}(\mathbb{R}),\ X^\top AX \geqslant 0 \iff \mathrm{Sp}_{\mathbb{R}}(A) \subseteq \mathbb{R}_+ \text{ (matrice \textbf{positive})}.}$$

Notez la ressemblance de la première propriété avec la propriété de positivité du produit scalaire. Nous introduisons aussi la notion de matrice définie positive :

$$\boxed{\begin{array}{c}\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \{0_{\mathrm{M}_{n,1}(\mathbb{R})}\},\ X^\top AX > 0 \iff \mathrm{Sp}_{\mathbb{R}}(A) \subseteq \mathbb{R}_+^* \iff A \\ \text{inversible et } \mathrm{Sp}_{\mathbb{R}}(A) \subseteq \mathbb{R}_+ \text{ (matrice \textbf{définie positive})}.\end{array}}$$

Notez la ressemblance de la première propriété avec le caractère défini du produit scalaire. La notion existe aussi pour les endomorphismes.

**Exercice 23.** Soit $f$ un endomorphisme autoadjoint d'un espace euclidien $(E, \langle \cdot, \cdot \rangle)$.

1. Montrer : $\mathrm{Sp}(f) \subseteq \mathbb{R}_+ \iff \forall \vec{x} \in E,\ \langle \vec{x}, f(\vec{x}) \rangle \geqslant 0$. On dit dans ce cas que $f$ est **positif**.
2. Montrer : $\mathrm{Sp}(f) \subseteq \mathbb{R}_+^* \iff \forall \vec{x} \in E \setminus \{\vec{0}\},\ \langle \vec{x}, f(\vec{x}) \rangle > 0$. On dit dans ce cas que $f$ est **défini positif**.

Cet exercice se traite EXACTEMENT de la même manière que dans le cas matriciel : en exprimant $\langle \vec{x}, f(\vec{x}) \rangle$ comme combinaison linéaire dépendant des valeurs propres de $f$ et des coordonnées de $\vec{x}$ au carré.

Absolument TOUT problème de concours sur les endomorphismes autoadjoints utilisera ces équivalences, dans le cas matriciel ou géométrique. Par conséquent :

**CES CARACTÉRISATIONS DES ENDOMORPHISMES ET MATRICES (DÉFINIS) POSITIFS SONT À CONNAÎTRE IMPÉRATIVEMENT !**

Notez que si vous savez que $A$ est une matrice positive, alors il vient avec d'autres quantités positives :

— la trace (puisque c'est la somme des valeurs propres avec multiplicités, qui sont positives) ;

— le déterminant (puisque c'est le produit des valeurs propres avec multiplicités, qui sont positives) ;

— *certains* coefficients de $A$ (puisque $X^\top AX \geqslant 0$ pour tout $X$, c'est en particulier si $X$ est le $i^{\text{e}}$ vecteur de la base canonique $E_i$ : quel coefficient de $A$ obtient-on alors en calculant $E_i^\top AE_i$ ?).

Un intérêt de cette notion est partiellement contenu dans l'exercice suivant (justifiant la terminologie) :

**Exercice 24.**

1. Montrer que si $f$ est un endomorphisme autoadjoint d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$, et si $f$ est défini positif, alors l'application $(\vec{x}, \vec{y}) \mapsto \langle f(\vec{x}), \vec{y}\rangle$ est un produit scalaire sur $E$.
2. Montrer que si $A \in \mathrm{S}_n(\mathbb{R})$ est une matrice définie positive, alors l'application $(X, Y) \mapsto X^\top AY$ est un produit scalaire sur $\mathrm{M}_{n,1}(\mathbb{R})$.

En fait, on peut démontrer que TOUT produit scalaire sur $\mathrm{M}_{n,1}(\mathbb{R})$ est de cette forme (exercice 21). ← page 26

C'est en particulier TRÈS utile pour pouvoir utiliser à nouveau la propriété de positivité ou le caractère défini : vous en déduisez que si $\langle f(\vec{x}), \vec{x}\rangle = 0$ (ou $X^\top AX = 0$), alors $\vec{x} = \vec{0}$ (ou $X = 0_{\mathrm{M}_{n,1}(\mathbb{R})}$). Même si $f$ (ou $A$) n'est pas définie positive, mais seulement positive, alors ce n'est pas inutile pour autant : l'égalité $\langle f(\vec{x}), \vec{x}\rangle = 0$ avec $\vec{x} \neq \vec{0}$ implique que 0 est une valeur propre de $f$. C'est un des exercices de travaux dirigés.

Ainsi nous en déduisons de nouvelles informations sur ces produits scalaires, sachant que (nous ne le répéterons jamais assez) quand on les évalue en des vecteurs d'une base orthonormée, ils donnent les coefficients de la matrice de $f$ dans cette même base. On peut donc étudier la matrice de $f$ (et $f$ par extension) à l'aide de produits scalaires de la forme $\langle f(\vec{x}), \vec{y}\rangle$ (n'oublions pas les identités de polarisation, au besoin). De même pour $A$.

**Exemple 21.** Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice définie positive. Montrons qu'alors : $a > 0$ et $ad - bc > 0$. En effet, puisque ses valeurs propres $\lambda_1$ et $\lambda_2$ sont strictement positives, on a : $\det(A) = \lambda_1\lambda_2 > 0$. Or $\det(A) = ad - bc$, donc $ad - bc > 0$. De plus, $a = E_1^\top AE_1 > 0$, où $E_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (on utilise le fait que tous les coefficients d'une matrice puissent s'écrire en termes de produits scalaires : revoir au besoin la section 2.3), d'où le résultat. ← page 16

**Exercice 25.** Montrer que pour une matrice symétrique définie positive, tous les coefficients diagonaux sont strictement positifs, mais que ce n'est pas forcément vrai de tous ses coefficients.

Mais un produit scalaire vérifie bien plus de propriétés que la seule positivité et le caractère défini :

**Exercice 26.**

1. On suppose que $f$ est un endomorphisme autoadjoint de $(E, \langle\cdot,\cdot\rangle)$, **positif**.
   (a) Montrer que pour tout $(\vec{x}, \vec{y}) \in E^2$, on a :
   $$(\langle f(\vec{x}), \vec{y}\rangle)^2 \leqslant \langle f(\vec{x}), \vec{x}\rangle \langle f(\vec{y}), \vec{y}\rangle \quad \text{(inégalité de Cauchy-Schwarz).}$$
   (b) Montrer que pour tout $(\vec{x}, \vec{y}) \in E^2$, on a :
   $$\sqrt{\langle f(\vec{x} + y), \vec{x} + \vec{y}\rangle} \leqslant \sqrt{\langle f(\vec{x}), \vec{x}\rangle} + \sqrt{\langle f(\vec{y}), \vec{y}\rangle} \quad \text{(inégalité triangulaire, ou de Minkowski).}$$

On pourrait facilement adapter la première question à la démonstration de l'inégalité de Cauchy-Schwarz vérifiée par le coefficient de corrélation, dans le chapitre de probabilités. Pour cela, adaptez-la à l'application $(X,Y) \mapsto \mathrm{E}(XY)$, qui est une forme bilinéaire symétrique positive sur l'espace vectoriel des variables aléatoires réelles discrètes admettant une variance.

2. Soit $A \in \mathrm{S}_n(\mathbb{R})$ **positive**. Montrer : $\forall (A,B) \in (\mathrm{M}_{n,1}(\mathbb{R}))^2$, $\left|X^\top AY\right| \leqslant \sqrt{X^\top AX \cdot Y^\top AY}$.

Notez qu'on ne dit rien des cas d'égalité (pour cela, il faudrait avoir un produit scalaire, et donc que $f$ et $A$ soient définis positifs d'après l'exercice 24 plus haut).

On en déduit également de nouvelles formes de réduction des matrices, par exemple en orthonormalisant *dans ces nouveaux produits scalaires* des bases qui étaient orthonormées *pour les produits scalaires usuels*. Par exemple, toute matrice symétrique $A$ définie positive peut être écrite de manière unique sous la forme $A = T^\top T$, où $T$ est une matrice triangulaire supérieure à coefficients diagonaux strictement positifs. Mais je n'en parle pas car nous déborderions beaucoup trop du programme de PSI.

### 4.3.1 Montrer (ou infirmer) qu'une matrice symétrique est positive ou définie positive

Soit $A \in \mathrm{S}_n(\mathbb{R})$. On se demande si $A$ est positive, définie positive, ou ni l'un ni l'autre. Il y a deux façons de l'étudier : soit avec le signe de $X^\top AX$, soit avec le signe des valeurs propres. Nous discutons ici de la méthode à privilégier, sachant qu'il faut avoir en tête que le bon choix à faire dépend de la facilité que l'on a à avoir les valeurs propres.

**Cas d'une matrice explicite d'ordre raisonnable (2 ou 3)** Pour une matrice symétrique $A$ d'ordre 2, il est toujours possible d'obtenir ses valeurs propres aisément : cela revient à déterminer les racines de son polynôme caractéristique, qui est de degré 2. C'est donc classique. Si $A$ est d'ordre 3, cela dépend car il est tout à fait possible qu'on ne soit pas en présence de trois racines aisément trouvables.

Si vous avez réussi à trouver les valeurs propres de $A$, alors il suffit de déterminer leur signe pour en déduire si $A$ est positive ou définie positive.

Si vous voyez que vous peinez à trouver les valeurs propres, alors vous oubliez la caractérisation avec les valeurs propres et montrez que pour tout vecteur colonne $X$, on a : $X^\top AX \geqslant 0$ (ou $X^\top AX > 0$ pour tout $X$ non nul, si vous voulez montrer que $A$ est définie positive). Pour y parvenir, il suffit de poser $X = \begin{pmatrix} x \\ y \end{pmatrix}$ (ou $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ si $A$ est d'ordre 3), et vous calculez explicitement $X^\top AX$ en fonction de $x$ et $y$ (et $z$ le cas échéant). Il vous suffira alors d'utiliser la méthode de la section 1.1.1 pour montrer que c'est positif.

Pour montrer que c'est **défini** positif par cette méthode, privilégiez la contraposée : montrez que si $X^\top AX = 0$, alors $X = 0_{\mathrm{M}_{n,1}(\mathbb{R})}$. Cela veut bien dire que si $X$ est non nul, alors $X^\top AX > 0$ (en partant du principe que vous avez déjà montré la positivité de cette quantité).

**Exemple 22.** Soit $A = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}$. C'est une matrice symétrique. On se demande si elle est (définie) positive. Je vous laisse vous convaincre (ordinateur à l'appui) que les racines de $\chi_A$ ne sont pas simples du tout à expliciter : vous n'y parviendriez certainement pas à la main ! Puisque nous sommes dans l'incapacité d'expliciter les valeurs propres de $A$ à la main, nous allons plutôt passer par le calcul de

$X^\top AX$. Soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{M}_{3,1}(\mathbb{R})$. On a :

$$\begin{aligned}
X^\top AX = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 5x^2 + 3y^2 + 4z^2 + 4xy + 2xz + 4yz \\
&= 5\left(x + \frac{2y}{5} + \frac{z}{5}\right)^2 + \frac{71}{25}y^2 + \frac{99}{25}z^2 + \frac{96}{25}yz \\
&= 5\left(x + \frac{2y}{5} + \frac{z}{5}\right)^2 + \frac{71}{25}\left(y^2 + \frac{96}{71}yz + \frac{99}{71}z^2\right) \\
&= 5\left(x + \frac{2y}{5} + \frac{z}{5}\right)^2 + \frac{71}{25}\left[\left(y + \frac{48}{71}z\right)^2 + \left(99 - \frac{48^2}{71^2}\right)z^2\right].
\end{aligned}$$

Nous avons une somme de carrés (multipliés par des scalaires positifs), donc : $X^\top AX \geqslant 0$. De plus, si $X^\top AX = 0$ alors, une somme de réels positifs étant nul si seulement si chaque terme est nul, on en déduit : $x + \frac{2y}{5} + \frac{z}{5} = 0$, $y + \frac{48}{71}z = 0$ et $z = 0$. De là il découle aisément : $x = y = z = 0$, donc : $X = 0_{\mathrm{M}_{3,1}(\mathbb{R})}$. Par contraposée, si $X$ est non nul, alors : $X^\top AX > 0$, ce qui montre que $A$ est **définie** positive. Nous y sommes parvenus sans calculer les valeurs propres.

**Cas d'une matrice d'ordre $n$** Il est dans ce cas assez rare que le calcul de $X^\top AX$ soit suffisamment simple pour en extraire le signe. Privilégiez la recherche des valeurs propres, en n'oubliant pas que grâce aux méthodes du chapitre de réduction des endomorphismes, on sait parfois trouver toutes les valeurs propres d'une matrice d'ordre $n$ par un usage adéquat du théorème du rang, de la trace, etc. Il vaut mieux, parce que le calcul du polynôme caractéristique d'une matrice d'ordre $n$ n'est pas forcément aisé.

**Exemple 23.** On veut déterminer si la matrice $A = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}$ est symétrique positive. Si l'on cherche à poser $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et à calculer $X^\top AX$, on trouvera certes l'expression : $X^\top AX = 2\sum_{1\leqslant i<j\leqslant n} x_i x_j$, mais bon courage pour en trouver le signe... La méthode de la section 1.1.1 se met très difficilement en œuvre. En revanche les valeurs propres de $A$ s'obtiennent aisément : en effet $A + \mathrm{I}_n$ est de rang 1 (toutes les colonnes sont proportionnelles à $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$), donc par le théorème du rang : $\dim(\ker(A + \mathrm{I}_n)) = n - 1$. On en déduit que $-1$ est valeur propre (d'ordre de multiplicité $n-1$), ce qui suffit à démontrer que $A$ n'est pas une matrice positive (elle admet une valeur propre strictement négative).

On peut montrer grâce à la somme des coefficients de chaque ligne, ou à la trace, que la dernière valeur propre de $A$ est $n - 1$. Ainsi $A$ admet une valeur propre de chaque signe, donc elle n'est ni positive, ni négative.

**Cas d'une matrice symétrique définie à l'aide d'une autre matrice symétrique (puissances...)** Dans le cas où l'on sait que $A$ est symétrique positive (par exemple), et qu'on nous demande de montrer que $B$ est positive avec $B = A^k$, il devient pertinent de passer par les valeurs propres : en effet, on sait que *via* diagonalisation, les valeurs propres de $B$ s'obtiennent à partir de celles de $A$ à la puissance $k$. Connaissant le signe des valeurs propres de $A$, on en déduit le signe des valeurs propres de $B$.

En revanche, le calcul de $X^\top BX = X^\top A^k X$ ne peut pas se ramener au calcul de $X^\top AX$ aisément : cette approche fait long feu.

**Exemple 24.** Soit $A$ une matrice symétrique positive. On veut justifier que $A^3$ est symétrique positive (il est facile de montrer qu'elle est symétrique : nous vous laissons le faire). Pour cela, nous allons exprimer les valeurs propres de $A^3$ en fonction de celles de $A$. Par le théorème spectral, il existe $P$ inversible et $D = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$ telles que : $A = PDP^{-1}$. Comme $A$ est positive, tous les $\lambda_i$ sont positifs. Alors :

$$A^3 = PD^3P^{-1} = P \begin{pmatrix} \lambda_1^3 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n^3 \end{pmatrix} P^{-1}.$$

On en déduit que les valeurs propres de $A^3$ sont les $\lambda_i^3$ avec $\lambda_i \in \mathrm{Sp}_{\mathbb{R}}(A)$. Or : $\forall i \in [\![1, n]\!]$, $\lambda_i \geqslant 0$, donc : $\forall i \in [\![1, n]\!]$, $\lambda_i^3 \geqslant 0$. On en déduit que toutes les valeurs propres de $A^3$ sont positives. D'où le résultat : $A^3$ est une matrice symétrique positive.

**Exercice 27.** Montrer que si $A$ est une matrice symétrique (non nécessairement positive), alors $A^k$ est symétrique positive pour tout entier *pair* $k$.

**Exercice 28.** Montrer que si $A$ est une matrice symétrique définie positive, alors $A^{-1}$ l'est aussi.

**Cas de matrices non explicites** Je pense par exemple à des exercices qui demandent si $AB$, ou $A+B$ (ou plus généralement : toute matrice $M$ définie à l'aide de $A$ et $B$), etc., est positive, sachant que $A$ et $B$ le sont. Dans ces cas-là il n'y a pas de réponse définitive concernant la méthode à adopter. Cela dépend. Néanmoins on retiendra que :

— l'application $M \mapsto X^\top MX$ est linéaire (c'est-à-dire notamment : $X^\top(M + \lambda N)X = X^\top MX + \lambda X^\top NX$), donc si $M$ est définie à l'aide d'une somme de matrices symétriques (disons $A$ et $B$ pour reprendre les notations ci-dessus), on peut aisément ramener l'étude de $X^\top MX$ à celle de $X^\top AX$ et $X^\top BX$ ;

— en revanche, pour le produit, on voit difficilement comment on exprimerait $X^\top ABX$ en fonction de $X^\top AX$ et $X^\top BX$ (ne parlez *surtout pas* d'écrire quelque chose comme : $X^\top ABX = X^\top AXX^{-1}BX = \cdots$ : puisque $X$ n'est pas une matrice carrée, elle n'est pas inversible).

**En résumé :** pour une SOMME, ou plus généralement une COMBINAISON LINÉAIRE de matrices carrées, il est pertinent d'étudier le caractère défini ou positif *via* le signe de $X^\top MX$. Mais pour le PRODUIT c'est à première vue impossible.

Pour le cas des valeurs propres, il faut être plus nuancé : c'est impossible *en général* (même si on connaît les valeurs propres de $A$ et $B$, on ne peut en déduire les valeurs propres de $A + B$ ou $AB$, par exemple parce qu'il n'y a pas de formule exprimant $\chi_{A+B}$ ou $\chi_{AB}$ en fonction de $\chi_A$ et $\chi_B$). Mais *parfois*, le miracle opère : **si $A$ et $B$ sont diagonalisables avec une même matrice de passage** $P$ (c'est le cas si $A$ et $B$ commutent, ce que nous ne démontrerons pas ici : c'est un exercice en soi). Dans ce cas, en effet, les opérations normalement incompatibles avec les éléments propres deviennent agréables. Si $A = PDP^{-1}$ et $B = PD'P^{-1}$ alors $D$, $D'$ diagonales, et $P$ inversible, alors :

$$A + B = P(D + D')P^{-1}, \quad \text{et :} \quad AB = PDP^{-1}PD'P^{-1} = PDD'P^{-1},$$

où $D + D'$ et $DD'$ sont diagonales : ainsi leurs coefficients diagonaux (qui dépendent de ceux de $D$ et $D'$) donnent les valeurs propres de $A + B$ et $AB$, et on est donc en mesure de trancher sur leur signe.

Comme on l'a dit plus haut, le cas de la somme s'étudie assez bien si l'on passe par le signe de $X^\top MX$. C'est donc plutôt en cas de produit qu'on songera à ce paragraphe.

## 4.4 Utiliser la réduction des endomorphismes autoadjoints pour *d'autres* endomorphismes

La situation très avantageuse des endomorphismes ou matrices symétriques (diagonalisation dans une base orthonormée) permet aussi d'en déduire des résultats très avantageux sur d'autres endomorphismes ou matrices, à condition de pouvoir s'y ramener. Il existe plusieurs façons de se ramener d'une matrice $M$ quelconque à une matrice symétrique, puisqu'en effet :

$$\forall M \in \mathrm{M}_n(\mathbb{R}), \quad M^\top M \in \mathrm{S}_n(\mathbb{R}), \quad M^\top + M \in \mathrm{S}_n(\mathbb{R}) \quad \text{(c'est quasiment sa projection orthogonale sur } \mathrm{S}_n(\mathbb{R})\text{)}$$

Avant de montrer comment tirer pleinement profit de cette méthode, il faut tout de même écarter les situations où elle N'est PAS intéressante du tout :

— si $M$ est une matrice orthogonale, alors $M^\top M = \mathrm{I}_n$ : matrice triviale qui ne nous permettra pas de déduire quoi que ce soit de $M$ ; donc si $M$ est une matrice orthogonale, on passera plutôt par $M + M^\top = M + M^{-1}$ ;

— si $M$ est une matrice antisymétrique, alors $M^\top + M = 0_{\mathrm{M}_n(\mathbb{R})}$ : matrice triviale qui ne nous permettra pas de déduire quoi que ce soit de $M$ ; donc si $M$ est une matrice antisymétrique, on passera plutôt par $M^\top M = -M^2$ (ou $M^2$, cela revient au même) ;

— dans les autres cas, ce n'est d'aucune aide si $M^\top$ ne vérifie rien de spécial (cela dépend donc des hypothèses).

Enfin, ce que nous utiliserons comme propriété dans ces cas-là n'est pas, *en général*, le fait qu'une matrice symétrique réelle soit diagonalisable, mais le résultat moins fort que **toutes ses valeurs propres sont réelles** (le polynôme caractéristique étant scindé sur $\mathbb{R}$). En effet, il n'y a pas (en général) de rapport entre la réduction de $M$ et celle de $M^\top M$ ou $M + M^\top$. Toutefois la relation $M^\top M X = \lambda X$ ou $(M + M^\top)X = \lambda X$ implique bien des relations entre $M$ et $\lambda$ ou $X$. On n'en déduit pas forcément des éléments propres de $M$, mais c'est mieux que rien.

**Exercice 29. (application de la méthode aux isométries)** Soit $f$ une isométrie d'un espace euclidien $E$. On suppose que $f$ n'admet pas de vecteur propre.

1. Montrer que pour tout $\vec{x} \in E$ non nul, la famille $(\vec{x}, f(\vec{x}))$ est libre.
2. Rappeler pourquoi une isométrie est inversible, et montrer que $g = f + f^{-1}$ est un endomorphisme autoadjoint.
3. En déduire l'existence de $\vec{x} \in E$ non nul, et de $\lambda \in \mathbb{R}$, tels que : $f^2(\vec{x}) = \lambda f(\vec{x}) - \vec{x}$.
4. Déduire de tout ce qui précède qu'il existe un plan (c'est-à-dire : un sous-espace vectoriel de dimension 2) de $E$ stable par $f$.

Ce résultat est EXTRÊMEMENT UTILE dans la réduction des isométries, dans le cas où elles n'ont pas de vecteur propre (et donc pas de droite stable). En effet, il nous permet d'étudier l'endomorphisme induit par cette isométrie sur un plan (et on sait que c'est une rotation planaire, grâce à la classification des isométries du plan), et en raisonnant par récurrence sur le supplémentaire orthogonal on en déduit une base de $E$ où la matrice de $f$ est diagonale par blocs, et les blocs diagonaux sont des matrices de rotation.

5. Le démontrer.

En vérité l'existence de plans stables reste vraie pour tout endomorphisme de $E$, mais la démonstration est plus délicate si l'on reste dans le cadre du programme.

**Exercice 30. (application de la méthode aux matrices antisymétriques réelles)** Soit $M$ une matrice antisymétrique *réelle*. Soit $\lambda \in \mathbb{C}$ une valeur propre de $M$.

1. Montrer que $M^2$ est une matrice symétrique réelle.
2. Montrer que $\lambda^2$ est une valeur propre de $M^2$ ; en déduire que les valeurs propres de $M$ sont soit des réels, soit des imaginaires purs.
3. On suppose dans cette question que $\lambda$ est une valeur propre **réelle** de $M$. Soit $X \in \mathrm{M}_{n,1}(\mathbb{R})$ un vecteur propre associé. Montrer : $\lambda^2 \cdot X^\top X = -(MX)^\top MX$, et en déduire : $\lambda = 0$.

Ainsi, toutes les valeurs propres d'une matrice antisymétrique réelle sont des imaginaires purs ! (notons que 0 en est un aussi) Ce résultat sur le spectre d'une matrice antisymétrique réelle est très pratique. Il rend trivial ce grand classique de l'oral :

**Exercice 31.** Montrer que si $M \in \mathrm{M}_n(\mathbb{R})$ est une matrice antisymétrique réelle, alors $\mathrm{I}_n + M$ est inversible (*on peut aussi faire cet exercice sans* AUCUN *résultat de réduction, mais en montrant que* $(\mathrm{I}_n + M)X = 0$ *implique* $X = 0$ *par des manipulations classiques*).

**Mise en garde 5.** Attention à ne pas avoir des raisonnements abusifs, et faire dire à cette méthode plus de choses qu'elle n'en promet. Par exemple, **ce serait un raisonnement faux** de dire que si $M$ est antisymétrique (par exemple), alors $M^2$ est symétrique d'après ce qui précède, donc diagonalisable sur $\mathbb{R}$ d'après le théorème spectral (pour l'instant tout est vrai), et que par conséquent $M$ l'est aussi. En effet, la matrice $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ n'est pas diagonalisable sur $\mathbb{R}$ (son polynôme caractéristique n'est même pas scindé), bien que son carré le soit (c'est la matrice $-\mathrm{I}_2$).

Ce n'est donc pas une méthode miracle qui permettrait de réduire n'importe quelle matrice à l'aide des matrices symétriques.

Enfin notons l'exercice intéressant suivant, apparu implicitement en exercice de travaux dirigés :

**Exercice 32.** Si $M \in \mathrm{M}_n(\mathbb{R})$, alors $M^\top M$ est symétrique et à valeurs propres positives ou nulles.

C'est donc un peu mieux que le simple fait d'être symétrique, à la lumière de tout ce qui fut prodigué sur les matrices positives dans la section 4.3. ← page 28

## 4.5 ☢ Utilisation de la racine carrée matricielle

Le résultat suivant est classique : pour toute matrice $M \in \mathrm{M}_n(\mathbb{R})$ symétrique **positive**, il existe une matrice $R \in \mathrm{M}_n(\mathbb{R})$ symétrique positive telle que : $M = R^2$. **Refaites cet exercice avant de lire cette section**.

L'énoncé est formulé dans le cas plus contraignant de valeurs propres strictement positives, mais en vérité elles ne servent que pour la démonstration de l'unicité.

L'intérêt des racines carrées matricielles apparaît lorsque nous étudions une relation entre deux matrices symétriques positives, **et que nous voulons remplacer l'une d'elles par la matrice identité sans perdre la symétrie de l'autre**. Donnons d'abord un exemple pour que le propos soit plus concret.

**Exemple 25.** Nous allons démontrer que si $A$ et $B$ sont deux matrices symétriques réelles **définies positives** :

$$\det(A+B) \geqslant \det(A) + \det(B).$$

Pour cela, je vais d'abord démontrer le résultat dans le cas plus simple où $A$ est la matrice identité, c'est-à-dire :

$$\det(\mathrm{I}_n + B) \geqslant \det(\mathrm{I}_n) + \det(B) = 1 + \det(B). \tag{7}$$

Pour la démontrer, on va se ramener au cas favorable d'une matrice diagonale, puisque dans ce cas le déterminant est simplement le produit des termes diagonaux. Or $B$ est symétrique réelle, donc diagonalisable : il existe $P$ inversible et $D$ diagonale telles que $B = PDP^{-1}$, et on en déduit (on écrit $\mathrm{I}_n = P\mathrm{I}_n P^{-1}$) :

$$\det(\mathrm{I}_n + B) = \det\left(P\left(\mathrm{I}_n + D\right)P^{-1}\right) = \det(P)\det\left(\mathrm{I}_n + D\right)\det(P)^{-1} = \det\left(\mathrm{I}_n + D\right).$$

Les coefficients diagonaux de $D$ sont les valeurs propres $\lambda_1, \ldots, \lambda_n$ de $B$, et elles sont strictement positives par hypothèse sur $B$. Donc :

$$\det\left(\mathrm{I}_n + D\right) = \det\left(\begin{pmatrix} 1+\lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1+\lambda_n \end{pmatrix}\right) = \prod_{i=1}^{n}\left(1+\lambda_i\right) = 1 + \prod_{i=1}^{n}\lambda_i + etc.,$$

où cette dernière égalité est obtenue en développant le produit : les deux termes isolés sont obtenus par multiplication des 1 de chaque facteur, ou des $\lambda_i$ de chaque facteur. Tous les autres termes du développement, cachés dans le « etc. », sont des produits des $\lambda_i$, donc sont positifs strictement. On en déduit : $\det(\mathrm{I}_n + D) > 1 + \prod_{i=1}^{n} \lambda_i = 1 + \det(D)$, ce qu'on voulait démontrer. Le raisonnement vaut pour toute matrice $B$ à valeurs propres strictement positives (précision d'importance pour ce qui suit).

Passons maintenant au cas général $\det(A+B) \geqslant \det(A) + \det(B)$. On l'a vu, il existe une matrice $R$ symétrique réelle définie positive telle que $R^2 = A$. Comme 0 n'en est pas valeur propre, on en déduit que $R$ est inversible, et on peut donc écrire :

$$\begin{aligned}\det(A+B) = \det(R^2+B) = \det\left(R\left(R+R^{-1}B\right)\right) &= \det(R)\det(R+R^{-1}B)\\ &= \det(R)\det\left(\left(\mathrm{I}_n + R^{-1}BR^{-1}\right)R\right)\\ &= \det(R)\det\left(\mathrm{I}_n + R^{-1}BR^{-1}\right)\det(R)\\ &= \det(R)^2\det\left(\mathrm{I}_n + R^{-1}BR^{-1}\right).\end{aligned}$$

Par des manipulations analogues, on a : $\det(A) + \det(B) = \det(R)^2\left(1 + \det\left(R^{-1}BR^{-1}\right)\right)$. Ainsi, si on reprend l'inégalité qu'on veut démontrer :

$$\begin{aligned}&\det(A+B) \geqslant \det(A) + \det(B)\\ \iff &\det(R)^2\det\left(\mathrm{I}_n + R^{-1}BR^{-1}\right) \geqslant \det(R)^2\left(1+\det\left(R^{-1}BR^{-1}\right)\right)\\ \iff &\det\left(\mathrm{I}_n + R^{-1}BR^{-1}\right) \geqslant 1 + \det\left(R^{-1}BR^{-1}\right), \qquad \text{car } \det(R)^2 > 0\end{aligned}$$

et on peut démontrer que $R^{-1}BR^{-1}$ est une matrice symétrique réelle définie positive (voir exercice ci-bas) ; elle vérifie donc l'inégalité (7) (en remplaçant $B$ par cette matrice). Ainsi cette équivalence est vraie, donc a bien $\det(A+B) \geqslant \det(A) + \det(B)$ : d'où le résultat.

**Exercice 33.** Vérifier le résultat admis : $R^{-1}BR^{-1}$ est symétrique réelle et définie positive ; pour cela, on vérifie que pour tout $X \in \mathrm{M}_{n,1}(\mathbb{R})$ non nul on a : $X^\top\left(R^{-1}BR^{-1}\right)X > 0$, en utilisant le fait que $B$ soit définie positive.

Si nous voulions nous ramener à la matrice identité sans utiliser la racine carrée, nous écririons : $\det(A+B) = \det\left(\left(A\left(\mathrm{I}_n + A^{-1}B\right)\right)\right) = \det(A)\det\left(\mathrm{I}_n + A^{-1}B\right)$. Effectivement, la matrice identité apparaît, mais on ne peut rien faire de $A^{-1}B$ : ce n'est *a priori* pas une matrice symétrique, et en particulier on ne sait pas si elle est diagonalisable ; même si elle était, rien n'assurerait que ses valeurs propres sont positives : elles n'ont pas de rapport *a priori* avec celles de $A$ et de $B$.

Autre problème : on pourrait se demander pourquoi il ne suffisait pas tout simplement de diagonaliser $A$ et $B$, de sorte à pouvoir écrire $\det(A+B)$ comme le produit des $\mu_i + \lambda_i$, où les $\mu_i$ et $\lambda_i$ sont les valeurs propres de $A$ et $B$ respectivement. Le problème est qu'*a priori* $A$ et $B$ ne sont pas diagonalisables dans les mêmes bases, et donc pas avec les mêmes matrices de passage : on ne peut donc pas *a priori* écrire $A+B = P(D'+D)P^{-1}$ pour en déduire $\det(A+B) = \det(D'+D) = \prod_{i=1}^{n}(\mu_i + \lambda_i)$. C'est faux en général !

On voit donc qu'en factorisant $\det(A+B)$ à l'aide de $A = R^2$, nous avons non seulement simplifié le problème en nous ramenant à la matrice identité, mais nous n'avons pas « perdu » la symétrie de la seconde matrice, puisque $R^{-1}BR^{-1}$ reste symétrique, réelle, et à valeurs propres positives. C'est pour conserver la symétrie que la racine carrée $R$ a grand intérêt, **en factorisant à gauche et à droite par** $R$ (là est l'astuce).

**Exercice 34.**

1. Montrer que le résultat reste vrai si $B$ est **positive**, et que $\det(A+B) = \det(A) + \det(B)$ si et seulement si $B$ est la matrice nulle.
2. Montrer que le résultat reste vrai si $A$ et $B$ sont **positives**. Si vous ne voulez pas passer par l'algèbre linéaire, un argument de continuité est possible : voir *Méthodes*, chapitre de topologie, section 7.3.2.

**Exercice 35.**

1. Soit $A$ une matrice symétrique réelle d'ordre $n$, définie positive (donc inversible en particulier). Montrer que si on a :
$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \{0\}, \quad X^\top AX > X^\top X,$$
alors :
$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \{0\}, \quad X^\top A^{-1}X < X^\top X$$
*(ramener ces inégalités à des inégalités sur les valeurs propres de $A$).*

2. En déduire, en utilisant la question précédente et la racine carrée de $A$, que si $A$ et $B$ sont deux matrices symétriques réelles d'ordre $n$, définies positives, telles que :
$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \{0\}, \quad X^\top AX > X^\top BX,$$
alors :
$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}) \setminus \{0\}, \quad X^\top A^{-1}X < X^\top B^{-1}X.$$

Cet exercice généralise l'implication $0 < b < a \Rightarrow 0 < \frac{1}{a} < \frac{1}{b}$ bien connue pour les nombres réels.

En principe, l'existence de racines carrées n'est facile à démontrer **que pour les matrices symétriques positives**. Mais pour toute matrice $M$ (même non symétrique), l'exercice 32 permet d'établir l'existence d'une matrice $R$ symétrique positive telle que $R^2 = M^\top M$. C'est parfois utile. ← page 34

**Exercice 36.** À l'aide de cette idée, montrer que toute matrice inversible $M$ peut s'écrire de manière unique sous la forme $M = OS$ avec $O$ orthogonale et $S$ symétrique définie positive (*déterminer $O$ et $S$ par analyse et synthèse*).

### 4.6 ☢ Étudier les matrices (anti)symétriques à l'aide des endomorphismes (anti)symétriques

Il est conseillé de lire cette section conjointement à la section 5.1 : il y a plusieurs techniques qui se font écho, et vous pourriez gagner en compréhension en les mettant en parallèle. → page 39

La stratégie n'est pas nouvelle : depuis le chapitre de réduction des endomorphismes, nous avons réduit des matrices en passant par leurs endomorphismes associés, et en écrivant leurs matrices dans des bases adaptées à une décomposition en sous-espaces stables (le meilleur cas de figure étant le cas où ces sous-espaces stables sont des sous-espaces propres : la restriction à ces sous-espaces est une homothétie, dont la matrice est diagonale).

La question se pose aussi dans le cas particulier des matrices symétriques et antisymétriques (on pourrait en discuter avec bien d'autres ; notamment les matrices orthogonales, voir section 5.1). Mais réduire n'a jamais été aussi souple qu'avec ces endomorphismes : ils vérifient en effet que si un sous-espace est stable, alors le supplémentaire orthogonal est stable aussi (c'est le cours pour les endomorphismes autoadjoints, et l'exercice 37 ci-dessous pour les antisymétriques). Voici alors une stratégie très efficace pour réduire une matrice $A$ de cette nature (je reste dans le vague sur le type de réduction, cela dépend de ce qu'on veut démontrer) : → page 39

1. On introduit l'endomorphisme $f$ de $\mathrm{M}_{n,1}(\mathbb{R})$ canoniquement associé, et on en détermine un « bon » sous-espace stable $F$ : un sous-espace propre si on en connaît un, ou seulement une droite stable si on ne connaît qu'un vecteur propre ; un plan stable faute de mieux (dans ce cas on doit vous donner des indications). L'endomorphisme induit par $f$ sur $F$ a alors une matrice « simple » que je note $M$ dans une base $\mathcal{B}_0$ de $F$.
2. L'endomorphisme $f$ laisse aussi stable $F^\perp$, et l'endomorphisme induit par $f$ sur $F^\perp$ vérifie les mêmes propriétés que $f$ ; comme $\dim(F^\perp) = \dim(\mathrm{M}_{n,1}(\mathbb{R})) - \dim(F) < n$, la dimension de $F^\perp$ est strictement inférieure à celle de l'espace entier, ce qui permet d'en déduire par récurrence forte une base $\mathcal{B}_1$ de $F^\perp$ où la matrice de l'endomorphisme induit, notée $N$, est « simple ».

3. Dans la base $\mathcal{B}_0 \cup \mathcal{B}_1$ adaptée à la décomposition $E = F \oplus F^\perp$, la matrice de $f$ est $\begin{pmatrix} M & \mathbf{0} \\ \mathbf{0} & N \end{pmatrix}$. En utilisant la formule du changement de base, on en déduit que $A$ est semblable à $\begin{pmatrix} M & \mathbf{0} \\ \mathbf{0} & N \end{pmatrix}$.

Voir l'exemple 26 pour une application très concrète de cette méthode. → page 37

Dans ce genre d'exercice, on préfèrera vous embêter avec des matrices autres que symétriques (je parlerai essentiellement des matrices antisymétriques dans cette section). En effet, une matrice symétrique est rendue trop facile à réduire avec le théorème spectral, alors que pour les autres types de matrices vous ne savez rien et tout est à démontrer. C'est d'ailleurs l'occasion de poser des questions de cours déguisées, vu que beaucoup de résultats pour les endomorphismes antisymétriques se démontrent de la même manière que pour les endomorphismes autoadjoints. Nous récapitulons ces démonstrations « communes » dans l'exercice suivant : si vous peinez, reprenez le cours sur les endomorphismes autoadjoints et adaptez-le *mutatis mutandis*.

**Exercice 37.** Soit $f$ un endomorphisme antisymétrique de $E$.

1. Montrer que sa matrice relativement à une base orthonormée est antisymétrique.
2. Réciproquement, si $M \in \mathrm{M}_n(\mathbb{R})$ est une matrice antisymétrique, montrer que son endomorphisme canoniquement associé est un endomorphisme antisymétrique de $\mathrm{M}_{n,1}(\mathbb{R})$.
3. Montrer que si $F$ est un sous-espace vectoriel de $E$ stable par $f$, alors $F^\perp$ est aussi stable par $f$.
4. Montrer que $\ker(f) = \mathrm{im}(f)^\perp$ et $\ker(f)^\perp = \mathrm{im}(f)$.

**Dans tout le reste de la section, on considère ces propriétés comme connues** (mais elles seraient bien sûr à redémontrer en exercice).

On en déduit une première piste pour réduire des endomorphismes antisymétriques : notons que $\ker(f)$ (et donc $\ker(f)^\perp = \mathrm{im}(f)$) est TOUJOURS un sous-espace stable. **Il ne coûte donc rien de toujours commencer à étudier l'endomorphisme induit par $f$ sur** $\ker(f)^\perp$, ce qui a deux avantages loin d'être négligeables :

— l'endomorphisme induit par $f$ sur $\ker(f)^\perp$ EST TOUJOURS BIJECTIF (en effet, un vecteur de son noyau serait à la fois dans $\ker(f)$ et $\ker(f)^\perp$, or $\ker(f) \cap \ker(f)^\perp = \{\vec{0}\}$), ce qui est intéressant ne serait-ce que pour la non nullité de $f(\vec{x})$ pour tout $\vec{x} \in \ker(f)^\perp$ ;

— l'image par $f$ de tout vecteur de $\ker(f)$ est nulle (c'est la définition), donc la matrice de l'endomorphisme induit par $f$ sur $\ker(f)$ est la matrice nulle.

L'intérêt de ce dernier point : une fois qu'on a trouvé une « bonne » base de $\ker(f)^\perp$ pour notre réduction, on en déduit une « bonne » base de $E$ en la complétant avec une base de $\ker(f)$ (du fait que $E = \ker(f) \oplus \ker(f)^\perp$). Alors, lorsqu'on écrit la matrice de $f$ dans cette base adaptée, les colonnes correspondant aux images par $f$ des vecteurs de $\ker(f)$ sont toutes nulles. On peut difficilement faire plus simple.

**Exemple 26.** Posons $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}$. Nous allons (un peu) réduire $A$ : soit $f$ l'endomorphisme de $\mathrm{M}_{3,1}(\mathbb{R})$ canoniquement associé à $A$.

1. Je détermine d'abord $\ker(f)$ ; la résolution du système $AX = 0_{\mathrm{M}_{3,1}(\mathbb{R})}$ donne : $\ker(f) = \mathrm{Vect}_{\mathbb{R}}\left(\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\right)$.
2. Ensuite, j'étudie l'endomorphisme induit par $f$ sur $\ker(f)^\perp$, qui est le plan de vecteur normal $U_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, c'est-à-dire le plan d'équation $-x+y+z=0$. Une base de ce plan est $\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)$ ;

mais je vais privilégier une base orthonormée, de sorte à pouvoir utiliser le fait que $f$ soit de matrice antisymétrique dans une base orthonormée. L'algorithme de Gram-Schmidt permet d'obtenir cette base orthonormée de $\ker(f)^\perp$ :

$$(U_2, U_3) = \left( \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right).$$

Pour déterminer la matrice de l'endomorphisme induit par $f$ sur $\ker(f)^\perp$ relativement à cette base, exprimons $f(U_2)$ et $f(U_3)$ en fonction de $U_2$ et $U_3$ :

$$f(U_2) = AU_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \sqrt{3}U_3, \quad \text{et :} \quad f(U_3) = AU_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -3 \\ -3 \\ 0 \end{pmatrix} = -\sqrt{3}U_2,$$

donc dans la base $(U_2, U_3)$, la matrice de l'endomorphisme induit par $f$ sur $\ker(f)^\perp$ est : $M = \begin{pmatrix} 0 & -\sqrt{3} \\ \sqrt{3} & 0 \end{pmatrix}$ (en vérité, en utilisant le fait que $M$ soit antisymétrique d'ordre 2, on savait déjà qu'elle serait de la forme $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ avec $a \in \mathbb{R}$ : il n'y a donc aucune surprise à trouver $f(U_2)$ proportionnel à $U_3$, et le calcul de $f(U_3)$ est rendu inutile par l'antisymétrie).

3. Concluons. La concaténation des bases : $\mathcal{B} = (U_1, U_2, U_3)$, est une base de $\mathrm{M}_{3,1}(\mathbb{R})$ du fait que $\ker(f) \oplus \ker(f)^\perp = \mathrm{M}_{3,1}(\mathbb{R})$ (elle n'est pas orthonormée à cause du premier vecteur, mais ce n'est pas grave). On a $f(U_1) = 0_{\mathrm{M}_{3,1}(\mathbb{R})}$ car $U_1 \in \ker(f)$, et l'étude du point précédent démontre que dans la base $\mathcal{B}$, la matrice de $f$ est : $\begin{pmatrix} 0 & 0_{\mathrm{M}_{1,2}(\mathbb{R})} \\ 0_{\mathrm{M}_{2,1}(\mathbb{R})} & M \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sqrt{3} \\ 0 & \sqrt{3} & 0 \end{pmatrix}$. Or $A$ est la matrice de $f$ dans la base canonique $\mathcal{B}_c$. Par conséquent, si l'on pose $P = \mathrm{M}_{\mathcal{B}_c}(\mathcal{B})$, alors la formule du changement de base implique :

$$\mathrm{M}_{\mathcal{B}_c}(f) = \mathrm{M}_{\mathcal{B}_c}(\mathcal{B})\mathrm{M}_{\mathcal{B}}(f)\mathrm{M}_{\mathcal{B}}(\mathcal{B}_c) \iff A = P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sqrt{3} \\ 0 & \sqrt{3} & 0 \end{pmatrix} P^{-1}.$$

On a réduit $A$.

On peut montrer que $\chi_A = X(X^2 + 3)$ très facilement en passant par la matrice réduite ci-dessus ; il n'est pas scindé, donc $A$ n'est pas diagonalisable (ni trigonalisable) sur $\mathbb{R}$, et la réduction proposée est la meilleure possible (sauf si l'on va dans $\mathbb{C}$).

Cet exemple est un cas simplifié de l'exercice suivant, qui n'est finalement qu'une formalisation matricielle du conseil ci-dessus :

**Exercice 38.** Soit $f$ un endomorphisme antisymétrique de $E$. On admet que sa matrice dans toute base orthonormée est antisymétrique : voir l'exercice 37 ci-dessus. ← page 37

1. On suppose dans cette question que $f$ est *bijective*. Soit $M$ sa matrice dans une base orthonormée. En simplifiant $\det(M^\top)$ de deux façons différentes, montrer que la dimension de $E$ est un entier pair.
2. On ne suppose plus que $f$ est bijective. Avec la question précédente et la décomposition $E = \ker(f) \oplus \ker(f)^\perp$, montrer que la dimension de $\ker(f)^\perp$ est un entier pair.
3. En déduire que le rang d'un endomorphisme antisymétrique est toujours un entier pair.
4. Montrer que toute matrice antisymétrique $M$ est semblable à une matrice de la forme (par blocs) $\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & N \end{pmatrix}$, où $N$ est une matrice antisymétrique, *inversible*, et de taille paire.

Il reste alors à réduire $N$ pour réduire $M$.

**Culture scientifique.** Le contenu de cette section se généraliserait à tout endomorphisme dont la matrice, dans une base orthonormée quelconque, commute avec sa transposée. Un tel endomorphisme est dit *normal*. Les endomorphismes symétriques et antisymétriques, et les isométries, sont des cas particuliers d'endomorphismes normaux.

**Exercice 39.** Démontrer les affirmations de la dernière phrase.

En voici un cas particulier.

**Exercice 40.** Soit $M$ une matrice réelle d'ordre 3 telle que : $M^\top M = MM^\top$. On note $f$ l'endomorphisme de $\mathrm{M}_{3,1}(\mathbb{R})$ canoniquement associé à $M$, et $g$ l'endomorphisme de $\mathrm{M}_{3,1}(\mathbb{R})$ canoniquement associé à $M^\top$.

1. Montrer que $f$ admet au moins une valeur propre réelle $\lambda$ (*penser à ce qu'on a établi dans le cours pour la classification des isométries de l'espace*). On pose $E_\lambda = \ker(f - \lambda \mathrm{Id}_E)$.
2. On note $\langle \cdot, \cdot \rangle$ le produit scalaire usuel sur $\mathrm{M}_{3,1}(\mathbb{R})$. Montrer que pour tout $(X, Y) \in (\mathrm{M}_{3,1}(\mathbb{R}))^2$, on a : $\langle f(X), Y \rangle = \langle X, g(Y) \rangle$.
3. En déduire que si $F$ est stable par $f$, alors $F^\perp$ est stable par $g$ (et inversement).
4. Grâce au fait que $f$ et $g$ commutent, montrer que $E_\lambda$ est stable par $g$, puis que $E_\lambda^\perp$ est stable par $f$ et $g$.
5. On suppose que $\dim(E_\lambda) = 2$. Montrer que $M$ est diagonalisable sur $\mathbb{R}$ dans une base orthonormée.
6. On suppose à présent que $\dim(E_\lambda) = 1$. Si on note $N = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ la matrice de l'endomorphisme induit $f_{E_\lambda^\perp}$ dans une base **orthonormée** $\mathcal{B}$ de $E_\lambda^\perp$, montrer que : $\mathrm{M}_\mathcal{B}\left(g_{E_\lambda^\perp}\right) = N^\top$, où $g_{E_\lambda^\perp}$ est l'endomorphisme induit par $g$ sur $E_\lambda^\perp$ (*utiliser la question 2*).
7. Résoudre $N^\top N = NN^\top$, et en déduire que $M$ est diagonalisable sur $\mathbb{R}$ ou $\mathbb{C}$, et que dans le premier cas on peut trouver une base orthonormée de vecteurs propres.

# 5 ✓ Isométries

## 5.1 ☢ Réduction des isométries en dimension supérieure

La stratégie de réduction d'une isométrie est la même que pour n'importe quel endomorphisme : on cherche une base adaptée à une décomposition en « bons » sous-espaces stables. Avec, comme dans le cas des endomorphismes autoadjoints et antisymétriques (voir la section 4.6), quelques propriétés qui facilitent la réduction : ← page 36

— le supplémentaire orthogonal d'un sous-espace stable est stable ;

— on connaît (parfois) des bons sous-espaces stables : les sous-espaces propres associés à 1 ou $-1$ ;

— on connaît la forme des isométries (et de leurs matrices) en dimension 2 ou 3.

À noter qu'on sait qu'il n'existe pas de vecteur propre associé à 1 dans $\ker(f - \mathrm{Id}_E)^\perp$, étant donné que $\ker(f - \mathrm{Id}_E)^\perp \cap \ker(f - \mathrm{Id}_E) = \{\vec{0}\}$ et qu'un vecteur propre associé à 1 est un vecteur non nul de $\ker(f - \mathrm{Id}_E)$. Cela permet d'exclure des possibilités d'isométries, au moment de déterminer l'endomorphisme induit par $f$ sur $\ker(f - \mathrm{Id}_E)^\perp$. De même si l'on raisonne sur $\ker(f + \mathrm{Id}_E)^\perp$.

Il vaut le coup de bien assimiler le théorème de classification en dimension 3 : il contient presque toute la stratégie valable en dimension quelconque. La seule subtilité est le cas où nous n'avons pas de valeur propre (impossible en dimension impaire). Auquel cas, n'ayant pas de sous-espace propre stable, nous trouvons mieux que rien en dénichant un plan stable $P$ : voir l'exercice 29. Comme nous savons expliciter la matrice d'une isométrie sur un plan stable (et ce doit être une rotation s'il n'y a pas de valeur propre), la description de la restriction à $P$ est facile, puis on raisonne de même sur $P^\perp$ : une ← page 33

récurrence sur la dimension achève le travail.

**Exercice 41.**

1. Utiliser ces remarques pour en déduire que le supplémentaire orthogonal de :
$$F = \ker(f - \mathrm{Id}_E) \oplus \ker(f + \mathrm{Id}_E)$$
est toujours de dimension paire (*montrer que dans le cas contraire, l'endomorphisme induit par $f$ sur $F^\perp$ aurait nécessairement des valeurs propres, et que c'est impossible*).
2. En déduire que si $f$ est une isométrie, alors $\dim(\ker(f - \mathrm{Id}_E)) + \dim(\ker(f + \mathrm{Id}_E))$ a toujours la parité de $\dim(E)$.

## 5.2 ✓ Reconnaître une isométrie en dimension 2

Dans ce cas la situation est très simple, parce qu'on voit à l'œil nu si l'on a une matrice de la forme :

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ ou } \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

avec $a^2 + b^2 = 1$. Dans le premier cas, c'est une matrice de rotation ; dans le second cas, c'est une matrice de réflexion (on peut écrire $a$ et $b$ sous la forme $a = \cos(\theta)$ et $b = \sin(\theta)$). Au moment de déterminer plus précisément l'angle de la rotation (si c'est une rotation) ou l'axe de symétrie (si c'est une réflexion), il y a deux petites subtilités.

**Déterminer l'angle d'une rotation.** Si la matrice est de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ avec $a^2 + b^2 = 1$, alors on sait qu'il existe $\theta \in \mathbb{R}$ tel que $a = \cos(\theta)$ et $b = \sin(\theta)$. Pour en déduire $\theta$, **une erreur de raisonnement est de dire simplement que** $\theta = \arccos(a)$ convient. **C'est faux si** $\sin(\theta) < 0$ ! En effet, si $\sin(\theta) < 0$, alors il est impossible de prendre $\theta \in [0, \pi]$ (le sinus est positif sur cet intervalle). Or c'est seulement pour $\theta$ dans cet intervalle qu'on peut écrire l'équivalence : $\cos(\theta) = a \iff \theta = \arccos(a)$ (parce que l'arc cosinus est la bijection réciproque de la RESTRICTION du cosinus à $[0, \pi]$). Retenez cette alternative :

— si $b > 0$, alors $\theta = \arccos(a)$ CONVIENT ;

— si $b < 0$, alors $\theta = -\arccos(a)$ CONVIENT.

Je justifie le second cas : si l'on prend $\theta = -\arccos(a)$, vérifions qu'on a $a = \cos(\theta)$ et $b = \sin(\theta)$ ; d'une part : $a = \cos(\arccos(a)) = \cos(-\theta) = \cos(\theta)$, et d'autre part : $(\sin(\theta))^2 = 1 - \cos(\theta)^2 = 1 - a^2 = b^2$, et donc en prenant la racine carrée : $|\sin(\theta)| = \sqrt{b^2} = |b| = -b$ (on est dans le cas où $b < 0$) ; comme $\theta = -\arccos(a) \in [-\pi, 0]$, on a $\sin(\theta) < 0$, et donc : $|\sin(\theta)| = -\sin(\theta)$. On déduit de tout cela que $b = \sin(\theta)$, comme voulu.

**Remarque.** N'oubliez pas le mot « convient » : c'est FAUX que si $\cos(\theta) = a$ alors $\theta = \arccos(a)$ : ce n'est vrai que modulo $2\pi$... Et de toute façon, c'est l'implication réciproque que l'on veut : ce que vous dites, c'est que vous avez fait UN choix de $\theta$ qui marche (en ajoutant des multiples de $2\pi$ vous avez d'autres choix fonctionnels).

**Déterminer l'axe de symétrie d'une réflexion.** On pourrait s'inspirer de la méthode ci-dessus, en résolvant $a = \cos(\theta)$ et $b = \sin(\theta)$. Le problème est qu'en général, on n'obtient pas ainsi un vecteur directeur explicite de l'axe de symétrie, à cause d'arc cosinus non simplifiables. Pour déterminer un vecteur directeur de l'axe de symétrie de $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ avec $a^2 + b^2 = 1$, **on cherche un vecteur propre pour la valeur propre** 1 **en résolvant** $AX = X$. Le vecteur propre retenu engendre l'axe de symétrie : il aura toujours une expression simple.

Dans le cas d'un endomorphisme, ce n'est pas très éloigné du cas matriciel : on cherche une base orthonormée de l'espace sur lequel est défini l'endomorphisme, et on écrit sa matrice relativement à cette base.

## 5.3 ✓ Reconnaître une isométrie en dimension 3

La stratégie pour déterminer une isométrie de $\mathbb{R}^3$ est la suivante. Soit $f$ un endomorphisme dont on connaît la matrice $M$ dans la base canonique. On veut identifier sa nature ; on suppose que $f \neq \pm \mathrm{Id}_E$ (sinon il n'y a rien à raconter).

1. On vérifie d'abord que la matrice $M$ est orthogonale, ce qui montre que $f$ est une isométrie. Il est plus efficace de le faire en vérifiant que les colonnes forment une famille orthonormée, plutôt qu'en regardant si $M^\top M = \mathrm{I}_3$ (trop coûteux en calculs).
2. On détermine le sous-espace $\ker(M - \mathrm{I}_3)$ des vecteurs invariants avec la méthode du pivot de Gauß.
3. (a) Si $\dim(E_1) = 2$, alors $f$ est la réflexion par rapport à $E_1$.
   (b) Si $\dim(E_1) = 1$, alors $f$ est une rotation d'axe $E_1$, dirigé et orienté par un vecteur unitaire $\vec{a}$. On détermine une mesure d'angle $\theta$ de $f$ avec la relation $\mathrm{tr}(f) = 1 + 2\cos(\theta)$, sachant que le signe du produit mixte $[\vec{a}, \vec{u}, r(\vec{u})]$, où $\vec{u}$ est un vecteur non nul orthogonal à $\vec{a}$, permet de déterminer le signe de $\sin(\theta)$, et donc de déterminer $\theta \bmod 2\pi$.
   (c) Si $\dim(E_1) = 0$, alors $-f$ est une rotation dont on détermine l'axe de rotation $D$ et une mesure d'angle $\vartheta$. Alors $f$ est la composée commutative de la réflexion par rapport à $P = D^\perp$ et de la rotation d'axe $D$ et de mesure d'angle $\vartheta + \pi$.

**Remarquons qu'il est inutile de calculer** $\det(M)$.

**Sur la vérification que les colonnes forment une base orthonormée de** $\mathrm{M}_{3,1}(\mathbb{R})$**.** Vous oubliez souvent une des deux conditions pour être une base orthonormée. Il faut des colonnes ORTHOGONALES ($\langle C_1, C_2\rangle = \langle C_1, C_3\rangle = \langle C_2, C_3\rangle = 0$) et UNITAIRES ($\langle C_1, C_1\rangle = \langle C_2, C_2\rangle = \langle C_3, C_3\rangle = 1$). Cela fait donc six produits scalaires à calculer. CETTE ÉTAPE N'EST PAS FACULTATIVE ! (même si vous pouvez l'abréger)

**Sur l'angle de la rotation.** La détermination de $\cos(\theta)$ se passe de commentaire. Mais elle est en général insuffisante pour obtenir $\theta$ (nous avons déjà fait ce constat pour les isométries du plan, ci-dessus). Elle n'est suffisante que si $\cos(\theta) = -1$ : dans ce cas, on sait que $\theta \equiv \pi \bmod 2\pi$ et il est inutile d'aller plus loin. Sinon, au moment de calculer le produit mixte $[\vec{a}, \vec{u}, f(\vec{u})]$ donnant le signe de $\sin(\theta)$ (et non $\sin(\theta)$ si $\vec{a}$ et $\vec{u}$ ne sont pas unitaires, mais de toute façon on n'en demande pas tant) : **prenez $\vec{u}$ aussi simple que possible parmi les vecteurs non nuls orthogonaux à $\vec{a}$. Privilégiez un vecteur avec plusieurs coordonnées nulles, par exemple les vecteurs de la base canonique, si c'est possible**. Non seulement cela vous simplifiera le calcul de $f(\vec{u})$, mais aussi celui du produit mixte (en développant par rapport à la deuxième colonne). Ensuite, si $\vec{u} = (a, b, c)$, alors pour obtenir $f(\vec{u})$, vous calculez le produit $M \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

Enfin, si vous obtenez une mesure d'angle opposée à celle proposée dans une correction d'exercice, ce n'est pas grave : vous avez sans doute orienté l'axe de la rotation par un vecteur de sens opposé à celui choisi dans l'exercice.

**Remarque.** Si $M$ est symétrique en plus d'être orthogonale, alors les égalités $M^\top = M$ et $MM^\top = \mathrm{I}_3$ impliquent : $M^2 = \mathrm{I}_3$. Vous savez donc déjà *a priori* que $f$ est une symétrie orthogonale : soit une réflexion (cas où $\dim(E_1) = 2$), soit une rotation d'angle $\pi$ modulo $2\pi$ (cas où $\dim(E_1) = 1$ : on parle dans ce cas de *retournement*). Vous en rendre compte vous permet de détecter une erreur dans vos calculs ultérieurs.

**Cas d'une composition commutative.** Nous traitons ci-dessous le cas d'une composition commutative, pour vous montrer qu'il n'y a finalement pas grand'chose de neuf. Pour comprendre « pourquoi ça marche de faire ainsi », voir le cours.

**Exemple 27.** Montrons que la matrice $A = \dfrac{1}{50}\begin{pmatrix} -25 & 20\sqrt{3} & -15\sqrt{3} \\ -20\sqrt{3} & -34 & -12 \\ 15\sqrt{3} & -12 & -41 \end{pmatrix}$ est la matrice dans la base canonique d'une isométrie $f$ de $\mathbb{R}^3$ que nous déterminerons. Nous vous laissons vérifier que les colonnes forment une base orthonormée de $\mathrm{M}_{3,1}(\mathbb{R})$ (ne pas oublier le facteur $\frac{1}{50}$). Soit $\vec{x} = (a, b, c) \in \mathbb{R}^3$. Alors :

$$\vec{x} \in \ker(f - \mathrm{Id}_{\mathbb{R}^3}) \iff A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \iff \left\{\begin{array}{rrrcl} -75a & +20\sqrt{3}b & -15\sqrt{3}c & = & 0 \\ -20\sqrt{3}a & -84b & -12c & = & 0 \\ 15\sqrt{3}a & -12b & -91c & = & 0 \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} -15a & +4\sqrt{3}b & -3\sqrt{3}c & = & 0 & \left(L_1 \leftarrow \frac{1}{5}L_1\right) \\ -5\sqrt{3}a & -21b & -3c & = & 0 & \left(L_2 \leftarrow \frac{1}{4}L_2\right) \\ 15\sqrt{3}a & -12b & -91c & = & 0 & \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} -15a & +4\sqrt{3}b & -3\sqrt{3}c & = & 0 & \\ & -25\sqrt{3}b & & = & 0 & \left(L_2 \leftarrow \sqrt{3}L_2 - L_1\right) \\ & & -100c & = & 0 & \left(L_3 \leftarrow L_3 + \sqrt{3}L_1\right) \end{array}\right.$$

On en déduit aisément $a = b = c = 0$, donc $\vec{x} = \vec{0}$. Autrement dit : $\dim(\ker(f - \mathrm{Id}_{\mathbb{R}^3})) = 0$, donc $f$ est la composition commutative d'une rotation et d'une réflexion, dont nous allons à présent déterminer les caractéristiques géométriques.

Pour cela, on se souvient que $-f$ est une rotation, dont on détermine l'axe en déterminant les vecteurs invariants, puis une mesure d'angle avec la trace et un produit mixte bien choisi. Déterminons l'axe :

$$-f(\vec{x}) = \vec{x} \iff A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = -\begin{pmatrix} a \\ b \\ c \end{pmatrix} \iff \left\{\begin{array}{rrrcl} 25a & +20\sqrt{3}b & -15\sqrt{3}c & = & 0 \\ -20\sqrt{3}a & +16b & -12c & = & 0 \\ 15\sqrt{3}a & -12b & +9c & = & 0 \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} 5a & +4\sqrt{3}b & -3\sqrt{3}c & = & 0 & (L_1 \leftarrow \frac{1}{5}L_1) \\ -5\sqrt{3}a & +4b & -3c & = & 0 & (L_2 \leftarrow \frac{1}{4}L_2) \\ 5\sqrt{3}a & -4b & +3c & = & 0 & (L_3 \leftarrow \frac{1}{3}L_3) \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} 5a & +4\sqrt{3}b & -3\sqrt{3}c & = & 0 & \\ -5\sqrt{3}a & +4b & -3c & = & 0 & \\ & & 0 & = & 0 & (L_3 \leftarrow L_3 + L_2) \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} 5a & +4\sqrt{3}b & -3\sqrt{3}c & = & 0 & \\ & 16b & -12c & = & 0 & (L_2 \leftarrow L_2 + \sqrt{3}L_1) \\ & & 0 & = & 0 & \end{array}\right.$$

$$\iff \exists \lambda \in \mathbb{R}, \quad \left\{\begin{array}{rcl} a & = & \sqrt{3}(3c - 4b) = 0 \\ b & = & \frac{3}{4}\lambda \\ c & = & \lambda \end{array}\right.$$

donc : $D = \ker(-f - \mathrm{Id}_{\mathbb{R}^3}) = \mathrm{Vect}_{\mathbb{R}}\left(\left(0, \frac{3}{4}, 1\right)\right) = \mathrm{Vect}_{\mathbb{R}}((0,3,4))$. On en déduit que $-f$ est une rotation d'axe $D$ qu'on oriente par $\vec{a} = (0,3,4)$. Pour déterminer une mesure d'angle $\theta \in \mathbb{R}$, on se souvient qu'elle vérifie $1 + 2\cos(\theta) = \mathrm{tr}(-f) = -\mathrm{tr}(A) = 2$, donc $\cos(\theta) = \frac{1}{2}$. Nous avons aussi besoin du signe de $\sin(\theta)$ ; il est de même signe que :

$$\left[\vec{a}, \vec{i}, f(\vec{i})\right] = \begin{vmatrix} 0 & 1 & \frac{1}{2} \\ 3 & 0 & \frac{2\sqrt{3}}{5} \\ 4 & 0 & -\frac{3\sqrt{3}}{10} \end{vmatrix} = -\begin{vmatrix} 3 & \frac{2\sqrt{3}}{5} \\ 4 & -\frac{3\sqrt{3}}{10} \end{vmatrix} = \frac{5\sqrt{3}}{2} > 0. \quad (\text{on a bien : } \vec{a} \perp \vec{i})$$

On en déduit $\sin(\theta) > 0$, et donc : $\theta \equiv \frac{\pi}{3} \bmod 2\pi$.

En conclusion, $-f$ est la rotation d'axe orienté par $\vec{a} = (0,3,4)$ et d'angle $\frac{\pi}{3}$ modulo $2\pi$. On en déduit que $f$ est la composition de la rotation d'axe orienté par $\vec{a} = (0,3,4)$ et d'angle $\frac{\pi}{3} + \pi = \frac{4\pi}{3} \bmod 2\pi$, et de la réflexion par rapport au plan $P = D^{\perp}$, qui est d'équation $3y + 4z = 0$ (parce que $\vec{a}$ en est un vecteur normal).

**Commentaire sur le plan de la réflexion, dans le cas de la composition commutative.** Une fois que vous avez un vecteur qui dirige et oriente l'axe de la rotation en présence, inutile de vous fatiguer pour trouver le plan de la réflexion : le vecteur qui oriente l'axe de la rotation est un vecteur normal. Ainsi, si $(a, b, c)$ dirige l'axe de la rotation, alors la réflexion est par rapport au plan d'équation $ax + by + cz = 0$.

**Commentaire sur les racines carrées dans les systèmes linéaires.** Dans l'exemple ci-dessus, les racines carrées de 3 se simplifient très heureusement. Vous n'aurez pas toujours de telles simplifications (s'il apparaît en coefficients des nombres de la forme $a + b\sqrt{d}$, et parfois pire), ce qui complique la résolution de $AX = X$. Je vous conseille alors de simplifier la quantité $a + b\sqrt{d}$ en facteur du pivot que vous utilisez à chaque étape, *en multipliant la ligne par son conjugué* $a - b\sqrt{d}$. Grâce à l'identité remarquable $(a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - b^2 d$, vous n'avez plus qu'un nombre rationnel en facteur et vous diminuez les risques de vous tromper. Vous pouvez aussi, si vous le préférez, en faire autant sur les lignes sur lesquelles vous allez opérer, mais ne perdez pas de vue que vous allez sans doute compliquer les colonnes suivantes ainsi.

**Exemple 28.** Soit $A = \dfrac{1}{28}\begin{pmatrix} -15 & 2\sqrt{42}+3 & 3\sqrt{42}-2 \\ -2\sqrt{42}+3 & -23 & \sqrt{42}+6 \\ -3\sqrt{42}-2 & -\sqrt{42}+6 & -18 \end{pmatrix}$. Nous vous laissons démontrer que $A$ est une matrice orthogonale, donc l'endomorphisme $f$ de $\mathbb{R}^3$ canoniquement associé à $A$ est une isométrie. Nous précisons sa nature à l'aide de la dimension de $\ker(f - \mathrm{Id}_{\mathbb{R}^3})$. Soit $(a, b, c) \in \mathbb{R}^3$. Alors :

$$A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \iff \left\{\begin{array}{rrrcl} -43a & +(2\sqrt{42}+3)b & +(3\sqrt{42}-2)c & = & 0 \\ (3-2\sqrt{42})a & -51b & +(\sqrt{42}+6)c & = & 0 \\ -(3\sqrt{42}+2)a & +(6-\sqrt{42})b & -46c & = & 0 \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrcll} -43a & +(2\sqrt{42}+3)b & +(3\sqrt{42}-2)c & = & 0 & \\ & -2352b & +\left(31\sqrt{42}+435\right)c & = & 0 & (L_2 \leftarrow 43L_2 + (3-2\sqrt{42})L_1) \\ & -56\sqrt{42}b & -2352c & = & 0 & (L_3 \leftarrow 43L_3 - (3\sqrt{42}+2)L_1) \end{array}\right.$$

où le facteur de $b$ en deuxième ligne (par exemple) s'est simplifié en écrivant : $(3 - 2\sqrt{42})(3 + 2\sqrt{42}) = 3^2 - \left(2\sqrt{42}\right)^2 = 9 - 4 \cdot 42 = -159$. On poursuit :

$$A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \iff \left\{\begin{array}{rrrcll} -43a & +(2\sqrt{42}+3)b & +(3\sqrt{42}-2)c & = & 0 & \\ & -2352b & +\left(31\sqrt{42}+435\right)c & = & 0 & \\ & -2352b & -2352\sqrt{42}c & = & 0 & (L_3 \leftarrow \sqrt{42}L_3) \end{array}\right.$$

Alors l'opération $L_3 \leftarrow L_3 - L_2$ nous mène clairement à un système échelonné avec trois pivots, donc $a = b = c = 0$ et $\dim(\ker(f - \mathrm{Id}_{\mathbb{R}^3})) = 0$. On en déduit que $f$ est la composition commutative d'une rotation et d'une réflexion.

**Exercice 42.** Achever cet exemple pour montrer que $f$ et la composée de la rotation d'axe orienté par $(1, -3{,}2)$ et d'angle $\frac{4\pi}{3} \bmod 2\pi$, et de la réflexion par rapport au plan d'équation $x - 3y + 2z = 0$.

**Mise en garde 6.** Notez bien que si nous avons fait deux théorèmes de classification différents, pour la dimension 2 et la dimension 3, c'est que les isométries sont différentes. Par conséquent, leur étude l'est aussi :

$\boxed{\text{N'UTILISEZ PAS LES MÉTHODES DE LA DIMENSION 3 POUR LA DIMENSION 2 !}}$

Par exemple, si vous trouvez qu'une matrice orthogonale **d'ordre deux** admet un sous-espace $E_1$ de vecteurs invariants de dimension 1, vous aurez l'air ridicules si vous parlez de rotation par rapport à l'axe $E_1$. Ridicules, parce qu'un dessin permet très rapidement de remarquer l'absurdité d'une rotation planaire par rapport à un axe.

### 5.4 Comment commencer un exercice théorique sur les rotations ?

Je parle des rotations d'un espace euclidien en dimension 2 ou 3 : comment utiliser les résultats du cours afin de résoudre les exercices portant sur les rotations ?

Le résultat le plus important, qui est souvent le seul utile pour traiter un exercice, est la forme de la matrice d'une rotation dans toute base orthonormée directe (ou « presque ») :

— **en dimension** 2 **:** la rotation planaire de mesure d'angle $\theta$ a pour matrice $\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$ dans TOUTE base orthonormée directe ;

— **en dimension** 3 **:** la rotation de l'espace d'axe $D$ et de mesure d'angle $\theta$ a pour matrice $\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\theta) & -\sin(\theta)\\ 0 & \sin(\theta) & \cos(\theta)\end{pmatrix}$ dans TOUTE base orthonormée directe de la forme $(\vec{a}, \vec{u}, \vec{v})$, où $\vec{a}$ dirige et oriente $D$, et $(\vec{u}, \vec{v})$ est une base orthonormée directe de $D^{\perp}$.

Notons en passant que la dernière colonne est nécessairement connue, si l'on connaît la ou les précédentes. Par exemple, si la première colonne de la matrice d'une rotation planaire (dans une base orthonormée directe) est : $\begin{pmatrix}\frac{\sqrt{3}}{2}\\ \frac{1}{2}\end{pmatrix}$, la classification implique que la matrice complète est nécessairement : $\begin{pmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}$ (et c'est donc la matrice de la rotation planaire de mesure d'angle $\frac{\pi}{6}$). Or la première colonne d'une matrice de rotation planaire est la même peu importe la base orthonormée directe choisie : **on a donc tout intérêt à choisir pour premier vecteur de la base, un vecteur unitaire dont l'image par** $r$ **est liée aux données de l'énoncé** (soit cette image est connue, soit cette image est à déterminer). Ou bien, un vecteur unitaire dont l'image est facile à calculer grâce à ces mêmes données. Les autres vecteurs peuvent être choisis arbitrairement (à moins que d'autres hypothèses de l'énoncé ne fournissent des vecteurs intéressants orthogonaux au premier), tant que cela donne une base orthonormée directe.

Par exemple, si l'énoncé commence par : « soit $r$ une rotation de $\mathbb{R}^2$, et soit $\vec{u} \in E$ non nul tel que $r(\vec{u})$ vérifie la propriété $\mathcal{P}$ », il est avisé de compléter la famille $\left(\frac{\vec{u}}{\|\vec{u}\|}\right)$ (la division par la norme est pour avoir un vecteur unitaire) en une base orthonormée directe $\left(\frac{\vec{u}}{\|\vec{u}\|}, \vec{v}\right)$ de $\mathbb{R}^2$. Le fait que $r(\vec{u})$ vérifie $\mathcal{P}$ devrait permettre d'obtenir la première colonne de $M_{\left(\frac{\vec{u}}{\|\vec{u}\|}, \vec{v}\right)}(r)$ (et donc aussi la seconde, d'après ce qui précède), puisque cette colonne s'obtient en écrivant les coordonnées de $r(\vec{u})$, sur lesquelles on a normalement des informations. Mais comme on sait que $M_{\left(\frac{\vec{u}}{\|\vec{u}\|}, \vec{v}\right)}(r)$ est aussi de la forme $\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$, avec $\theta$ une mesure d'angle de $r$, en comparant les deux expressions obtenues de cette matrice on peut tirer des informations non triviales sur $r$ (selon ce qu'on nous demande de montrer).

**Exemple 29. (exemple où le choix explicite du premier vecteur suffit)** Soit $r$ une rotation d'un espace vectoriel euclidien orienté $E$ de dimension 2. On suppose qu'il existe $\vec{u} \in E$ non nul tel que : $r(\vec{u}) = -\vec{u}$. On veut déterminer $r$ (c'est-à-dire : trouver son angle).

Pour cela, conformément au conseil ci-dessus (on a une hypothèse sur l'image de $\vec{u}$ par $r$), écrivons la matrice de $r$ dans une base orthonormée directe de la forme $\mathcal{B} = \left(\frac{\vec{u}}{\|\vec{u}\|}, \vec{v}\right)$ (le vecteur $\vec{v}$ n'a pas d'importance ; tout ce qui importe est qu'il existe). Comme $r(\vec{u}) = -\vec{u}$, on a par linéarité : $r\left(\frac{\vec{u}}{\|\vec{u}\|}\right) = -\frac{\vec{u}}{\|\vec{u}\|}$, donc la matrice de $r$ dans la base $\mathcal{B}$ est de la forme : $\begin{pmatrix}-1 & *\\ 0 & *\end{pmatrix}$. Mais, si $\theta \in \mathbb{R}$ est une mesure d'angle de $r$, on sait que sa matrice dans $\mathcal{B}$ est aussi égale à : $\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$. Par identification : $\cos(\theta) = -1$, et donc : $\theta \equiv \pi \bmod 2\pi$. On a donc montré que $r$ est la rotation de mesure d'angle $\pi$ (ou encore : $r = -\mathrm{Id}_E$), d'où le résultat attendu.

**Exemple 30. (exemple où le choix du second vecteur compte aussi)** Soit $r$ une rotation d'un espace vectoriel euclidien orienté $E$ de dimension 2. On suppose qu'il existe deux vecteurs non nuls et orthogonaux $\vec{i}$ et $\vec{j}$ tels que : $r(\vec{i}) = \vec{j}$. On veut montrer qu'on a : $r(\vec{j}) = -\vec{i}$.

On note que l'hypothèse de l'énoncé nous donne explicitement $r(\vec{i})$ : il serait donc avisé de considérer une base orthonormée directe commençant par $\frac{\vec{i}}{\|\vec{i}\|}$, puisque cela permettrait d'expliciter la première colonne... Mais à condition d'inclure $\vec{j}$ à la base également, vu que $r(\vec{i})$ s'exprime à l'aide de $\vec{j}$ (un autre argument qui motive d'intégrer $\vec{j}$ au choix d'une base orthonormée directe, est qu'on veut $r(\vec{j})$ : on a donc tout intérêt à prendre $\vec{j}$ pour second vecteur, de sorte que la deuxième colonne de la matrice de $r$ donne $r(\vec{j})$). Heureusement, le fait que $\vec{i}$ et $\vec{j}$ soient orthogonaux assure que la base $\mathcal{B} = \left(\frac{\vec{i}}{\|\vec{i}\|}, \frac{\vec{j}}{\|\vec{j}\|}\right)$ est orthonormée (éventuellement indirecte ; si elle est indirecte, il suffit de remplacer $\vec{j}$ par $-\vec{j}$, et le lecteur en exercice s'assurera que cela ne change pas le résultat final).

Écrivons la matrice de $r$ dans $\mathcal{B}$ : la deuxième colonne de $r$ nous donnera $r(\vec{j})$, qui est ce qu'on cherche à déterminer. Pour cela, déterminons d'abord la première colonne grâce à $r(\vec{i})$. On a par hypothèse : $r(\vec{i}) = \vec{j}$. Donc, par linéarité : $r\left(\frac{\vec{i}}{\|\vec{i}\|}\right) = \frac{\vec{j}}{\|\vec{i}\|}$. Or $\|\vec{i}\| = \|\vec{j}\|$ car $\vec{j}$ est l'image par l'isométrie $r$ de $\vec{i}$, et une isométrie conserve la norme. Donc : $r\left(\frac{\vec{i}}{\|\vec{i}\|}\right) = \frac{\vec{j}}{\|\vec{j}\|}$. On en déduit que la matrice de $r$ dans $\mathcal{B}$ est de la forme : $\begin{pmatrix} 0 & * \\ 1 & * \end{pmatrix}$. Mais, on l'a vu plus haut, la connaissance de la première colonne permet d'obtenir la seconde. On a nécessairement : $\mathrm{M}_{\mathcal{B}}(r) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. La seconde colonne nous dit alors que $r\left(\frac{\vec{j}}{\|\vec{j}\|}\right) = -\frac{\vec{i}}{\|\vec{i}\|}$, ce qui, après multiplication par $\|\vec{i}\| = \|\vec{j}\|$, implique : $r(\vec{i}) = -\vec{j}$. Ce qu'on voulait démontrer.

**Exemple 31. (exemple où le choix du premier vecteur n'est pas directement soufflé par les hypothèses)** Soient $r$ la rotation de $\mathbb{R}^2$ (muni de l'orientation usuelle) de mesure d'angle $\frac{\pi}{2}$, et $s$ la réflexion par rapport à l'axe $(Oy)$, définie par $(x, y) \mapsto (-x, y)$. Un calcul de déterminant montre que : $\det(s \circ r \circ s) = 1$, et donc $s \circ r \circ s$ est une rotation. On cherche une mesure d'angle de cette rotation.

On pourrait passer par le pur calcul matriciel, en notant que $S(\pi)$ est la matrice de la réflexion $s$ (pourquoi ?), et en calculant $S(\pi)\, R\left(\frac{\pi}{2}\right) S(\pi)$ pour obtenir la matrice de $s \circ r \circ s$. Mais voyons comment l'idée de cette section permet de répondre à cette question.

Construisons une base orthonormée directe de $\mathbb{R}^2$ qui facilite le calcul de la première colonne de la matrice de $s \circ r \circ s$. Pour cela, il vaut mieux choisir un vecteur invariant par $s$ : considérons $\vec{j} = (0{,}1)$ par exemple, comme premier vecteur de notre base orthonormée ; pour le second, $-\vec{i} = (-1{,}0)$ fera l'affaire. La base $\mathcal{B} = \left(\vec{j}, -\vec{i}\right)$ est bien orthonormée directe ; calculons $s \circ r \circ s(\vec{j})$ pour avoir sa première colonne dans la base $\mathcal{B}$. On a : $s(\vec{j}) = \vec{j}$ car $\vec{j}$ est sur l'axe $(Oy)$, donc : $r \circ s(\vec{j}) = r(\vec{j})$. Comme $r$ est la rotation de mesure d'angle $\frac{\pi}{2}$, il est facile de se convaincre qu'on a : $r(\vec{j}) = -\vec{i}$. Alors : $s \circ r \circ s(\vec{i}) = s(-\vec{i}) = -(-\vec{i})$, puisque $-\vec{i}$ est orthogonal à l'axe $(Oy)$. On en déduit que la matrice de $s \circ r \circ s$ dans la base $\mathcal{B}$ est de la forme : $\begin{pmatrix} 0 & * \\ -1 & * \end{pmatrix}$. Mais, si $\theta \in \mathbb{R}$ est une mesure d'angle de $s \circ r \circ s$, on sait que sa matrice dans $\mathcal{B}$ est aussi égale à : $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. Par identification : $\cos(\theta) = 0$, $\sin(\theta) = -1$, donc : $\theta \equiv -\frac{\pi}{2} \bmod 2\pi$. Ceci démontre que $s \circ r \circ s$ est la rotation de mesure d'angle $-\frac{\pi}{2}$.

**Exercice 43.**

1. Redémontrer ce résultat en passant par les matrices de rotation et de réflexion.
2. Refaire cette étude avec une réflexion quelconque de $\mathbb{R}^2$ à la place de $s$, et une rotation de mesure d'angle $\theta \in \mathbb{R}$ quelconque à la place de $r$.

On procède également ainsi dans le cours, pour ces deux résultats :

— montrer que pour tous vecteurs non nuls de même norme, il existe une rotation envoyant l'un sur l'autre ;

— si $r$ est une rotation de mesure d'angle $\theta$, alors pour tout $\vec{u}$ unitaire, on a : $\langle \vec{u}, r(\vec{u}) \rangle = \cos(\theta)$ et $[\vec{u}, r(\vec{u})] = \sin(\theta)$.

Le principe est le même en dimension 3, à ceci près que le premier vecteur $\vec{a}$ de la base orthonormée directe doit orienter l'axe $D$ (pas le choix), tandis que le second doit appartenir au plan orthogonal à l'axe, qu'on note $P = D^{\perp}$ ci-dessous. Alors, même s'il est question d'un vecteur $\vec{u} \in E$ dans l'énoncé dont l'image par $r$ est connue, on ne peut pas l'utiliser pour former une base orthonormée directe de $P$, puis de $\mathbb{R}^3$, à moins que :

— l'on parvienne à démontrer qu'il appartient à $P$ (pour cela, on effectue le produit scalaire avec un vecteur orientant $D$, et on vérifie qu'il est nul ; ceci prouve l'appartenance à $D^{\perp} = P$) ;

— sa décomposition dans la somme directe $E = D \oplus P$ soit connue, ainsi que son image par $r$ (auquel cas, on construit une base orthonormée en commençant par $\vec{a}$, et la composante de $\vec{u}$ dans $P$).

En tous les cas, trouver le dernier vecteur pour obtenir une base orthonormée directe n'est pas le problème majeur : il suffit souvent de prendre $\vec{v} = \vec{a} \wedge \vec{u}$ (si $\vec{u} \in P$), puis de rendre $\vec{v}$ unitaire.

**Exemple 32. (exemple où le vecteur de l'hypothèse est dans $P$)** Soit $r$ une rotation de $\mathbb{R}^3$. On suppose qu'il existe $\vec{u}$ unitaire tel que : $r(\vec{u}) = -\vec{u}$. Comme dans l'exemple 29, on aimerait en déduire que $r$ est la rotation de mesure d'angle $\pi$. Problème : dans le cas de la dimension 2, on n'a aucune contrainte sur la base orthonormée directe où la matrice d'une rotation est connue. Ici, il faut se placer dans une base orthonormée directe dont le premier vecteur oriente l'axe de rotation $D$, et dont les deux autres vecteurs forment une base orthonormée directe de $P = D^{\perp}$. Si $\vec{u} \notin P$, alors le raisonnement de l'exemple 29 ne peut donc pas se transposer tel quel...

Heureusement, nous allons démontrer qu'on a bien $\vec{u} \in P$ (et nous vous invitons à faire un dessin pour vous en convaincre). Soit $\vec{a}$ un vecteur unitaire qui oriente $D$. Pour montrer que $\vec{u} \in P$, il suffit de montrer qu'on a : $\langle \vec{u}, \vec{a} \rangle = 0$. Or une isométrie (ce qu'est une rotation) conserve les produits scalaires, donc :

$$\langle \vec{a}, \vec{u} \rangle = \langle r(\vec{a}), r(\vec{u}) \rangle .$$

On sait que $r(\vec{a}) = \vec{a}$ par définition de l'axe d'une rotation ; on a aussi $r(\vec{u}) = -\vec{u}$ par hypothèse. L'égalité ci-dessus devient donc : $\langle \vec{a}, \vec{u} \rangle = -\langle \vec{a}, \vec{u} \rangle$, puis : $\langle \vec{a}, \vec{u} \rangle = 0$. Ce qu'on voulait démontrer. Ainsi $\vec{u} \in D^{\perp} = P$.

Complétons la famille $(\vec{u})$ de $P$ en une base orthonormée directe de $(\vec{u}, \vec{v})$ de $P$. Alors $\mathcal{B} = (\vec{a}, \vec{u}, \vec{v})$ est une base orthonormée directe de $P$, et la théorie nous dit que la matrice de $r$ dans cette base est de la forme : $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$, où $\theta$ est une mesure d'angle de $r$. Mais on a aussi : $r(\vec{u}) = -\vec{u}$, ce qui nous donne une autre description de la deuxième colonne de la matrice ; on en déduit qu'elle est de la forme : $\begin{pmatrix} 1 & 0 & * \\ 0 & -1 & * \\ 0 & 0 & * \end{pmatrix}$. Par identification : $\cos(\theta) = -1$, donc : $\theta \equiv \pi \bmod 2\pi$, ce qu'on voulait démontrer.

Pour un exemple où le vecteur de l'énoncé n'est pas dans $P$, vous pouvez regarder la démonstration de la formule de Rodrigues, où l'on décompose d'abord un vecteur $\vec{x}$ sous la forme $\vec{x} = \vec{x}_P + \vec{x}_D$, avec $\vec{x}_P \in P$ et $\vec{x}_D \in D$, puis on forme une base orthonormée directe de $P$ en complétant la famille $\left( \frac{\vec{x}_P}{\|\vec{x}_P\|} \right)$ à l'aide d'un produit vectoriel, comme expliqué plus haut.

# Table des matières



# Table des figures