CORRIGÉ – CHAPITRE 19

ESPACES PRÉHILBERTIENS RÉELS

⋆ application directe ⋆⋆ mise en œuvre ⋆⋆⋆ problème ⋆⋆⋆⋆ problème long, en autonomie

Les renvois « proposition n », « théorème n » désignent les énoncés du chapitre 19 du cours ; ceux des chapitres antérieurs sont précisés. On reprend les notations de la feuille : $\langle x, y\rangle$ est le produit scalaire, $\|x\|$ la norme associée, $A^\perp$ l'orthogonal d'une partie A ; $\mathbb{R}^n$ et $\mathcal{M}_{n,1}(\mathbb{R})$ sont munis du produit scalaire canonique $\langle X, Y\rangle = X^TY$, et $\mathcal{M}_{n,p}(\mathbb{R})$ du produit scalaire $\langle A, B\rangle = \mathrm{tr}(A^TB)$ (proposition 1). Pour montrer qu'une application est un produit scalaire, on suit l'ordre recommandé par le cours : symétrie, linéarité par rapport à la première variable, positivité, et le caractère défini en dernier, puisque c'est lui qui demande un argument. Les énoncés n'indiquent pas la technique d'intégration à employer : lorsqu'une intégrale doit être calculée, le corrigé l'annonce au début du calcul.

Produit scalaire

EXERCICE 1 ⋆

1. *L'application φ_1.* Si φ_1 était bilinéaire, l'application $y \mapsto \varphi_1(x, y)$ serait linéaire, donc nulle en $y = 0$. Or, pour $x = (1, 0)$,

$$\varphi_1\big((1,0),(0,0)\big) = \sqrt{1} = 1 \neq 0.$$

φ_1 n'est pas bilinéaire : ce n'est pas un produit scalaire

La racine carrée fait penser à une norme, mais un produit scalaire est une expression du second degré en (x, y), du premier degré en chacune des variables.

L'application φ_2. Elle est de la forme X^TSY avec $S = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$ symétrique, donc bilinéaire et symétrique (section 19.2.3 du cours). Mais, pour $e_2 = (0, 1)$,

$\varphi_2(e_2, e_2) = -1 < 0$: φ_2 n'est pas positive, ce n'est pas un produit scalaire

L'application φ_3. Elle s'écrit X^TSY avec $S = \begin{pmatrix} 1 & -3 \\ -3 & 10 \end{pmatrix}$ symétrique : elle est bilinéaire et symétrique. Pour la positivité, on met $\varphi_3(x, x)$ sous forme canonique, en regroupant les termes en x_1 :

$$\varphi_3(x, x) = x_1^2 - 6x_1x_2 + 10x_2^2 = (x_1 - 3x_2)^2 + x_2^2 \geqslant 0.$$

Si $\varphi_3(x, x) = 0$, les deux carrés sont nuls : $x_2 = 0$, puis $x_1 = 3x_2 = 0$.

φ_3 est un produit scalaire sur $\mathbb{R}^2$

2. L'application φ s'écrit X^TSY avec $S = \begin{pmatrix} 1 & 4 \\ b & a \end{pmatrix}$: elle est bilinéaire quels que soient a et b. Elle est symétrique si et seulement si S l'est, c'est-à-dire si $\varphi(e_1, e_2) = \varphi(e_2, e_1)$, soit $b = 4$. Supposons-le ; la forme canonique donne

$$\varphi(x, x) = x_1^2 + 8x_1x_2 + ax_2^2 = (x_1 + 4x_2)^2 + (a - 16)x_2^2.$$

Si $a > 16$, cette quantité est positive, et nulle seulement si $x_2 = 0$ et $x_1 = -4x_2 = 0$. Si $a \leqslant 16$, le vecteur non nul $x = (-4, 1)$ annule le premier carré et donne $\varphi(x, x) = a - 16 \leqslant 0$: la forme n'est pas définie positive.

φ est un produit scalaire $\iff b = 4$ et $a > 16$

Piège. L'erreur la plus répandue est d'oublier la symétrie et de ne regarder que

$$\varphi(x,x) = x_1^2 + (4+b)x_1x_2 + ax_2^2.$$

Pour $b=0$ et $a=5$, cette quantité vaut $(x_1+2x_2)^2 + x_2^2 > 0$ pour $x \neq 0$, et pourtant φ n'est pas un produit scalaire, puisque $\varphi(e_1,e_2) = 4 \neq 0 = \varphi(e_2,e_1)$. Comme le dit la remarque de la section 19.2.3, la quantité $\varphi(x,x)$ ne voit pas la partie antisymétrique de S.

EXERCICE 2 ⋆

Dans les trois cas, l'expression est symétrique en les deux variables ; elle est linéaire par rapport à la première, car la dérivation $f \mapsto f'$ (ou $P \mapsto P^{(k)}$), l'évaluation en un point et l'intégrale sont linéaires. Il reste à étudier $\langle f,f\rangle$.

(a) Pour $f \in E$, la fonction f' est continue sur $[-1\,,1]$, et

$$\langle f,f\rangle = f(0)^2 + \int_{-1}^{1} f'(t)^2\,\mathrm{d}t \geqslant 0$$

par positivité de l'intégrale (chapitre 18, théorème 3). Si $\langle f,f\rangle = 0$, les deux termes, positifs, sont nuls : $f(0)=0$, et la fonction f'^2, continue, positive et d'intégrale nulle, est nulle (chapitre 18, théorème 4). Alors $f'=0$ sur l'intervalle $[-1\,,1]$, donc f est constante, égale à $f(0)=0$.

$$\boxed{\langle f,g\rangle = f(0)g(0) + \int_{-1}^{1} f'g' \text{ est un produit scalaire sur } \mathcal{C}^1([-1\,,1],\mathbb{R})}$$

Le terme $f(0)g(0)$ est indispensable : sans lui, toute fonction constante non nulle f vérifierait $\langle f,f\rangle = 0$. L'intégrale de f'^2 ne voit f qu'à une constante près, et l'évaluation en 0 fixe cette constante.

(b) Si P est de degré d (ou nul), on a $P^{(k)} = 0$ pour $k > d$: dans la somme, tous les termes d'indice $k > \min(\deg P, \deg Q)$ sont nuls, et elle ne comporte qu'un nombre fini de termes non nuls. Pour $P = \sum_k a_k X^k$, on a $P^{(k)}(0) = k!\,a_k$, donc

$$\langle P,P\rangle = \sum_{k=0}^{+\infty} \left(P^{(k)}(0)\right)^2 = \sum_{k\geqslant 0} (k!)^2 a_k^2 \geqslant 0,$$

et $\langle P,P\rangle = 0$ impose $a_k = 0$ pour tout k, c'est-à-dire $P=0$; c'est la formule de Taylor pour les polynômes (chapitre 13, théorème 6), qui dit que les $P^{(k)}(0)$ déterminent P.

$$\boxed{\langle P,Q\rangle = \sum_{k\geqslant 0} P^{(k)}(0)Q^{(k)}(0) = \sum_{k\geqslant 0} (k!)^2 a_k b_k \text{ est un produit scalaire sur } \mathbb{R}[X]}$$

Ici (b_k) désigne la suite des coefficients de Q. La seconde écriture montre que c'est le produit canonique des suites de coefficients, pondéré par les facteurs $(k!)^2$: la famille $(X^k)_{k\in\mathbb{N}}$ est orthogonale, et $\|X^k\| = k!$.

(c) On a $\langle P,P\rangle = \sum_{k=0}^{n} \left(P^{(k)}(a_k)\right)^2 \geqslant 0$. Supposons $\langle P,P\rangle = 0$: on a $P^{(k)}(a_k) = 0$ pour tout $k \in [\![0\,;n]\!]$. Montrons par récurrence descendante sur k que $P^{(k)} = 0$.

- Comme $\deg P \leqslant n$, le polynôme $P^{(n)}$ est constant ; il est nul en a_n, donc nul.
- Si $P^{(k+1)} = 0$, avec $k \leqslant n-1$, le polynôme $P^{(k)}$ a une dérivée nulle : il est constant, et sa valeur en a_k est nulle, donc $P^{(k)} = 0$.

Pour $k=0$, on obtient $P=0$.

$$\boxed{\langle P,Q\rangle = \sum_{k=0}^{n} P^{(k)}(a_k)Q^{(k)}(a_k) \text{ est un produit scalaire sur } \mathbb{R}_n[X]}$$

Piège. L'hypothèse $\deg P \leqslant n$ sert à l'initialisation, et c'est là que l'espace de travail intervient : sur $\mathbb{R}_{n+1}[X]$, le polynôme X^{n+1} annule tous les $P^{(k)}(a_k)$ si $a_0 = \cdots = a_n = 0$, et la formule n'est plus un produit scalaire. Comme le disait le cours après la proposition 2, l'espace sur lequel on travaille fait partie de la donnée.

EXERCICE 3 ⋆⋆

Quel que soit w, l'application $\langle\cdot,\cdot\rangle_w$ est symétrique, et linéaire par rapport à la première variable par linéarité de l'intégrale : elle est bilinéaire symétrique. Toute la question porte sur le caractère défini positif. On utilisera la fonction « chapeau » associée à un segment $[\alpha\,,\beta]\subset[0\,,1]$, avec $\alpha<\beta$,

$$h(t)=(t-\alpha)(\beta-t)\ \text{ si } t\in[\alpha\,,\beta],\qquad h(t)=0\ \text{ sinon,}$$

qui est continue sur $[0\,,1]$ (les deux expressions valent 0 en α et en β), strictement positive sur $]\alpha\,,\beta[$ et nulle en dehors.

Condition nécessaire. Supposons que $\langle\cdot,\cdot\rangle_w$ est un produit scalaire.

— Si $w(t_0)<0$ pour un $t_0\in[0\,,1]$, la continuité de w en t_0 fournit un segment $[\alpha\,,\beta]\subset[0\,,1]$, avec $\alpha<\beta$ et $t_0\in[\alpha\,,\beta]$, sur lequel $w<0$ (on prend $\varepsilon=\dfrac{|w(t_0)|}{2}$ dans la définition de la continuité). La fonction $-h^2w$ est alors continue, positive, et non nulle au milieu de $[\alpha\,,\beta]$: son intégrale est strictement positive (chapitre 18, théorème 4, sous forme contraposée). Ainsi $\langle h,h\rangle_w<0$, ce qui contredit la positivité.

— Si w est nulle sur un segment $[\alpha\,,\beta]$ avec $\alpha<\beta$, la fonction h associée n'est pas nulle, et h^2w est nulle partout, donc $\langle h,h\rangle_w=0$, ce qui contredit le caractère défini.

Condition suffisante. Supposons $w\geqslant 0$, et que l'ensemble de ses zéros ne contient aucun segment non réduit à un point. Pour $f\in\mathcal{C}([0\,,1],\mathbb{R})$, la fonction f^2w est continue et positive, donc $\langle f,f\rangle_w\geqslant 0$. Si $\langle f,f\rangle_w=0$, le théorème 4 du chapitre 18 donne $f(t)^2w(t)=0$ pour tout t : f est nulle en tout point où w ne s'annule pas. Soit $t_0\in[0\,,1]$. Pour tout $k\geqslant 1$, le segment $J_k=\left[\max(0,t_0-\dfrac{1}{k})\,,\min(1,t_0+\dfrac{1}{k})\right]$ n'est pas réduit à un point, donc il n'est pas contenu dans l'ensemble des zéros de w : il contient un point t_k tel que $w(t_k)\neq 0$, d'où $f(t_k)=0$. Comme $|t_k-t_0|\leqslant\dfrac{1}{k}$, la suite (t_k) tend vers t_0, et la continuité de f donne $f(t_0)=\lim f(t_k)=0$. Ainsi $f=0$.

$$\boxed{\langle\cdot,\cdot\rangle_w \text{ produit scalaire } \iff w\geqslant 0 \text{ et } w \text{ n'est nulle sur aucun segment } [\alpha\,,\beta],\ \alpha<\beta}$$

Comparaison avec la proposition 2. La proposition 2 demande que w ne s'annule qu'en un nombre fini de points : c'est une condition suffisante, plus simple à vérifier, mais pas nécessaire. Le poids défini par $w(t)=t^2\sin^2\left(\dfrac{1}{t}\right)$ pour $t\in]0\,,1]$ et $w(0)=0$ est continu (il est majoré par t^2 en valeur absolue), positif, et s'annule en 0 et en chacun des points $\dfrac{1}{k\pi}$, $k\geqslant 1$: une infinité de zéros, mais aucun segment, puisque ces zéros sont isolés les uns des autres. L'exercice montre qu'il fournit pourtant un produit scalaire. La démonstration de la proposition 2 utilisait d'ailleurs exactement cet argument : chaque zéro de w est limite de points où w ne s'annule pas.

Norme associée et inégalité de Cauchy-Schwarz

EXERCICE 4 ⋆

1. On applique l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$ (proposition 4 (i)) aux n-uplets (x_k) et (2^{-k}) :

$$\Big(\sum_{k=1}^{n} \frac{x_k}{2^k}\Big)^2 \leqslant \Big(\sum_{k=1}^{n} x_k^2\Big)\Big(\sum_{k=1}^{n} \frac{1}{4^k}\Big).$$

La seconde somme est celle d'une suite géométrique de raison $\dfrac{1}{4}$ (chapitre 3) :

$$\sum_{k=1}^{n} \frac{1}{4^k} = \frac{1}{4} \cdot \frac{1-4^{-n}}{1-\frac{1}{4}} = \frac{1-4^{-n}}{3} < \frac{1}{3}.$$

$$\boxed{\Big(\sum_{k=1}^{n} \frac{x_k}{2^k}\Big)^2 \leqslant \frac{1-4^{-n}}{3} \sum_{k=1}^{n} x_k^2 \leqslant \frac{1}{3} \sum_{k=1}^{n} x_k^2}$$

La constante $\dfrac{1}{3}$ ne dépend pas de n : c'est ce qui fait l'intérêt de l'énoncé, et c'est la somme de toute la suite géométrique, que la majoration par $\dfrac{1}{3}$ dispense de calculer exactement.

2. Le choix des deux vecteurs est dicté par la contrainte : on fait apparaître $2x^2+y^2+5z^2$ comme le carré de la norme du premier, en écrivant

$$x+y+z = \sqrt{2}x \cdot \frac{1}{\sqrt{2}} + y \cdot 1 + \sqrt{5}z \cdot \frac{1}{\sqrt{5}}.$$

La proposition 4 (i), appliquée à $(\sqrt{2}x, y, \sqrt{5}z)$ et $\Big(\dfrac{1}{\sqrt{2}}, 1, \dfrac{1}{\sqrt{5}}\Big)$, donne

$$(x+y+z)^2 \leqslant (2x^2+y^2+5z^2)\Big(\frac{1}{2}+1+\frac{1}{5}\Big) \leqslant 1 \times \frac{17}{10}, \qquad \boxed{(x+y+z)^2 \leqslant \frac{17}{10}}$$

La majoration ne peut pas être améliorée : il y a égalité dans l'inégalité de Cauchy-Schwarz lorsque $(\sqrt{2}x, y, \sqrt{5}z)$ est proportionnel à $\Big(\dfrac{1}{\sqrt{2}}, 1, \dfrac{1}{\sqrt{5}}\Big)$, c'est-à-dire $(x,y,z) = c\Big(\dfrac{1}{2}, 1, \dfrac{1}{5}\Big)$, et égalité dans la contrainte lorsque $2x^2+y^2+5z^2 = \dfrac{17c^2}{10} = 1$. Pour $c = \sqrt{\dfrac{10}{17}}$,

$$\boxed{(x,y,z) = \sqrt{\frac{10}{17}}\Big(\frac{1}{2}, 1, \frac{1}{5}\Big) : \quad 2x^2+y^2+5z^2 = 1, \quad (x+y+z)^2 = \frac{10}{17} \cdot \frac{289}{100} = \frac{17}{10}}$$

À retenir. Le cas d'égalité du théorème 1 transforme une majoration en un maximum : $\dfrac{17}{10}$ est la plus grande valeur de $(x+y+z)^2$ sur l'ellipsoïde plein d'équation $2x^2+y^2+5z^2 \leqslant 1$. L'erreur la plus répandue est d'appliquer Cauchy-Schwarz à (x,y,z) et $(1,1,1)$, ce qui fait apparaître $x^2+y^2+z^2$, que la contrainte ne contrôle pas bien.

EXERCICE 5 ⋆

1. On applique la proposition 4 (i) dans $\mathbb{R}^{n+1}$ aux $(n+1)$-uplets $\Big(\sqrt{\binom{n}{k}}\Big)_{0\leqslant k\leqslant n}$ et $(1,\dots,1)$, puis la formule du binôme (chapitre 3) :

$$\Big(\sum_{k=0}^{n}\sqrt{\binom{n}{k}}\cdot 1\Big)^2 \leqslant \Big(\sum_{k=0}^{n}\binom{n}{k}\Big)(n+1) = 2^n(n+1).$$

On conclut par croissance de la racine carrée.

$$\boxed{\sum_{k=0}^{n}\sqrt{\binom{n}{k}} \leqslant 2^{n/2}\sqrt{n+1}}$$

Il y a égalité si et seulement si les $\binom{n}{k}$ sont tous égaux, c'est-à-dire pour $n=0$ et $n=1$; pour $n=2$, on compare $2+\sqrt{2}\approx 3{,}414$ et $2\sqrt{3}\approx 3{,}464$ ✓.

2. La trace s'écrit comme un produit scalaire avec la matrice identité : $\langle I_n, A\rangle = \operatorname{tr}({I_n}^T A) = \operatorname{tr} A$, et $\|I_n\|^2 = \operatorname{tr}(I_n) = n$. Le théorème 1 donne

$$\boxed{|\operatorname{tr} A| = |\langle I_n, A\rangle| \leqslant \|I_n\|\,\|A\| = \sqrt{n}\,\|A\|}$$

avec égalité si et seulement si A est colinéaire à I_n, c'est-à-dire si A est une matrice scalaire λI_n.

3. Pour A, $B\in\mathcal{S}_n(\mathbb{R})$, on a $A^T=A$, donc

$$\operatorname{tr}(AB) = \operatorname{tr}(A^TB) = \langle A,B\rangle \qquad \text{et} \qquad \operatorname{tr}(A^2) = \operatorname{tr}(A^TA) = \|A\|^2;$$

de plus $\operatorname{tr}(BA)=\operatorname{tr}(AB)$ (chapitre 16, proposition 14). Ainsi

$$\big(\operatorname{tr}(AB+BA)\big)^2 = \big(2\langle A,B\rangle\big)^2 = 4\langle A,B\rangle^2 \leqslant 4\|A\|^2\|B\|^2, \qquad \boxed{\big(\operatorname{tr}(AB+BA)\big)^2 \leqslant 4\operatorname{tr}(A^2)\operatorname{tr}(B^2)}$$

Piège. La symétrie de A et de B est indispensable : pour $A=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$, on trouve $\operatorname{tr}(A^2)=-2<0$, et l'inégalité avec $B=A$ s'écrirait $16\leqslant 16$, mais avec $B=I_2$ elle donnerait $0\leqslant 4\times(-2)\times 2$, ce qui est faux. L'erreur la plus répandue est de croire que $\operatorname{tr}(A^2)$ est un carré de norme pour toute matrice ; c'est $\operatorname{tr}(A^TA)$ qui l'est.

EXERCICE 6 ⋆

On développe chaque terme en faisant apparaître $x-m$, par la proposition 3 (ii) :

$$\|x-x_i\|^2 = \|(x-m)+(m-x_i)\|^2 = \|x-m\|^2 + 2\langle x-m, m-x_i\rangle + \|m-x_i\|^2.$$

En sommant sur i, par linéarité du produit scalaire par rapport à la seconde variable,

$$f(x) = p\|x-m\|^2 + 2\Big\langle x-m, \sum_{i=1}^{p}(m-x_i)\Big\rangle + f(m), \qquad \text{avec} \qquad \sum_{i=1}^{p}(m-x_i) = pm - \sum_{i=1}^{p}x_i = 0.$$

$$\boxed{f(x) = f(m) + p\|x-m\|^2 \quad \text{pour tout } x\in E}$$

On en déduit $f(x)\geqslant f(m)$ pour tout x, avec égalité si et seulement si $\|x-m\|=0$, c'est-à-dire $x=m$ (proposition 3 (i)).

$$\boxed{f \text{ atteint son minimum } f(m) \text{ en l'unique point } m = \frac{1}{p}\sum_{i=1}^{p}x_i}$$

Par exemple, dans $\mathbb{R}^2$, avec $x_1=(0,0)$, $x_2=(4,0)$ et $x_3=(0,2)$, on trouve $m=\left(\frac{4}{3},\frac{2}{3}\right)$, $f(m)=\frac{40}{3}$, et $f(0)=0+16+4=20=\frac{40}{3}+3\times\frac{20}{9}$ ✓.

À retenir. L'isobarycentre est le point qui minimise la somme des carrés des distances aux x_i. Le même calcul, avec des masses, est le théorème de Huygens de la mécanique : le moment d'inertie par rapport à un point est minimal au centre de gravité. Le seul ingrédient est le double produit de la proposition 3 (ii), qui disparaît parce que $\sum_i (m-x_i)=0$; la droite des moindres carrés de la proposition 14 repose sur la même idée.

EXERCICE 7 ⋆⋆

1. Comme f est de classe $\mathcal{C}^1$ et $f(a)=0$, le théorème fondamental de l'analyse (chapitre 18, théorème 6) donne, pour $t\in[a\,,b]$,

$$f(t)=f(a)+\int_a^t f'(u)\,\mathrm{d}u=\int_a^t 1\cdot f'(u)\,\mathrm{d}u.$$

Pour $t=a$, l'inégalité demandée s'écrit $0\leqslant 0$. Pour $t>a$, l'inégalité de Cauchy-Schwarz sur $[a\,,t]$ (proposition 4 (ii)), appliquée aux fonctions continues $u\mapsto 1$ et f', donne

$$\boxed{f(t)^2\leqslant\Big(\int_a^t 1\,\mathrm{d}u\Big)\Big(\int_a^t f'(u)^2\,\mathrm{d}u\Big)=(t-a)\int_a^t f'(u)^2\,\mathrm{d}u}$$

2. Notons $K=\displaystyle\int_a^b f'^2$. Comme f'^2 est positive, la relation de Chasles et la positivité de l'intégrale donnent $\displaystyle\int_a^t f'^2\leqslant K$ pour tout $t\in[a\,,b]$, donc $f(t)^2\leqslant(t-a)K$. On intègre cette inégalité entre fonctions continues sur $[a\,,b]$ (chapitre 18, théorème 3) :

$$\int_a^b f(t)^2\,\mathrm{d}t\leqslant K\int_a^b(t-a)\,\mathrm{d}t=K\,\frac{(b-a)^2}{2},\qquad\boxed{\int_a^b f^2\leqslant\frac{(b-a)^2}{2}\int_a^b f'^2}$$

À retenir. Une fonction nulle en a ne peut pas être grande sans que sa dérivée le soit : elle est contrôlée par l'intégrale de f'^2. La constante $\frac{1}{2}$ n'est pas la meilleure possible. Pour $a=0$, $b=1$ et $f(t)=\sin\frac{\pi t}{2}$, le rapport $\displaystyle\int_0^1 f^2\Big/\int_0^1 f'^2$ vaut $\frac{4}{\pi^2}\approx 0{,}405$: la meilleure constante est donc comprise entre $\frac{4}{\pi^2}$ et $\frac{1}{2}$. Le facteur $(b-a)^2$ est en revanche le bon : il se lit par homogénéité, en remplaçant f par $t\mapsto f(a+(b-a)t)$.

EXERCICE 8 ⋆⋆

L'application est symétrique, puisque l'échange de f et g échange les deux termes $f(1)g(0)$ et $f(0)g(1)$; elle est linéaire par rapport à f, par linéarité de la dérivation, de l'évaluation et de l'intégrale. Pour $f\in\mathcal{C}^1([0\,,1],\mathbb{R})$,

$$\langle f,f\rangle=\int_0^1 f'^2+2f(0)f(1).$$

Piège. Le terme $2f(0)f(1)$ peut être négatif : la positivité n'est pas évidente, et l'erreur la plus répandue est de l'affirmer en invoquant la seule positivité de l'intégrale. Il faut minorer $\displaystyle\int_0^1 f'^2$ à l'aide de $f(0)$ et $f(1)$.

Par le théorème fondamental de l'analyse, $\displaystyle\int_0^1 f'=f(1)-f(0)$, et l'inégalité de Cauchy-Schwarz (proposition 4 (ii)), appliquée à 1 et f', donne

$$\big(f(1)-f(0)\big)^2=\Big(\int_0^1 1\cdot f'\Big)^2\leqslant\int_0^1 1\cdot\int_0^1 f'^2=\int_0^1 f'^2.$$

Par conséquent,

$$\boxed{\langle f,f\rangle \geqslant \big(f(1)-f(0)\big)^2 + 2f(0)f(1) = f(0)^2+f(1)^2 \geqslant 0}$$

Si $\langle f,f\rangle = 0$, on a $f(0)=f(1)=0$, puis $\langle f,f\rangle = \displaystyle\int_0^1 f'^2 = 0$; la fonction f'^2, continue, positive, d'intégrale nulle, est nulle (chapitre 18, théorème 4), donc f est constante, égale à $f(0)=0$.

$$\boxed{\langle f,g\rangle = \int_0^1 f'g' + f(1)g(0) + f(0)g(1) \ \text{ est un produit scalaire sur } \mathcal{C}^1([0\,,1],\mathbb{R})}$$

Pour $f(t)=2t-1$, on trouve $\langle f,f\rangle = 4-2 = 2 = f(0)^2+f(1)^2$ ✓ : la minoration est atteinte, parce que f' est constante, donc proportionnelle à 1 (cas d'égalité de Cauchy-Schwarz). L'exercice 2 (a) traitait un produit du même type, où la positivité était immédiate ; ici, c'est Cauchy-Schwarz qui la fournit.

EXERCICE 9 ⋆⋆

1. Soit $X=(x_0,\dots,x_n)^T$. En développant le produit matriciel, puis par linéarité de l'intégrale,

$$X^T H_n X = \sum_{i=0}^n\sum_{j=0}^n x_i x_j s_{i+j} = \int_0^1 \sum_{i=0}^n\sum_{j=0}^n x_i x_j t^{i+j} f(t)\,\mathrm{d}t = \int_0^1 \Big(\sum_{i=0}^n x_i t^i\Big)^2 f(t)\,\mathrm{d}t.$$

La fonction intégrée est continue et positive, puisque $f\geqslant 0$.

$$\boxed{X^T H_n X = \int_0^1 P(t)^2 f(t)\,\mathrm{d}t \geqslant 0, \qquad P=\sum_{i=0}^n x_i X^i}$$

Autrement dit, H_n est la matrice des $\langle X^i, X^j\rangle$ pour l'application bilinéaire, symétrique et positive

$$(P,Q)\longmapsto \int_0^1 P(t)Q(t)f(t)\,\mathrm{d}t.$$

2. Soit $p,q\in\mathbb{N}$. Si $p=q$, l'inégalité est une égalité. Sinon, on prend $n=\max(p,q)$ et la colonne X dont la coordonnée d'indice p vaut $\lambda\in\mathbb{R}$, celle d'indice q vaut 1, et les autres sont nulles. La question 1 donne

$$X^T H_n X = \lambda^2 s_{2p} + 2\lambda s_{p+q} + s_{2q} \geqslant 0 \qquad \text{pour tout réel } \lambda.$$

Si $s_{2p}>0$, c'est un trinôme du second degré de signe constant : son discriminant est négatif ou nul, $4s_{p+q}^2 - 4s_{2p}s_{2q}\leqslant 0$. Si $s_{2p}=0$, la fonction affine $\lambda\mapsto 2\lambda s_{p+q}+s_{2q}$ est positive sur $\mathbb{R}$, donc constante, et $s_{p+q}=0$: l'inégalité est encore vraie.

$$\boxed{s_{p+q}^2 \leqslant s_{2p}\,s_{2q}}$$

C'est la démonstration du théorème 1 du cours, refaite sur la matrice H_n. Directement : la fonction $\sqrt{f}$ est continue sur $[0\,,1]$, et la proposition 4 (ii), appliquée aux fonctions $t\mapsto t^p\sqrt{f(t)}$ et $t\mapsto t^q\sqrt{f(t)}$, donne

$$s_{p+q}^2 = \Big(\int_0^1 t^p\sqrt{f(t)}\cdot t^q\sqrt{f(t)}\,\mathrm{d}t\Big)^2 \leqslant \int_0^1 t^{2p}f(t)\,\mathrm{d}t\int_0^1 t^{2q}f(t)\,\mathrm{d}t = s_{2p}\,s_{2q}.$$

Pour $f=1$, on a $s_k=\dfrac{1}{k+1}$, et l'inégalité s'écrit $(2p+1)(2q+1)\leqslant (p+q+1)^2$, c'est-à-dire $0\leqslant (p-q)^2$ ✓.

À retenir. Lorsque f ne s'annule qu'en un nombre fini de points, $(P,Q)\mapsto \displaystyle\int_0^1 PQf$ est un produit scalaire sur $\mathbb{R}[X]$ (proposition 2) et l'inégalité de la question 1 est stricte pour $X\neq 0$. Pour $f=1$, H_n est la matrice de Hilbert $\Big(\dfrac{1}{i+j+1}\Big)$, dont l'exercice 27 exploitera ce point de vue.

EXERCICE 10 ⋆⋆

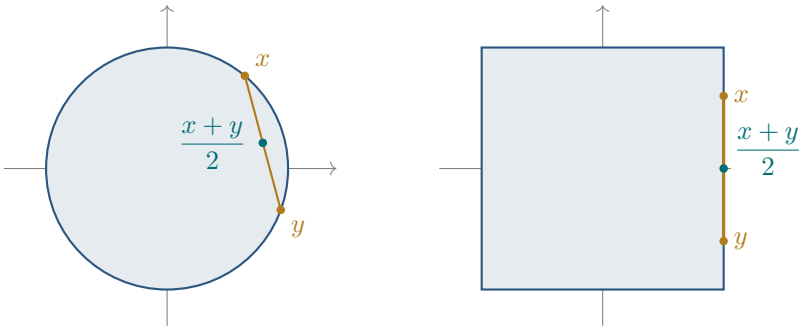
1. Par l'inégalité triangulaire (théorème 2) et l'homogénéité de la norme,

$$\|(1-t)x+ty\| \leqslant (1-t)\|x\| + t\|y\| \leqslant (1-t)+t = 1.$$

Supposons qu'il y ait égalité. La seconde inégalité est alors une égalité, et, comme $1-t>0$ et $t>0$, cela impose $\|x\| = \|y\| = 1$. La première est aussi une égalité : par le cas d'égalité du théorème 2, appliqué aux vecteurs $(1-t)x \neq 0$ et ty, il existe un réel $\lambda \geqslant 0$ tel que $ty = \lambda(1-t)x$, donc $y = \mu x$ avec $\mu \geqslant 0$. Comme $\|y\| = \|x\| = 1$, on obtient $\mu = 1$ et $y = x$, ce qui est exclu.

$$\boxed{\|(1-t)x+ty\| < 1}$$

Géométriquement, les points $(1-t)x+ty$, pour $t \in]0\,,1[$, décrivent le segment ouvert d'extrémités x et y. La boule unité fermée est donc strictement convexe : tout segment joignant deux de ses points est contenu, extrémités exceptées, dans la boule ouverte $\{z \mid \|z\| < 1\}$, et la sphère unité ne contient aucun segment.



norme euclidienne $\qquad\qquad$ $N(x) = \max(|x_1|, |x_2|)$

À gauche, la boule unité de la norme euclidienne de $\mathbb{R}^2$: le milieu de deux points du cercle est strictement à l'intérieur. À droite, celle de la norme $N(x) = \max(|x_1|, |x_2|)$, qui est un carré : le segment joignant $x = (1,0{,}6)$ et $y = (1,-0{,}6)$ reste sur le bord, ce qui interdit à N de dériver d'un produit scalaire. C'est l'argument de la question 2.

2. Pour f continue sur le segment $[0\,,1]$, la fonction $|f|$ est continue et atteint son maximum (chapitre 8, théorème 10) : $N_\infty(f)$ est un réel positif bien défini. Si $N_\infty(f) = 0$, alors $|f(t)| \leqslant 0$ pour tout t, et $f = 0$. Pour $\lambda \in \mathbb{R}$, $|\lambda f(t)| = |\lambda|\,|f(t)|$ et le maximum de $|\lambda|\,|f|$ est $|\lambda| N_\infty(f)$. Enfin, pour tout t,

$$|f(t)+g(t)| \leqslant |f(t)| + |g(t)| \leqslant N_\infty(f) + N_\infty(g),$$

et, en prenant pour t un point où $|f+g|$ atteint son maximum, $N_\infty(f+g) \leqslant N_\infty(f) + N_\infty(g)$.

$$\boxed{N_\infty \text{ est une norme sur } \mathcal{C}([0\,,1],\mathbb{R})}$$

Soit $f = 1$ et $g : t \mapsto t$. Ces deux fonctions sont distinctes, $N_\infty(f) = N_\infty(g) = 1$, et, pour $t = \dfrac{1}{2}$, la fonction $\dfrac{1}{2}f + \dfrac{1}{2}g : t \mapsto \dfrac{1+t}{2}$ a pour maximum 1, atteint en $t = 1$. Si N_∞ dérivait d'un produit scalaire, la question 1 donnerait $N_\infty\Big(\dfrac{f+g}{2}\Big) < 1$: contradiction.

$$\boxed{N_\infty \text{ ne dérive pas d'un produit scalaire}}$$

On peut aussi le voir par l'identité du parallélogramme (proposition 3 (iv)), qui est en défaut :

$$N_\infty(f+g)^2 + N_\infty(f-g)^2 = 4+1 = 5 \neq 4 = 2N_\infty(f)^2 + 2N_\infty(g)^2.$$

3. *N_1 est une norme.* Pour $P \in \mathbb{R}[X]$, la fonction $|P|$ est continue sur $[a\,,b]$ et positive, donc $N_1(P) \geqslant 0$. Si $N_1(P) = 0$, le théorème 4 du chapitre 18 montre que P s'annule en tout point de $[a\,,b]$: il a une infinité de racines, donc il est nul (chapitre 13, théorème 3). L'homogénéité vient de la linéarité de l'intégrale, et l'inégalité triangulaire de $|P+Q| \leqslant |P| + |Q|$ et de la croissance de l'intégrale.

N_1 ne dérive pas d'un produit scalaire. Si c'était le cas, elle vérifierait l'identité du parallélogramme (proposition 3 (iv)). Prenons $P = 1$ et le polynôme $Q = \dfrac{2X-a-b}{b-a}$, qui croît de -1 à 1 sur $[a\,,b]$ et s'annule au milieu. Les fonctions $|Q|$, $1+Q = \dfrac{2(X-a)}{b-a}$ et $1-Q = \dfrac{2(b-X)}{b-a}$ sont positives et affines par morceaux sur $[a\,,b]$; leurs intégrales sont des aires de triangles :

$$N_1(1) = b-a, \qquad N_1(Q) = 2\times\frac{1}{2}\cdot\frac{b-a}{2}\cdot 1 = \frac{b-a}{2}, \qquad N_1(1+Q) = N_1(1-Q) = \frac{1}{2}(b-a)\times 2 = b-a.$$

$$\boxed{N_1(P+Q)^2 + N_1(P-Q)^2 = 2(b-a)^2 \neq \frac{5}{2}(b-a)^2 = 2N_1(P)^2 + 2N_1(Q)^2}$$

La norme du maximum. Notons $N(P) = \max\limits_{[a,b]} |P|$, bien défini par le théorème des bornes atteintes. Les arguments de la question 2 montrent l'homogénéité et l'inégalité triangulaire, et $N(P) = 0$ entraîne que P s'annule sur $[a\,,b]$, donc que $P = 0$, comme pour N_1. Avec les mêmes P et Q, on a $N(P) = N(Q) = 1$ et $N(P+Q) = N(P-Q) = 2$, donc

$$\boxed{N(P+Q)^2 + N(P-Q)^2 = 8 \neq 4 = 2N(P)^2 + 2N(Q)^2}$$

et N est une norme sur $\mathbb{R}[X]$ qui ne dérive pas d'un produit scalaire.
Piège. Sur $\mathbb{R}[X]$, c'est le théorème 3 du chapitre 13 qui rend N_1 et N définies : un polynôme est déterminé par ses valeurs sur un segment. Sur l'espace de toutes les fonctions continues par morceaux, N_1 ne serait plus définie, pour la même raison qu'à la remarque qui suit la proposition 2. L'identité du parallélogramme est une condition nécessaire pour qu'une norme dérive d'un produit scalaire ; l'exercice suivant montre qu'elle est suffisante.

EXERCICE 11 ⋆⋆⋆

On note d'abord que $N(0) = N(0\cdot 0) = 0\cdot N(0) = 0$ et que $N(-x) = |-1|\,N(x) = N(x)$.
1. Comme $N(y-x) = N(x-y)$,

$$\varphi(y,x) = \frac{1}{4}(N(y+x)^2 - N(y-x)^2) = \frac{1}{4}(N(x+y)^2 - N(x-y)^2) = \varphi(x,y).$$

Ensuite $\varphi(x,x) = \dfrac{1}{4}(N(2x)^2 - N(0)^2) = \dfrac{1}{4}\cdot 4N(x)^2$, et $\varphi(0,y) = \dfrac{1}{4}(N(y)^2 - N(-y)^2) = 0$.

$$\boxed{\varphi(y,x) = \varphi(x,y), \qquad \varphi(x,x) = N(x)^2, \qquad \varphi(0,y) = 0}$$

2. On applique l'identité du parallélogramme deux fois : aux vecteurs $x+y$ et $x'+y$, puis aux vecteurs $x-y$ et $x'-y$. Leur somme et leur différence valent respectivement $x+x' \pm 2y$ et $x-x'$:

$$N(x+x'+2y)^2 + N(x-x')^2 = 2N(x+y)^2 + 2N(x'+y)^2,$$

$$N(x+x'-2y)^2 + N(x-x')^2 = 2N(x-y)^2 + 2N(x'-y)^2.$$

Le terme $N(x-x')^2$, qui ne contient pas y, disparaît par soustraction :

$$N(x+x'+2y)^2 - N(x+x'-2y)^2 = 2(N(x+y)^2 - N(x-y)^2) + 2(N(x'+y)^2 - N(x'-y)^2),$$

c'est-à-dire

$$4\varphi(x+x',2y) = 8\varphi(x,y) + 8\varphi(x',y).$$

Enfin, par homogénéité, $N(x+x'\pm 2y) = 2N\Big(\dfrac{x+x'}{2}\pm y\Big)$, donc $\varphi(x+x',2y) = 4\varphi\Big(\dfrac{x+x'}{2},y\Big)$, et

$$\boxed{\varphi(x,y)+\varphi(x',y) = \frac{1}{2}\varphi(x+x',2y) = 2\varphi\Big(\frac{x+x'}{2},y\Big)}$$

Pour $x'=0$, comme $\varphi(0,y)=0$, on obtient $\varphi(x,y) = 2\varphi\Big(\dfrac{x}{2},y\Big)$ pour tout x ; appliquée au vecteur $x+x'$, cette relation donne $2\varphi\Big(\dfrac{x+x'}{2},y\Big) = \varphi(x+x',y)$, et

$$\boxed{\varphi(x+x',y) = \varphi(x,y)+\varphi(x',y)}$$

3. Fixons x et y. Pour $\lambda\in\mathbb{N}$, on raisonne par récurrence : $\varphi(0\cdot x,y)=0$, et, si $\varphi(\lambda x,y)=\lambda\varphi(x,y)$, l'additivité donne $\varphi((\lambda+1)x,y) = \varphi(\lambda x,y)+\varphi(x,y) = (\lambda+1)\varphi(x,y)$. Pour $\lambda\in\mathbb{Z}$ négatif, l'additivité donne $\varphi(\lambda x,y)+\varphi(-\lambda x,y) = \varphi(0,y) = 0$, d'où $\varphi(\lambda x,y) = -(-\lambda)\varphi(x,y) = \lambda\varphi(x,y)$. Pour $\lambda = \dfrac{r}{s}$, avec $r\in\mathbb{Z}$ et $s\in\mathbb{N}^*$, le cas entier appliqué au vecteur λx et à l'entier s, puis au vecteur x et à l'entier r, donne

$$s\,\varphi(\lambda x,y) = \varphi(s\lambda x,y) = \varphi(rx,y) = r\,\varphi(x,y), \qquad \boxed{\varphi(\lambda x,y) = \lambda\varphi(x,y) \quad (\lambda\in\mathbb{Q})}$$

4. L'inégalité triangulaire donne $N(u) = N\big((u-v)+v\big) \leqslant N(u-v)+N(v)$, et, en échangeant u et v, $N(v)-N(u)\leqslant N(v-u) = N(u-v)$.

$$\boxed{|N(u)-N(v)| \leqslant N(u-v)}$$

Pour λ, μ réels, on en déduit $|N(\lambda x+y)-N(\mu x+y)| \leqslant N\big((\lambda-\mu)x\big) = |\lambda-\mu|\,N(x)$: la fonction $\lambda\mapsto N(\lambda x+y)$ est $N(x)$-lipschitzienne, donc continue sur $\mathbb{R}$, et de même pour $\lambda\mapsto N(\lambda x-y)$. Par opérations sur les fonctions continues,

$$\boxed{\lambda\longmapsto \varphi(\lambda x,y) = \frac{1}{4}\big(N(\lambda x+y)^2 - N(\lambda x-y)^2\big) \text{ est continue sur } \mathbb{R}}$$

Conclusion. Soit $\lambda\in\mathbb{R}$. Par densité de $\mathbb{Q}$ (chapitre 7, théorème 2 et proposition 12), il existe une suite (r_k) de rationnels qui tend vers λ. Par la question 3, $\varphi(r_k x,y) = r_k\varphi(x,y)$, et le passage à la limite, licite par continuité des deux membres en λ, donne $\varphi(\lambda x,y) = \lambda\varphi(x,y)$. Avec l'additivité, φ est linéaire par rapport à sa première variable, donc bilinéaire, puisqu'elle est symétrique. Enfin $\varphi(x,x) = N(x)^2 \geqslant 0$, avec égalité seulement si $x=0$.

$$\boxed{\varphi \text{ est un produit scalaire sur } E, \text{ de norme associée } \sqrt{\varphi(x,x)} = N(x)}$$

À retenir. Une norme dérive d'un produit scalaire si et seulement si elle vérifie l'identité du parallélogramme, et ce produit scalaire est alors donné par la seconde identité de polarisation (proposition 3 (iii)), qui était la seule définition possible. La difficulté est l'homogénéité : l'additivité ne donne la linéarité que sur $\mathbb{Q}$, et c'est la continuité, tirée de l'inégalité triangulaire, qui permet de passer à $\mathbb{R}$. Le même schéma « entiers, rationnels, puis réels par densité » résolvait, dans l'exemple qui suit la proposition 11 du chapitre 8, l'équation $f(x+y) = f(x)+f(y)$ pour f continue.

EXERCICE 12 ⋆

1. Notons $a = (1,-1,2,1)$ et $b = (-1,2,3,-1)$: les équations de F s'écrivent $\langle a,x\rangle = 0$ et $\langle b,x\rangle = 0$, donc $F = \{a,b\}^\perp$, et a, b sont orthogonaux à tout vecteur de F : $\mathrm{Vect}(a,b) \subset F^\perp$. Inversement, les vecteurs a et b ne sont pas colinéaires, donc le système est de rang 2, et $\dim F = 4-2 = 2$ (chapitre 15, proposition 9). Comme la somme $F + F^\perp$ est directe (proposition 5 (iv)), la formule de Grassmann (chapitre 15, théorème 7) donne $\dim F + \dim F^\perp \leqslant 4$, donc $\dim F^\perp \leqslant 2 = \dim \mathrm{Vect}(a,b)$, et l'inclusion est une égalité (chapitre 15, théorème 5).

$$\boxed{F^\perp = \mathrm{Vect}\big((1,-1,2,1),(-1,2,3,-1)\big), \quad \text{de base } \big((1,-1,2,1),(-1,2,3,-1)\big)}$$

Les coefficients des équations de F fournissent directement une base de $F^\perp$: c'est le premier exemple qui suit la proposition 5, qui lisait l'équation d'un hyperplan comme une condition d'orthogonalité.

2. Par la proposition 5 (ii), $G^\perp$ est l'orthogonal des deux vecteurs générateurs : un vecteur (x,y,z,t) appartient à $G^\perp$ si et seulement si

$$\begin{cases} x+y+z=0 \\ 2x+y+z-t=0 \end{cases} \iff \begin{cases} y=-x-z \\ t=x \end{cases} \iff (x,y,z,t) = x(1,-1,0,1) + z(0,-1,1,0).$$

$$\boxed{G^\perp = \mathrm{Vect}\big((0,-1,1,0),(1,-1,0,1)\big)}$$

Ces deux vecteurs ne sont pas colinéaires, et l'on contrôle qu'ils sont orthogonaux à $(1,1,1,0)$ et à $(2,1,1,-1)$ ✓.

3. On procède en sens inverse de la question 1 : on cherche d'abord une base de H, dont les vecteurs fourniront les équations de $H^\perp$. Par soustraction des deux équations, $y + 4z = 0$, puis $x = -y + z - t = 5z - t$:

$$(x,y,z,t) \in H \iff (x,y,z,t) = z(5,-4,1,0) + t(-1,0,0,1), \qquad H = \mathrm{Vect}\big((5,-4,1,0),(-1,0,0,1)\big).$$

Par la proposition 5 (ii), $H^\perp$ est l'ensemble des vecteurs orthogonaux à ces deux générateurs :

$$\boxed{H^\perp : \begin{cases} 5x-4y+z=0 \\ -x+t=0 \end{cases}}$$

Par la question 1, $H^\perp = \mathrm{Vect}\big((1,1,-1,1),(1,2,3,1)\big)$, et ces deux vecteurs vérifient bien les deux équations : $5-4-1=0$, $-1+1=0$, $5-8+3=0$ et $-1+1=0$ ✓.

À retenir. Le passage à l'orthogonal échange les deux descriptions d'un sous-espace : des équations de F on lit une famille génératrice de $F^\perp$, et d'une famille génératrice de F on lit des équations de $F^\perp$. L'erreur la plus répandue est de recopier les équations de H comme équations de $H^\perp$: ce sont les coordonnées de vecteurs de $H^\perp$, et non des équations.

EXERCICE 13 ⋆

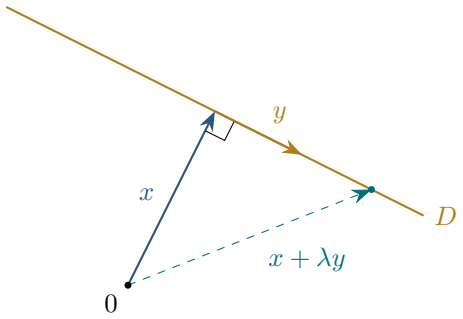
1. Par la proposition 3 (ii), pour tout réel λ,

$$\|x+\lambda y\|^2 - \|x\|^2 = 2\lambda\langle x,y\rangle + \lambda^2\|y\|^2.$$

Si $\langle x,y\rangle = 0$, cette quantité vaut $\lambda^2\|y\|^2 \geqslant 0$ pour tout λ. Réciproquement, supposons-la positive pour tout λ. Si $y=0$, on a $\langle x,y\rangle = 0$. Sinon, on choisit $\lambda = -\dfrac{\langle x,y\rangle}{\|y\|^2}$, valeur qui minimise le trinôme, et l'on obtient $-\dfrac{\langle x,y\rangle^2}{\|y\|^2} \geqslant 0$, donc $\langle x,y\rangle = 0$.

$$\boxed{x \perp y \iff \|x+\lambda y\| \geqslant \|x\| \quad \text{pour tout } \lambda \in \mathbb{R}}$$

Géométriquement, les vecteurs $x+\lambda y$ décrivent la droite D passant par l'extrémité de x et dirigée par y ; $\|x+\lambda y\|$ est la distance de l'origine au point de paramètre λ. L'énoncé dit que ce point est le plus proche de l'origine pour $\lambda = 0$ si et seulement si x est perpendiculaire à D : la perpendiculaire est le plus court chemin d'un point à une droite.



Avec $x=(1,2)$ et $y=(2,-1)$, orthogonaux, le point $x+0{,}9y = (2{,}8\,;1{,}1)$ de la droite D est à la distance $\sqrt{9{,}05} \approx 3{,}01$ de l'origine, plus grande que $\|x\| = \sqrt{5} \approx 2{,}24$: c'est le théorème de Pythagore, $\|x+\lambda y\|^2 = \|x\|^2 + \lambda^2\|y\|^2$.

2. Si F et G sont orthogonaux, pour $u \in F$ et $v \in G$, le théorème de Pythagore donne

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 \geqslant \|u\|^2.$$

Réciproquement, supposons $\|u+v\| \geqslant \|u\|$ pour tous $u \in F$, $v \in G$. Soit $u \in F$ et $v \in G$; pour tout réel λ, le vecteur λv appartient au sous-espace G, donc $\|u+\lambda v\| \geqslant \|u\|$, et la question 1 donne $u \perp v$.

$$\boxed{F \text{ et } G \text{ orthogonaux} \iff \|u+v\| \geqslant \|u\| \quad \text{pour tous } u \in F,\ v \in G}$$

Piège. La question 2 ne se réduit pas à la question 1 appliquée au seul couple (u,v) : il faut tous les multiples λv, et c'est parce que G est un sous-espace qu'ils sont disponibles. Pour un couple isolé, l'inégalité n'entraîne rien : avec $u=(1,0)$ et $v=(1,1)$, on a $\|u+v\| = \sqrt{5} \geqslant 1 = \|u\|$, et pourtant $\langle u,v\rangle = 1$. Cette caractérisation par une distance minimale annonce le théorème 7 de meilleure approximation.

EXERCICE 14 ⋆

1. La famille (u, v, w) est libre, puisque son déterminant dans la base canonique vaut $1 \neq 0$ (chapitre 17). On applique le théorème 4. D'abord $\|u\| = \sqrt{2}$ et $u_1 = \dfrac{1}{\sqrt{2}}(1,0,1)$. Ensuite $\langle v, u_1\rangle = \dfrac{2}{\sqrt{2}} = \sqrt{2}$, et

$$v_2 = v - \sqrt{2}\,u_1 = (1,1,1) - (1,0,1) = (0,1,0), \qquad u_2 = (0,1,0).$$

Enfin $\langle w, u_1\rangle = -\dfrac{1}{\sqrt{2}}$ et $\langle w, u_2\rangle = -1$, d'où

$$v_3 = w + \frac{1}{\sqrt{2}}u_1 + u_2 = (-1,-1,0) + \Big(\frac{1}{2},0,\frac{1}{2}\Big) + (0,1,0) = \Big(-\frac{1}{2},0,\frac{1}{2}\Big), \qquad \|v_3\| = \frac{1}{\sqrt{2}}.$$

$$\boxed{u_1 = \frac{1}{\sqrt{2}}(1,0,1), \qquad u_2 = (0,1,0), \qquad u_3 = \frac{1}{\sqrt{2}}(-1,0,1)}$$

On contrôle que $\langle u_3, u\rangle = 0$, $\langle u_3, v\rangle = 0$ et $\langle u_3, w\rangle = \dfrac{1}{\sqrt{2}} > 0$ ✓, conformément au théorème 4.

2. On cherche d'abord une base de F en résolvant le système : la seconde équation donne $y = x+z$, la première $t = -x - z$, donc

$$F = \{(x, x+z, z, -x-z) \mid x, z \in \mathbb{R}\} = \mathrm{Vect}(f_1, f_2), \qquad f_1 = (1,1,0,-1), \quad f_2 = (0,1,1,-1).$$

On applique la méthode du cours, sans racine carrée : $w_1 = f_1$, de norme $\sqrt{3}$, puis, avec $\langle f_2, f_1\rangle = 2$,

$$w_2 = f_2 - \frac{2}{3}f_1 = \Big(-\frac{2}{3},\frac{1}{3},1,-\frac{1}{3}\Big) = \frac{1}{3}(-2,1,3,-1).$$

On contrôle $\langle(-2,1,3,-1), f_1\rangle = -2+1+0+1 = 0$ ✓, et que $(-2,1,3,-1)$ vérifie les deux équations de F. Il reste à normaliser, avec $\|(-2,1,3,-1)\| = \sqrt{15}$:

$$\boxed{\Big(\frac{1}{\sqrt{3}}(1,1,0,-1),\ \frac{1}{\sqrt{15}}(-2,1,3,-1)\Big) \text{ est une base orthonormée de } F}$$

Piège. La réponse dépend de la base de départ : partir de $(-1,0,1,0)$ et $(-1,-1,0,1)$ conduit à $\dfrac{1}{\sqrt{2}}(-1,0,1,0)$ et $\dfrac{1}{\sqrt{10}}(-1,-2,-1,2)$, qui conviennent tout autant. Un sous-espace de dimension au moins 2 possède une infinité de bases orthonormées ; c'est seulement à base de départ fixée que le théorème 4 donne l'unicité.

EXERCICE 15 ⋆

(a) L'application s'écrit $X^T S Y$ avec S diagonale de coefficients 1, 2, 1 : elle est bilinéaire et symétrique, et $\langle x, x\rangle = x_1^2 + 2x_2^2 + x_3^2$ est positif, nul seulement si $x = 0$. Pour ce produit, $\langle e_i, e_j\rangle = 0$ si $i \neq j$: la base canonique est déjà orthogonale, et l'algorithme se réduit à la normalisation, avec $\|e_1\| = \|e_3\| = 1$ et $\|e_2\| = \sqrt{2}$.

$$\boxed{u_1 = e_1 = (1,0,0), \qquad u_2 = \frac{e_2}{\sqrt{2}} = \Big(0,\frac{1}{\sqrt{2}},0\Big), \qquad u_3 = e_3 = (0,0,1)}$$

Le vecteur $(0,1,0)$, de longueur 1 pour le produit canonique, a pour norme $\sqrt{2}$ ici : être unitaire dépend du produit scalaire, comme le soulignait l'exemple qui suit le théorème 5.

(b) Le poids $t \mapsto t$ est continu et positif sur $[0\,,2]$, et ne s'annule qu'en 0 : par la proposition 2 (ii), la formule définit un produit scalaire sur $\mathbb{R}[X]$, donc sur $\mathbb{R}_2[X]$. Les produits utiles se calculent par primitives usuelles :

$$\langle X^i, X^j\rangle = \int_0^2 t^{i+j+1}\,\mathrm{d}t = \frac{2^{i+j+2}}{i+j+2},$$

d'où

$$\|1\|^2 = 2, \qquad \langle X, 1\rangle = \frac{8}{3}, \qquad \|X\|^2 = \langle X^2, 1\rangle = 4, \qquad \langle X^2, X\rangle = \frac{32}{5}.$$

On applique la méthode du cours sans racine carrée. D'abord $w_0 = 1$, puis

$$w_1 = X - \frac{\langle X, 1\rangle}{\|1\|^2} = X - \frac{4}{3}, \qquad \|w_1\|^2 = \|X\|^2 - \frac{8}{3}\langle X, 1\rangle + \frac{16}{9}\|1\|^2 = 4 - \frac{64}{9} + \frac{32}{9} = \frac{4}{9}.$$

Ensuite $\langle X^2, w_1\rangle = \frac{32}{5} - \frac{4}{3}\cdot 4 = \frac{16}{15}$ et $\frac{\langle X^2, w_1\rangle}{\|w_1\|^2} = \frac{16}{15}\cdot\frac{9}{4} = \frac{12}{5}$, d'où

$$w_2 = X^2 - \frac{4}{2} - \frac{12}{5}\Big(X - \frac{4}{3}\Big) = X^2 - \frac{12}{5}X + \frac{6}{5}, \qquad \|w_2\|^2 = \frac{8}{75}.$$

On contrôle $\langle w_2, 1\rangle = 4 - \frac{12}{5}\cdot\frac{8}{3} + \frac{6}{5}\cdot 2 = 4 - \frac{32}{5} + \frac{12}{5} = 0$ ✓. En divisant par les normes $\sqrt{2}$, $\frac{2}{3}$ et $\sqrt{\frac{8}{75}} = \frac{2\sqrt{6}}{15}$:

$$\boxed{u_0 = \frac{1}{\sqrt{2}}, \qquad u_1 = \frac{3X-4}{2}, \qquad u_2 = \frac{\sqrt{6}}{4}(5X^2 - 12X + 6)}$$

Le polynôme w_1 s'annule en $\frac{4}{3}$, au-delà du milieu de $[0\,,2]$: le poids t favorise la partie droite de l'intervalle, et le « centre de gravité » $\frac{\langle X, 1\rangle}{\|1\|^2}$ s'y déplace. Les racines de w_2, $\frac{6 \pm \sqrt{6}}{5} \approx 0{,}71$ et $1{,}69$, sont réelles, simples et dans $]0\,,2[$, comme l'annonce la proposition 7.

EXERCICE 16 ⋆⋆

1. L'application est symétrique, et linéaire par rapport à P, puisque $P \mapsto P(\cos\theta)$ est linéaire et que l'intégrale l'est. Pour $P \in \mathbb{R}[X]$, la fonction $\theta \mapsto P(\cos\theta)^2$ est continue et positive sur $[0\,,\pi]$, donc $\langle P, P\rangle \geqslant 0$. Si $\langle P, P\rangle = 0$, cette fonction est nulle (chapitre 18, théorème 4) : $P(\cos\theta) = 0$ pour tout $\theta \in [0\,,\pi]$. Comme cos réalise une bijection de $[0\,,\pi]$ sur $[-1\,,1]$, le polynôme P s'annule sur tout le segment $[-1\,,1]$: il a une infinité de racines, donc il est nul (chapitre 13, théorème 3).

$$\boxed{\langle P, Q\rangle = \int_0^\pi P(\cos\theta)\,Q(\cos\theta)\,\mathrm{d}\theta \ \text{ est un produit scalaire sur } \mathbb{R}[X]}$$

2. Par définition de T_k, on a $\langle T_k, T_\ell\rangle = \displaystyle\int_0^\pi \cos(k\theta)\cos(\ell\theta)\,\mathrm{d}\theta$. On linéarise le produit (chapitre 4) :

$$\cos(k\theta)\cos(\ell\theta) = \frac{1}{2}\big(\cos((k-\ell)\theta) + \cos((k+\ell)\theta)\big).$$

Pour un entier $m \neq 0$, $\displaystyle\int_0^\pi \cos(m\theta)\,\mathrm{d}\theta = \Big[\frac{\sin(m\theta)}{m}\Big]_0^\pi = 0$, et pour $m = 0$ l'intégrale vaut π. Si $k \neq \ell$, les deux entiers $k - \ell$ et $k + \ell$ sont non nuls ; si $k = \ell \geqslant 1$, seul $k - \ell$ est nul ; si $k = \ell = 0$, les deux le sont.

$$\boxed{\langle T_k, T_\ell\rangle = 0 \ (k \neq \ell), \qquad \|T_0\|^2 = \pi, \qquad \|T_k\|^2 = \frac{\pi}{2} \ (k \geqslant 1)}$$

La famille $(T_k)_{k\in\mathbb{N}}$ est orthogonale : c'est l'orthogonalité des fonctions $\cos(k\,\cdot)$ de l'exemple qui suit le théorème 3, transportée par le changement de variable $t = \cos\theta$.

3. Posons $u_0 = \frac{1}{\sqrt{\pi}}T_0$ et $u_k = \sqrt{\frac{2}{\pi}}\,T_k$ pour $k \in [\![1\,;N]\!]$. Par la question 2, la famille $(u_0, \dots, u_N)$ est orthonormée. Par l'exercice 53 de la feuille du chapitre 13, T_k est de degré k, de coefficient dominant 2^{k-1} pour $k \geqslant 1$; la famille $(u_0, \dots, u_k)$ est donc à degrés échelonnés, et elle engendre $\mathbb{R}_k[X]$, c'est-à-dire $\mathrm{Vect}(1, X, \dots, X^k)$ (chapitre 14, proposition 10). Enfin, pour $k \geqslant 1$, on écrit

$$X^k = 2^{1-k}T_k + R_k, \qquad R_k \in \mathbb{R}_{k-1}[X] = \mathrm{Vect}(u_0, \dots, u_{k-1}),$$

où R_k est orthogonal à u_k, d'où

$$\langle u_k, X^k\rangle = 2^{1-k}\sqrt{\frac{2}{\pi}}\,\|T_k\|^2 > 0, \qquad \text{et} \qquad \langle u_0, 1\rangle = \sqrt{\pi} > 0.$$

Les conditions d'unicité du théorème 4 sont remplies : la famille (u_k) est celle que fournit l'algorithme.

$$\boxed{u_0 = \frac{1}{\sqrt{\pi}}, \qquad u_k = \sqrt{\frac{2}{\pi}}\,T_k \quad (1 \leqslant k \leqslant N)}$$

Par exemple, $u_2 = \sqrt{\frac{2}{\pi}}(2X^2-1)$ et $u_3 = \sqrt{\frac{2}{\pi}}(4X^3-3X)$, ce que confirme le calcul direct de l'algorithme pour $N = 3$ ✓.

À retenir. Pour orthonormaliser, il n'est pas toujours nécessaire d'exécuter l'algorithme : il suffit de deviner une famille orthonormée et de vérifier les deux conditions d'unicité du théorème 4, l'égalité des sous-espaces engendrés et la positivité des $\langle u_k, e_k\rangle$. L'erreur la plus répandue est d'oublier la seconde : $-u_k$ conviendrait aussi pour la première. Les polynômes de Tchebychev sont ainsi les polynômes orthogonaux de ce produit scalaire, comme les polynômes de Legendre pour $\displaystyle\int_{-1}^{1} PQ$; on retrouve, par l'exercice 53 de la feuille du chapitre 13, que T_k a ses k racines simples dans $]-1\,,1[$, en accord avec la proposition 7.

EXERCICE 17 ⋆⋆⋆

1. Les ensembles I_+ et I_- sont disjoints. Comme $\displaystyle\sum_{i=1}^{p-1} \lambda_i x_i = 0$ et que les termes d'indices hors de $I_+ \cup I_-$ sont nuls,

$$u = \sum_{i\in I_+} \lambda_i x_i = -\sum_{j\in I_-} \lambda_j x_j = \sum_{j\in I_-} |\lambda_j|\, x_j.$$

On calcule $\|u\|^2$ comme le produit scalaire de ces deux écritures, en développant par bilinéarité :

$$\|u\|^2 = \Big\langle \sum_{i\in I_+} \lambda_i x_i, \sum_{j\in I_-} |\lambda_j|\, x_j \Big\rangle = \sum_{i\in I_+}\sum_{j\in I_-} \lambda_i\,|\lambda_j|\,\langle x_i, x_j\rangle.$$

Dans chaque terme, $\lambda_i > 0$, $|\lambda_j| > 0$ et $\langle x_i, x_j\rangle < 0$, puisque $i \neq j$. Une somme vide étant nulle, on obtient dans tous les cas

$$\boxed{\|u\|^2 \leqslant 0}$$

Le point clé est d'avoir regroupé les coefficients par signe : développer directement $\|\sum \lambda_i x_i\|^2$ ferait apparaître les $\|x_i\|^2$, positifs, et l'on ne pourrait rien conclure.

2. Par la question 1, $u = 0$, donc $\langle u, x_p\rangle = 0$. Or

$$\langle u, x_p\rangle = \sum_{i\in I_+} \lambda_i \langle x_i, x_p\rangle,$$

où chaque terme est strictement négatif, puisque $\lambda_i > 0$ et $i \neq p$. Si I_+ n'était pas vide, cette somme serait strictement négative : donc $I_+ = \varnothing$. La famille $(-\lambda_1, \dots, -\lambda_{p-1})$ vérifie la même relation $\displaystyle\sum_i (-\lambda_i) x_i = 0$, et son ensemble d'indices positifs est I_- : le même raisonnement donne $I_- = \varnothing$.

$$\boxed{\lambda_1 = \cdots = \lambda_{p-1} = 0 : \text{ la famille } (x_1, \dots, x_{p-1}) \text{ est libre}}$$

C'est ici que sert le vecteur x_p, qui n'intervient pas dans la relation : il « voit » u sous un angle obtus.

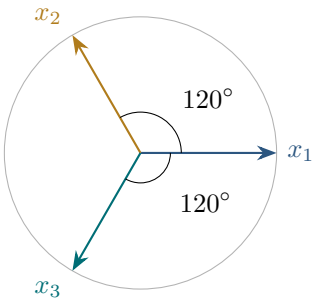
3. Une famille libre de l'espace E, de dimension n, a au plus n éléments (chapitre 15, théorème 4). Donc $p - 1 \leqslant n$.

$$\boxed{p \leqslant n+1}$$

4. Trois vecteurs unitaires de $\mathbb{R}^2$ faisant deux à deux un angle de $\dfrac{2\pi}{3}$ conviennent :

$$\boxed{x_1 = (1,0), \quad x_2 = \Big(-\frac{1}{2}, \frac{\sqrt{3}}{2}\Big), \quad x_3 = \Big(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\Big), \qquad \langle x_i, x_j\rangle = -\frac{1}{2}\ (i \neq j)}$$

On peut aussi prendre des coordonnées entières : $(1,0)$, $(-1,2)$ et $(-1,-2)$ ont pour produits scalaires -1, -1 et -3.



Les trois vecteurs ont une somme nulle ; un quatrième vecteur ne peut pas faire un angle obtus avec chacun d'eux, ce que la question 3 confirme pour $n = 2$.
À retenir. La borne $n+1$ est atteinte dans tout espace euclidien : les sommets d'un simplexe régulier centré en l'origine (triangle équilatéral dans le plan, tétraèdre régulier dans l'espace) en fournissent un exemple, qu'étudiera l'exercice 51. En revanche, si l'on demande seulement $\langle x_i, x_j\rangle \leqslant 0$, la borne $n+1$ tombe : les $2n$ vecteurs $\pm e_1$, …, $\pm e_n$ d'une base orthonormée vérifient cette condition large. L'inégalité stricte est essentielle dans la question 2.

EXERCICE 18 ⋆⋆⋆

1. Pour $k \in \mathbb{N}$, on a $(1-X^2)(X^k)' = kX^{k-1} - kX^{k+1}$, et, en dérivant,

$$\boxed{\varphi(X^k) = k(k-1)X^{k-2} - k(k+1)X^k}$$

avec la convention que le premier terme est nul pour $k \in \{0,1\}$. L'application φ est linéaire, comme composée de dérivations et d'une multiplication par $1-X^2$. Pour $P \in \mathbb{R}_n[X]$, combinaison de 1, …, X^n, le polynôme $\varphi(P)$ est combinaison des $\varphi(X^k)$, $k \leqslant n$, qui sont de degré au plus $k \leqslant n$.

$$\boxed{\varphi(\mathbb{R}_n[X]) \subset \mathbb{R}_n[X] : \ \varphi \text{ induit un endomorphisme de } \mathbb{R}_n[X]}$$

2. Notons ψ l'endomorphisme de $\mathbb{R}_n[X]$ induit par $\varphi + n(n+1)\mathrm{id}$, et M sa matrice dans la base $(1, X, \dots, X^n)$. Par la question 1,

$$\psi(X^k) = k(k-1)X^{k-2} + d_k X^k, \qquad d_k = n(n+1) - k(k+1) = (n-k)(n+k+1).$$

La matrice M est triangulaire supérieure, de coefficients diagonaux d_0, …, d_n, et $d_k = 0$ si et seulement si $k = n$. Pour $n = 3$, par exemple,

$$M = \begin{pmatrix} 12 & 0 & 2 & 0 \\ 0 & 10 & 0 & 6 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Le noyau est une droite. La dernière ligne de M est nulle, donc $\mathrm{rg}\psi \leqslant n$. Les n premières colonnes sont celles de $\psi(1)$, …, $\psi(X^{n-1})$: leur bloc supérieur est triangulaire de diagonale d_0, …, d_{n-1} non nulle, donc inversible (chapitre 16, proposition 12), et elles forment une famille libre ; ainsi $\mathrm{rg}\psi = n$, et le théorème du rang (chapitre 15, théorème 9) donne $\dim \mathrm{Ker}\,\psi = 1$.

Degré des éléments du noyau. L'endomorphisme ψ envoie $\mathbb{R}_{n-1}[X]$ dans lui-même, avec une matrice triangulaire de diagonale $d_0, \ldots, d_{n-1}$ non nulle : sa restriction à $\mathbb{R}_{n-1}[X]$ est injective. Un élément non nul de $\operatorname{Ker}\psi$ n'appartient donc pas à $\mathbb{R}_{n-1}[X]$: il est de degré exactement n, et l'on peut le diviser par son coefficient dominant. Deux polynômes unitaires du noyau, droite vectorielle, sont proportionnels, donc égaux.

$$\boxed{\text{Il existe un unique } P_n \text{ unitaire, de degré au plus } n, \text{ tel que } \varphi(P_n) + n(n+1)P_n = 0;\ \deg P_n = n}$$

On trouve $P_0 = 1$, $P_1 = X$, $P_2 = X^2 - \dfrac{1}{3}$, $P_3 = X^3 - \dfrac{3}{5}X$ et $P_4 = X^4 - \dfrac{6}{7}X^2 + \dfrac{3}{35}$. Pour $n = 3$, le système $MC = 0$ avec $c_3 = 1$ donne $c_2 = 0$, $10c_1 + 6 = 0$ et $12c_0 + 2c_2 = 0$, soit $c_1 = -\dfrac{3}{5}$ et $c_0 = 0$ ✓.

3. On intègre par parties (chapitre 6), en dérivant Q et en intégrant $\big((1-t^2)P'\big)'$; le crochet est nul, puisque $1 - t^2$ s'annule en ± 1 :

$$\langle \varphi(P), Q\rangle = \Big[(1-t^2)P'(t)Q(t)\Big]_{-1}^{1} - \int_{-1}^{1}(1-t^2)P'(t)Q'(t)\,\mathrm{d}t = -\int_{-1}^{1}(1-t^2)P'(t)Q'(t)\,\mathrm{d}t.$$

La dernière expression est symétrique en P et Q.

$$\boxed{\langle \varphi(P), Q\rangle = \langle P, \varphi(Q)\rangle = -\int_{-1}^{1}(1-t^2)P'(t)Q'(t)\,\mathrm{d}t}$$

Soit $m \neq n$. Par la question 2, $\varphi(P_n) = -n(n+1)P_n$ et $\varphi(P_m) = -m(m+1)P_m$, donc

$$-n(n+1)\langle P_n, P_m\rangle = \langle \varphi(P_n), P_m\rangle = \langle P_n, \varphi(P_m)\rangle = -m(m+1)\langle P_n, P_m\rangle.$$

Comme $x \mapsto x(x+1)$ est strictement croissante sur $\mathbb{N}$, $n(n+1) \neq m(m+1)$, et

$$\boxed{\langle P_n, P_m\rangle = 0 \quad (n \neq m) :\ \text{la famille } (P_n)_{n\in\mathbb{N}} \text{ est orthogonale}}$$

4. Soit $n \geqslant 1$. Les polynômes $P_0, \ldots, P_{n-1}$ sont de degrés $0, \ldots, n-1$: ils forment une base de $\mathbb{R}_{n-1}[X]$ (chapitre 14, proposition 10), et P_n, orthogonal à chacun d'eux, est orthogonal à $\mathbb{R}_{n-1}[X]$ (proposition 5 (ii)). Par la proposition 7 (i), Q_n est de degré n et orthogonal à $\mathbb{R}_{n-1}[X]$; il est même unitaire, puisque la méthode construit $Q_n = X^n - \displaystyle\sum_{j<n} \frac{\langle X^n, Q_j\rangle}{\|Q_j\|^2} Q_j$. Le polynôme $R = P_n - Q_n$ est alors de degré au plus $n-1$, les termes en X^n se compensant, et orthogonal à $\mathbb{R}_{n-1}[X]$, donc à lui-même : $\|R\|^2 = 0$ et $R = 0$.

$$\boxed{P_n = Q_n}$$

En particulier P_n est un multiple non nul de Q_n, avec le facteur 1. Par la proposition 7 (ii), les n racines de P_n sont réelles, simples et situées dans $]-1\,,1[$; comme $\deg P_n = n$, il n'en a pas d'autre, et 1 n'en est pas une (pour $n = 0$, $P_0 = 1$ ne s'annule pas non plus).

$$\boxed{P_n(1) \neq 0}$$

Enfin, le polynôme $L_n = \big[(X^2-1)^n\big]^{(n)}$ est de degré n, de coefficient dominant $\dfrac{(2n)!}{n!}$, puisque $(X^2-1)^n$ est unitaire de degré $2n$, et il est orthogonal à $\mathbb{R}_{n-1}[X]$ par la question 2 de l'exercice 59 de la feuille du chapitre 13. Le même argument, appliqué à $L_n - \dfrac{(2n)!}{n!}P_n$, donne

$$\boxed{\big[(X^2-1)^n\big]^{(n)} = \frac{(2n)!}{n!}\,P_n}$$

Pour $n = 2$, $\big[(X^2-1)^2\big]'' = 12X^2 - 4 = 12(X^2 - \dfrac{1}{3})$ ✓.

À retenir. Un polynôme de degré n orthogonal à $\mathbb{R}_{n-1}[X]$ est déterminé à un facteur près : c'est l'argument de la question 4, utilisé trois fois. Les polynômes de Legendre du cours, normalisés par $P_n(1) = 1$, sont les $\dfrac{P_n}{P_n(1)}$; la question 4 assure que cette normalisation est possible. La question 3 est une propriété de symétrie de φ par rapport au produit scalaire : c'est elle qui rend orthogonales deux solutions associées à des constantes $n(n+1)$ différentes, idée que l'on retrouvera en seconde année.

Bases orthonormées

EXERCICE 19 ⋆

Par hypothèse, $\|A\|^2 = \mathrm{tr}(A^TA) = \mathrm{tr}(I_n) = n$, c'est-à-dire $\displaystyle\sum_{i,j} a_{ij}^2 = n$.

1. On applique l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^{n^2}$ (proposition 4 (i)) aux familles des $|a_{ij}|$ et des coefficients égaux à 1 :

$$\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}| \cdot 1 \leqslant \Big(\sum_{i,j} a_{ij}^2\Big)^{1/2}\Big(\sum_{i,j} 1\Big)^{1/2} = \sqrt{n}\cdot\sqrt{n^2}.$$

$$\boxed{\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}| \leqslant n\sqrt{n}}$$

Il y a égalité si et seulement si tous les $|a_{ij}|$ sont égaux, donc égaux à $\dfrac{1}{\sqrt{n}}$. C'est possible pour $n = 2$, avec $A = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, mais pas pour $n = 3$: deux colonnes à coefficients $\pm\dfrac{1}{\sqrt{3}}$ ont un produit scalaire égal à $\dfrac{\pm 1 \pm 1 \pm 1}{3}$, dont le numérateur est impair, donc non nul.

2. La colonne AU a pour coefficients les sommes des lignes de A, et $U^T(AU)$ est la somme de ces sommes : $\displaystyle\sum_{i,j} a_{ij} = \langle U, AU\rangle$. Par ailleurs, $\|AU\|^2 = U^TA^TAU = U^TU = n$. L'inégalité de Cauchy-Schwarz dans $\mathcal{M}_{n,1}(\mathbb{R})$ donne

$$\boxed{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = \langle U, AU\rangle \leqslant \|U\|\,\|AU\| = \sqrt{n}\cdot\sqrt{n} = n}$$

avec égalité si et seulement si $AU = U$, comme pour $A = I_n$: par le cas d'égalité, AU est un multiple positif de U, de même norme.

À retenir. Une matrice telle que $A^TA = I_n$ conserve la norme des colonnes : $\|AX\| = \|X\|$. C'est ce qui rend la question 2 bien plus précise que la question 1, dont elle ne se déduit pas : la majoration $n\sqrt{n}$ de la question 1 ne tient pas compte des signes, et elle est atteinte pour $n = 2$, alors que la somme des coefficients ne dépasse jamais n.

EXERCICE 20 ⋆⋆

La famille est orthonormée. Soit $j \in [\![1\,;n]\!]$. L'hypothèse, appliquée à $x = e_j$, donne, puisque e_j est unitaire,

$$1 = \|e_j\|^2 = \sum_{i=1}^{n} \langle e_j, e_i\rangle^2 = 1 + \sum_{i\neq j} \langle e_j, e_i\rangle^2.$$

Une somme de carrés nulle a tous ses termes nuls : $\langle e_j, e_i\rangle = 0$ pour tout $i \neq j$. La famille est orthonormée, donc libre (proposition 6).

La famille est génératrice. Soit $x \in E$; on pose

$$y = x - \sum_{i=1}^{n} \langle x, e_i\rangle e_i.$$

Par orthonormalité, $\langle y, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle = 0$ pour tout j, donc y est orthogonal à chacun des vecteurs $\langle x, e_i\rangle e_i$. Le théorème de Pythagore (théorème 3 (ii)), appliqué à la famille orthogonale $(y, \langle x, e_1\rangle e_1, \ldots, \langle x, e_n\rangle e_n)$, donne

$$\|x\|^2 = \|y\|^2 + \sum_{i=1}^{n} \langle x, e_i\rangle^2 = \|y\|^2 + \|x\|^2,$$

donc $\|y\| = 0$, $y = 0$, et $x = \sum_i \langle x, e_i\rangle e_i \in \mathrm{Vect}(e_1, \ldots, e_n)$.

$$\boxed{(e_1, \ldots, e_n) \text{ est une base orthonormée de } E, \text{ et } \dim E = n}$$

Piège. L'hypothèse « unitaires » est indispensable. Dans $\mathbb{R}^2$, les trois vecteurs $\sqrt{\frac{2}{3}}\,x_1$, $\sqrt{\frac{2}{3}}\,x_2$, $\sqrt{\frac{2}{3}}\,x_3$, où x_1, x_2, x_3 sont les vecteurs à 120° de l'exercice 17, vérifient $\sum_i \langle x, e_i\rangle^2 = \|x\|^2$ pour tout x, sans être orthogonaux ni former une base ; ils ont pour norme $\sqrt{\frac{2}{3}}$. Plus simplement, $(e_1, \frac{e_2}{\sqrt{2}}, \frac{e_2}{\sqrt{2}})$ convient aussi. La réciproque est la proposition 8 : dans une base orthonormée, $\|x\|^2 = \sum_i \langle x, e_i\rangle^2$. L'exercice caractérise donc les bases orthonormées parmi les familles de vecteurs unitaires par cette seule identité, et l'inégalité de Bessel (proposition 10 (iv)) en éclairera la signification : pour une famille orthonormée quelconque, on a seulement $\sum_i \langle x, e_i\rangle^2 \leqslant \|x\|^2$.

EXERCICE 21 ⋆⋆

1. Une famille orthonormée est libre, quel que soit le produit scalaire (proposition 6) : la liberté de $\mathcal{F}$ est donc nécessaire. Montrons qu'elle est suffisante. Si $\mathcal{F}$ est libre, le théorème de la base incomplète (chapitre 15, théorème 2) la complète en une base $(f_1, \ldots, f_n)$ de E. Pour x, $y \in E$ de coordonnées (x_k) et (y_k) dans cette base, on pose

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

Les coordonnées dépendent linéairement du vecteur (chapitre 14, théorème 2) : cette application est bilinéaire, et elle est symétrique. On a $\langle x, x\rangle = \sum x_k^2 \geqslant 0$, avec égalité seulement si toutes les coordonnées sont nulles, c'est-à-dire si $x = 0$: c'est un produit scalaire, et $\langle f_i, f_j\rangle = \delta_{ij}$ puisque les coordonnées de f_i sont nulles sauf la i-ième, égale à 1.

$$\boxed{\text{il existe un produit scalaire rendant } \mathcal{F} \text{ orthonormée} \iff \mathcal{F} \text{ est libre}}$$

On transporte ainsi le produit canonique de $\mathbb{R}^n$ par l'isomorphisme « coordonnées » : c'est la remarque qui suit la proposition 8, lue à l'envers.

2. La famille $\mathcal{F}$ est libre, et c'est une base de $\mathbb{R}^2$. Un produit scalaire pour lequel elle est orthonormée est unique : par bilinéarité, si $x = af_1 + bf_2$ et $y = a'f_1 + b'f_2$, on a nécessairement $\langle x, y\rangle = aa' + bb'$. On cherche donc les coordonnées de $x = (x_1, x_2)$ dans $\mathcal{F}$: l'égalité $(x_1, x_2) = a(1,1) + b(0,2)$ donne $a = x_1$ et $a + 2b = x_2$.

$$\boxed{x = x_1 f_1 + \frac{x_2 - x_1}{2} f_2, \qquad \langle x, y\rangle = x_1 y_1 + \frac{(x_2 - x_1)(y_2 - y_1)}{4}}$$

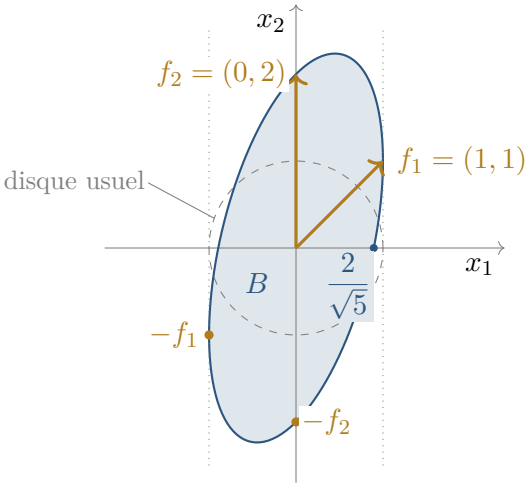
On contrôle que $\langle f_1, f_1\rangle = 1 + 0 = 1$, que $\langle f_2, f_2\rangle = 0 + \dfrac{4}{4} = 1$ et que $\langle f_1, f_2\rangle = 0 + 0 = 0$. ✓ La condition $\|x\|^2 \leqslant 1$ s'écrit $4x_1^2 + (x_2 - x_1)^2 \leqslant 4$, soit, après développement :

$$\boxed{B = \{(x_1, x_2) \in \mathbb{R}^2 \mid 5x_1^2 - 2x_1x_2 + x_2^2 \leqslant 4\}}$$

Pour la tracer, on remarque que B est l'image du disque unité usuel par l'application

$$(a, b) \longmapsto af_1 + bf_2 = (a, a + 2b)\,;$$

son bord est la courbe paramétrée $\theta \mapsto (\cos\theta, \cos\theta + 2\sin\theta)$, une ellipse. Elle passe par $\pm f_1 = \pm(1,1)$ et $\pm f_2 = \pm(0,2)$, coupe l'axe des abscisses en $\pm\dfrac{2}{\sqrt{5}}$, et x_2 y varie entre $-\sqrt{5}$ et $\sqrt{5}$.



La boule B, en bleu, est une ellipse : pour ce produit scalaire, f_1 et f_2 sont unitaires et orthogonaux, alors qu'ils ne le sont pas pour le produit canonique, dont la boule unité est le disque en pointillés. L'ellipse est tangente aux droites $x_1 = \pm 1$ en $\pm f_1$, puisque $x_1 = \cos\theta$ y est maximal.
À retenir. Être orthonormée n'est pas une propriété d'une famille, mais d'une famille *et* d'un produit scalaire : toute base est orthonormée pour un produit bien choisi, comme le disait le second exemple qui suit le théorème 5. Changer de produit scalaire, c'est déformer les cercles en ellipses.

EXERCICE 22 ⋆⋆

1. Pour $a \in E$, l'application $x \mapsto \langle a, x\rangle$ est linéaire, par linéarité du produit scalaire par rapport à la seconde variable : $\Phi(a)$ appartient bien à $\mathcal{L}(E, \mathbb{R})$. La linéarité par rapport à la première variable donne $\Phi(\lambda a + \mu b) = \lambda\Phi(a) + \mu\Phi(b)$: l'application Φ est linéaire. Elle est injective : si $\Phi(a) = 0$, alors en particulier $\langle a, a\rangle = \Phi(a)(a) = 0$, donc $a = 0$ par le caractère défini. Enfin, $\dim \mathcal{L}(E, \mathbb{R}) = \dim E \times 1 = \dim E$ (chapitre 15, proposition 8), et une application linéaire injective entre deux espaces de même dimension finie est bijective (chapitre 15, théorème 10).

$$\boxed{\Phi : a \longmapsto \langle a, \cdot\rangle \text{ est un isomorphisme de } E \text{ sur } \mathcal{L}(E, \mathbb{R})}$$

2. L'existence de u est la surjectivité de Φ, et son unicité est l'injectivité : $u = \Phi^{-1}(\varphi)$.

$$\boxed{\forall\, \varphi \in \mathcal{L}(E, \mathbb{R}),\ \exists!\, u \in E,\ \forall\, x \in E, \quad \varphi(x) = \langle u, x\rangle}$$

Piège. L'hypothèse « E de dimension finie » sert à l'argument de dimension. En dimension infinie, Φ reste injective mais n'est plus surjective en général : l'exercice 40 en donne un exemple sur $\mathbb{R}[X]$.

3. *(a)* Pour $M \in \mathcal{M}_n(\mathbb{R})$, on a $\langle I_n, M\rangle = \operatorname{tr}({I_n}^T M) = \operatorname{tr} M$. Par l'unicité de la question 2 :

$$\boxed{u = I_n}$$

(b) Le déterminant est linéaire par rapport à sa dernière colonne, donc φ est une forme linéaire. On le développe suivant cette colonne (chapitre 17, théorème 7) :

$$\det(a,b,x) = \begin{vmatrix} a_1 & b_1 & x_1 \\ a_2 & b_2 & x_2 \\ a_3 & b_3 & x_3 \end{vmatrix} = (a_2b_3 - a_3b_2)x_1 + (a_3b_1 - a_1b_3)x_2 + (a_1b_2 - a_2b_1)x_3.$$

C'est le produit scalaire canonique de x avec le vecteur suivant, qui est donc le vecteur u cherché :

$$\boxed{u = (a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1)}$$

On contrôle que $\langle u, a\rangle = \det(a,b,a) = 0$ et $\langle u, b\rangle = 0$, puisqu'un déterminant à deux colonnes égales est nul : u est orthogonal à a et à b. ✓ De plus, $u = 0$ si et seulement si $\det(a,b,x) = 0$ pour tout x, c'est-à-dire si et seulement si (a,b) est liée : si elle est libre, on la complète en une base (a,b,x), de déterminant non nul (chapitre 17, théorème 4). Ce vecteur sera étudié pour lui-même en seconde année.

4. Par la proposition 2 (ii), appliquée sur $[0\,,\pi]$ avec le poids $w = \sin$, continu, positif et nul seulement en 0 et en π, la formule $\langle P, Q\rangle = \displaystyle\int_0^\pi P(t)Q(t)\sin t\,\mathrm{d}t$ définit un produit scalaire sur $\mathbb{R}[X]$, donc sur le sous-espace $\mathbb{R}_n[X]$, qui est de dimension $n+1$: $\mathbb{R}_n[X]$ est euclidien. La fonction $t \mapsto \dfrac{1}{1+\ln^2 t}$ est continue sur $[2\,,7]$, et la linéarité de l'intégrale (chapitre 18, théorème 3) montre que

$$\varphi : Q \longmapsto \int_2^7 \frac{Q(t)}{1+\ln^2 t}\,\mathrm{d}t$$

est une forme linéaire sur $\mathbb{R}_n[X]$. La question 2 fournit un unique $P \in \mathbb{R}_n[X]$ tel que $\varphi(Q) = \langle P, Q\rangle$ pour tout $Q \in \mathbb{R}_n[X]$: c'est l'égalité demandée.

$$\boxed{\exists!\, P \in \mathbb{R}_n[X],\ \forall\, Q \in \mathbb{R}_n[X],\quad \int_2^7 \frac{Q(t)}{1+\ln^2 t}\,\mathrm{d}t = \int_0^\pi P(t)Q(t)\sin t\,\mathrm{d}t}$$

Piège. L'erreur la plus répandue est de chercher à calculer P, ou à « identifier » les deux intégrales : elles portent sur des intervalles différents, et aucune manipulation de l'intégrale n'y mène. L'existence est purement géométrique. Remarquons aussi que l'unicité tient à la condition $P \in \mathbb{R}_n[X]$: dans $\mathbb{R}[X]$, on peut ajouter à P n'importe quel polynôme orthogonal à $\mathbb{R}_n[X]$, par exemple le polynôme Q_{n+1} de degré $n+1$ fourni par la proposition 7 pour ce poids.

EXERCICE 23 ⋆⋆

1. On raisonne par l'absurde. Si $(f_1, \ldots, f_n)$ n'est pas une base, cette famille de n vecteurs de l'espace E de dimension n est liée (chapitre 15, théorème 4) : il existe des réels λ_1, …, λ_n non tous nuls tels que $\displaystyle\sum_k \lambda_k f_k = 0$, et, quitte à les diviser tous par $\sqrt{\displaystyle\sum_k \lambda_k^2}$, on peut supposer $\displaystyle\sum_k \lambda_k^2 = 1$. Alors

$$\sum_{k=1}^n \lambda_k e_k = \sum_{k=1}^n \lambda_k e_k - \sum_{k=1}^n \lambda_k f_k = \sum_{k=1}^n \lambda_k (e_k - f_k).$$

La base (e_k) étant orthonormée, la proposition 8 donne au membre de gauche la norme $\sqrt{\displaystyle\sum_k \lambda_k^2} = 1$. Au membre de droite, on applique l'inégalité triangulaire (théorème 2, étendue à n termes par récurrence),

puis l'hypothèse. Comme l'un au moins des λ_k n'est pas nul, et que pour lui $|\lambda_k|\,\|e_k - f_k\| < \dfrac{|\lambda_k|}{\sqrt{n}}$, l'inégalité est stricte :

$$1 = \Big\| \sum_{k=1}^{n} \lambda_k (e_k - f_k) \Big\| \leqslant \sum_{k=1}^{n} |\lambda_k|\,\|e_k - f_k\| < \frac{1}{\sqrt{n}} \sum_{k=1}^{n} |\lambda_k|.$$

Enfin, l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$ (proposition 4 (i)), appliquée à $(|\lambda_k|)$ et à $(1, \dots, 1)$, donne $\displaystyle\sum_k |\lambda_k| \leqslant \sqrt{\sum_k \lambda_k^2}\,\sqrt{n} = \sqrt{n}$. On obtient $1 < 1$, ce qui est absurde.

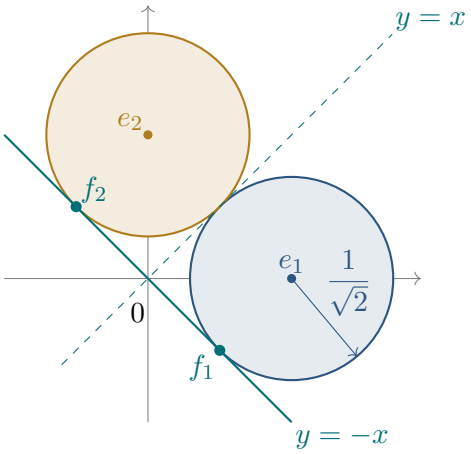
$$\boxed{\|f_k - e_k\| < \frac{1}{\sqrt{n}} \text{ pour tout } k \implies (f_1, \dots, f_n) \text{ est une base de } E}$$

2. Non. On note $s = e_1 + \cdots + e_n$, de norme $\sqrt{n}$ par la proposition 8, et l'on pose $f_k = e_k - \dfrac{1}{n}s$. Alors

$$\|f_k - e_k\| = \frac{\|s\|}{n} = \frac{1}{\sqrt{n}} \qquad \text{et} \qquad \sum_{k=1}^{n} f_k = s - n \cdot \frac{s}{n} = 0.$$

$$\boxed{\text{avec } f_k = e_k - \frac{e_1 + \cdots + e_n}{n},\ \|f_k - e_k\| = \frac{1}{\sqrt{n}} \text{ et } (f_1, \dots, f_n) \text{ est liée}}$$

Pour $n = 1$, c'est $f_1 = 0$; pour $n = 2$, dans $\mathbb{R}^2$ muni de sa base canonique, $f_1 = (\frac{1}{2}, -\frac{1}{2})$ et $f_2 = (-\frac{1}{2}, \frac{1}{2}) = -f_1$.



Si f_1 est dans le disque ouvert de centre e_1 et f_2 dans celui de centre e_2, de rayon $\dfrac{1}{\sqrt{2}}$, aucune droite passant par 0 ne rencontre les deux disques : la distance de e_1 et celle de e_2 à une droite d'angle θ valent $|\sin\theta|$ et $|\cos\theta|$, qui ne peuvent être toutes deux strictement inférieures à $\dfrac{1}{\sqrt{2}}$. Les droites $y = \pm x$ sont tangentes aux deux cercles, et le contre-exemple place f_1 et f_2 aux points de contact avec $y = -x$.

À retenir. Une famille assez proche d'une base orthonormée est encore une base : la liberté est une propriété « stable par petite perturbation ». La borne $\dfrac{1}{\sqrt{n}}$ ne peut pas être élargie, et elle se dégrade quand la dimension augmente.

EXERCICE 24 ⋆⋆

Soit $\mathcal{E} = (e_1, \dots, e_n)$ une base dans laquelle la matrice T de u est triangulaire supérieure. Pour $k \in [\![1\,;n]\!]$, on note $F_k = \text{Vect}(e_1, \dots, e_k)$. Dire que la matrice est triangulaire supérieure, c'est dire que la k-ième colonne, formée des coordonnées de $u(e_k)$, n'a de coefficients non nuls que sur ses k premières lignes : $u(e_k) \in F_k$ pour tout k. Comme $F_j \subset F_k$ pour $j \leqslant k$, on en déduit que $u(e_j) \in F_k$ pour tout $j \leqslant k$, donc, par linéarité, que $u(F_k) \subset F_k$: chaque F_k est stable par u.

On applique l'algorithme de Gram-Schmidt (théorème 4) à la base $\mathcal{E}$: il fournit une base orthonormée $(\varepsilon_1, \dots, \varepsilon_n)$ telle que, pour tout k, $F_k = \text{Vect}(\varepsilon_1, \dots, \varepsilon_k)$. Alors

$$u(\varepsilon_k) \in u(F_k) \subset F_k = \text{Vect}(\varepsilon_1, \dots, \varepsilon_k) \ :$$

les coordonnées de $u(\varepsilon_k)$ dans la base (ε_i) sont nulles au-delà de la k-ième.

la matrice de u dans la base orthonormée $(\varepsilon_1, \dots, \varepsilon_n)$ est triangulaire supérieure

Les coefficients diagonaux sont d'ailleurs les mêmes dans les deux bases. La matrice de passage P de $\mathcal{E}$ à (ε_k) est triangulaire supérieure, puisque $\varepsilon_k \in F_k$; la nouvelle matrice est $P^{-1}TP$ (chapitre 16, théorème 4), et le coefficient diagonal d'un produit de matrices triangulaires supérieures est le produit des coefficients diagonaux (chapitre 16, propositions 6 et 12) : $\dfrac{1}{p_{kk}} \cdot t_{kk} \cdot p_{kk} = t_{kk}$. Sur un exemple de $\mathbb{R}^3$, un calcul symbolique confirme les deux propriétés. ✓

À retenir. L'algorithme de Gram-Schmidt conserve les sous-espaces $\text{Vect}(e_1, \dots, e_k)$: toute propriété qui ne s'exprime qu'à travers ces sous-espaces, comme la stabilité par u, passe de la base de départ à la base orthonormée. L'exercice 26 en tirera la décomposition $A = OT$.

EXERCICE 25 ⋆⋆⋆

1. Soit i_0 un indice tel que $\lambda_{i_0} \neq 0$. Le vecteur x étant non nul, la famille (x) est libre, et le théorème de la base incomplète (chapitre 15, théorème 2) la complète en une base $(x, h_2, \dots, h_n)$ de E. On numérote b_i, pour $i \neq i_0$, les vecteurs h_2, …, h_n, dans cet ordre, et l'on pose

$$b_{i_0} = \frac{1}{\lambda_{i_0}}\Big(x - \sum_{i \neq i_0} \lambda_i b_i\Big), \qquad \text{de sorte que} \qquad x = \sum_{i=1}^{n} \lambda_i b_i.$$

La famille $(b_1, \dots, b_n)$ engendre un sous-espace qui contient les h_j et $x = \displaystyle\sum_i \lambda_i b_i$, donc la base $(x, h_2, \dots, h_n)$: elle est génératrice, et, formée de n vecteurs, c'est une base (chapitre 15, théorème 4).

il existe une base $(b_1, \dots, b_n)$ de E dans laquelle x a pour coordonnées $(\lambda_1, \dots, \lambda_n)$

L'hypothèse $(\lambda_i) \neq 0$ est indispensable, puisque $x \neq 0$; c'est la seule condition.

2. *Condition nécessaire.* Si x a pour coordonnées (λ_i) dans une base orthonormée, la proposition 8 donne $\|x\|^2 = \displaystyle\sum_i \lambda_i^2$.

Condition suffisante. Supposons que $\displaystyle\sum_i \lambda_i^2 = \|x\|^2$, et notons $r = \|x\|$, qui est strictement positif ; on note encore $\lambda = (\lambda_1, \dots, \lambda_n)$, vecteur de $\mathbb{R}^n$ de norme canonique r. Par le théorème de la base orthonormée incomplète (théorème 5 (ii)), on dispose :

— d'une base orthonormée $(f_1, \dots, f_n)$ de E telle que $f_1 = \dfrac{x}{r}$;

— d'une base orthonormée $(g_1, \dots, g_n)$ de $\mathbb{R}^n$, muni du produit canonique, telle que $g_1 = \dfrac{\lambda}{r}$.

Soit $G \in \mathcal{M}_n(\mathbb{R})$ la matrice dont les colonnes sont les g_k. Sa matrice de Gram est $G^TG = I_n$ (proposition 9 (i)) : G est inversible d'inverse G^T (chapitre 16, théorème 3), donc on a aussi $GG^T = I_n$, et les *lignes* de G sont orthonormées. On pose alors

$$e_i = \sum_{k=1}^{n} G_{ik}\, f_k \qquad (1 \leqslant i \leqslant n),$$

vecteur dont les coordonnées dans la base orthonormée (f_k) forment la ligne i de G. Par la proposition 8, $\langle e_i, e_j\rangle = \sum_k G_{ik}G_{jk} = (GG^T)_{ij} = \delta_{ij}$: la famille $(e_1, \dots, e_n)$ est orthonormée, donc c'est une base orthonormée de E. Enfin, la coordonnée d'indice i de $x = rf_1$ dans cette base vaut

$$\langle x, e_i\rangle = r\langle f_1, e_i\rangle = rG_{i1} = r \times \frac{\lambda_i}{r} = \lambda_i,$$

puisque la première colonne de G est $g_1 = \dfrac{\lambda}{r}$.

$$\boxed{\text{il existe une base orthonormée dans laquelle } x \text{ a pour coordonnées } (\lambda_i) \iff \sum_{i=1}^{n} \lambda_i^2 = \|x\|^2}$$

Sur des données tirées au hasard, pour n de 2 à 5, un calcul numérique vérifie que la famille construite est orthonormée et donne bien les coordonnées (λ_i). ✓
Piège. L'erreur la plus répandue est de croire que $G^TG = I_n$ ne dit rien des lignes. Pour une matrice *carrée*, c'est l'inversibilité qui permet de passer de $G^TG = I_n$ à $GG^T = I_n$; pour une matrice rectangulaire à colonnes orthonormées, les lignes ne sont pas orthonormées en général.
À retenir. Dans une base quelconque, n'importe quel n-uplet non nul peut servir de coordonnées à x ; dans une base orthonormée, la norme est la seule contrainte. C'est la traduction de la proposition 8 : un espace euclidien « ressemble » à $\mathbb{R}^n$, et ses bases orthonormées sont toutes équivalentes.

EXERCICE 26 ⋆⋆⋆

1. La matrice A est inversible, donc son rang vaut n, et ses colonnes engendrent $\mathcal{M}_{n,1}(\mathbb{R})$ (chapitre 16, théorème 3) : ces n colonnes forment une base. La colonne j de la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$ contient les coordonnées de C_j dans la base canonique, c'est-à-dire C_j elle-même.

$$\boxed{\mathcal{C} \text{ est une base de } \mathcal{M}_{n,1}(\mathbb{R}), \qquad P_{\mathcal{B}}^{\mathcal{C}} = A}$$

La colonne j de $T = P_{\mathcal{D}}^{\mathcal{C}}$ contient les coordonnées de C_j dans la base orthonormée $\mathcal{D}$, qui sont les produits scalaires $\langle C_j, D_i\rangle$ (proposition 8). Par le théorème 4, $C_j \in \text{Vect}(C_1, \dots, C_j) = \text{Vect}(D_1, \dots, D_j)$, donc $\langle C_j, D_i\rangle = 0$ pour $i > j$; et $t_{jj} = \langle C_j, D_j\rangle > 0$, toujours par le théorème 4.

$$\boxed{T = P_{\mathcal{D}}^{\mathcal{C}} = \left(\langle C_j, D_i\rangle\right)_{i,j} \text{ est triangulaire supérieure, et } t_{jj} = \langle C_j, D_j\rangle > 0}$$

2. Par la formule de composition des matrices de passage (chapitre 16, proposition 13), on a $P_{\mathcal{B}}^{\mathcal{C}} = P_{\mathcal{B}}^{\mathcal{D}} P_{\mathcal{D}}^{\mathcal{C}}$, c'est-à-dire $A = OT$ avec $O = P_{\mathcal{B}}^{\mathcal{D}}$, dont les colonnes sont D_1, …, D_n. La matrice O^TO est la matrice de Gram de la famille orthonormée $\mathcal{D}$ (proposition 9 (i)), c'est-à-dire I_n.

$$\boxed{A = OT, \qquad O^TO = I_n, \qquad T \text{ triangulaire supérieure à diagonale } > 0}$$

3. Soit $A = O'T'$ une autre décomposition de ce type, et D'_1, …, D'_n les colonnes de O', qui forment une famille orthonormée puisque ${O'}^T O' = I_n$. L'égalité $A = O'T'$ s'écrit, colonne par colonne, $C_j = \sum_{i \leqslant j} t'_{ij} D'_i$, puisque T' est triangulaire supérieure. Ainsi $\text{Vect}(C_1, \dots, C_j) \subset \text{Vect}(D'_1, \dots, D'_j)$, et les deux sous-espaces sont de dimension j, puisque ces familles sont libres : ils sont égaux (chapitre 15, théorème 5). De plus, $\langle D'_j, C_j\rangle = t'_{jj} > 0$ par orthonormalité. La famille (D'_j) vérifie donc les deux propriétés qui caractérisent la famille de Gram-Schmidt dans le théorème 4 : par l'unicité affirmée par ce théorème, $D'_j = D_j$ pour tout j, donc $O' = O$. Enfin, $T' = O^TO'T' = O^TA = O^TOT = T$.

$$\boxed{\text{le couple } (O, T) \text{ est unique}}$$

Piège. Sans la condition de signe sur la diagonale, l'unicité tombe : on peut changer le signe d'une colonne de O et de la ligne correspondante de T. C'est la condition $\langle u_k, e_k\rangle > 0$ du théorème 4 qui fixe ces signes.

4. On orthonormalise $C_1 = (1,-1)$ et $C_2 = (2,1)$ par la méthode du cours. On a $\|C_1\| = \sqrt{2}$ et $D_1 = \dfrac{1}{\sqrt{2}}(1,-1)$; puis $\langle C_2, D_1\rangle = \dfrac{1}{\sqrt{2}}$, et

$$v_2 = C_2 - \langle C_2, D_1\rangle D_1 = (2,1) - \frac{1}{2}(1,-1) = \Big(\frac{3}{2}, \frac{3}{2}\Big), \qquad \|v_2\| = \frac{3}{\sqrt{2}}, \qquad D_2 = \frac{1}{\sqrt{2}}(1,1).$$

La matrice T contient les produits scalaires $\langle C_j, D_i\rangle$: $t_{11} = \sqrt{2}$, $t_{12} = \dfrac{1}{\sqrt{2}}$ et $t_{22} = \langle C_2, D_2\rangle = \dfrac{3}{\sqrt{2}}$.

$$\boxed{O = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad T = \begin{pmatrix} \sqrt{2} & \dfrac{\sqrt{2}}{2} \\ 0 & \dfrac{3\sqrt{2}}{2} \end{pmatrix}}$$

On vérifie le produit : $OT = \dfrac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2} & \dfrac{1}{\sqrt{2}} + \dfrac{3}{\sqrt{2}} \\ -\sqrt{2} & -\dfrac{1}{\sqrt{2}} + \dfrac{3}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = A$. ✓

5. Par le théorème 6 du chapitre 17, $\det(O^T O) = (\det O)^2 = 1$, donc $|\det O| = 1$, et le déterminant d'une matrice triangulaire est le produit de ses coefficients diagonaux. Ainsi

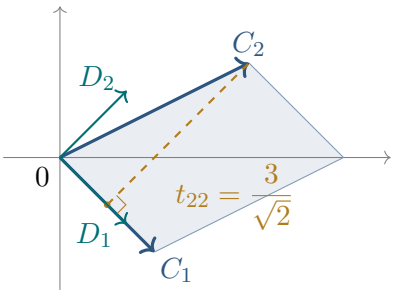
$$|\det A| = |\det O| \det T = t_{11} t_{22} \cdots t_{nn}.$$

Chaque $t_{jj} = \langle C_j, D_j\rangle$ est majoré par $\|C_j\|\,\|D_j\| = \|C_j\|$, par l'inégalité de Cauchy-Schwarz.

$$\boxed{|\det A| = \prod_{j=1}^{n} t_{jj} \leqslant \|C_1\|\,\|C_2\| \cdots \|C_n\|}$$

Pour la matrice de la question 4, $|\det A| = 3 \leqslant \sqrt{2}\cdot\sqrt{5} = \sqrt{10} \approx 3{,}16$. ✓ Si A n'est pas inversible, $\det A = 0$ et l'inégalité est évidente : elle reste vraie. Pour A inversible, la colonne j de T donne $\|C_j\|^2 = \sum_{i\leqslant j} t_{ij}^2$ (proposition 8), donc $t_{jj} = \|C_j\|$ si et seulement si $t_{ij} = 0$ pour $i < j$: il y a égalité si et seulement si T est diagonale, c'est-à-dire si les colonnes de A sont deux à deux orthogonales.

Pour $n = 3$, $|\det A|$ est le volume du parallélépipède construit sur C_1, C_2, C_3 (chapitre 17, et remarque qui suit la proposition 9) : il est au plus égal au produit des longueurs des arêtes, avec égalité pour un pavé droit. La décomposition $A = OT$ en donne la raison : t_{11} est la longueur de la première arête, t_{22} la hauteur du parallélogramme de base (C_1, C_2), et t_{33} la hauteur du parallélépipède au-dessus de ce parallélogramme. La figure illustre le cas $n = 2$ de la question 4.



L'aire du parallélogramme vaut la base $\|C_1\| = t_{11} = \sqrt{2}$ multipliée par la hauteur $t_{22} = \dfrac{3}{\sqrt{2}}$, soit $3 = |\det A|$; la hauteur, en pointillés, est plus courte que le côté $\|C_2\| = \sqrt{5} \approx 2{,}24$, et c'est toute l'inégalité de Hadamard.

À retenir. La décomposition $A = OT$ est l'algorithme de Gram-Schmidt écrit en matrices : O contient la base orthonormée obtenue, T les coefficients $\langle C_j, D_i\rangle$ calculés en chemin. L'inégalité de Hadamard dit qu'un volume est maximal, à longueurs d'arêtes fixées, quand les arêtes sont orthogonales.

EXERCICE 27 ⋆⋆⋆

1. Pour $0 \leqslant i, j \leqslant n$, la primitive usuelle de $t \mapsto t^{i+j}$ donne

$$\langle X^i, X^j \rangle = \int_0^1 t^{i+j}\,\mathrm{d}t = \frac{1}{i+j+1},$$

et H_n est bien la matrice de Gram $\left(\langle X^i, X^j\rangle\right)_{i,j}$ de la famille $(1, X, \dots, X^n)$. Soit $X = (x_0, \dots, x_n)$ une colonne, et $P = \sum_i x_i X^i$. Par bilinéarité,

$$X^T H_n X = \sum_{i=0}^n \sum_{j=0}^n x_i x_j \langle X^i, X^j\rangle = \Big\langle \sum_i x_i X^i, \sum_j x_j X^j \Big\rangle = \int_0^1 P(t)^2\,\mathrm{d}t.$$

Si $H_n X = 0$, alors $X^T H_n X = 0$, donc $\|P\|^2 = 0$, et $P = 0$, puisque la formule définit un produit scalaire sur $\mathbb{R}[X]$ (proposition 2 (ii)) : tous les x_i sont nuls. Le noyau de H_n est réduit à $\{0\}$, et H_n est inversible (chapitre 16, théorème 3).

$$\boxed{H_n = \left(\langle X^i, X^j\rangle\right)_{0\leqslant i,j\leqslant n}, \qquad X^T H_n X = \int_0^1 P^2 > 0 \text{ si } X \neq 0, \qquad H_n \in \mathrm{GL}_{n+1}(\mathbb{R})}$$

L'exercice 48 de la feuille du chapitre 17 obtenait l'inversibilité en calculant le déterminant de Cauchy, après de lourdes manipulations de lignes et de colonnes. Ici, aucun calcul : l'inversibilité vient de ce que $(1, X, \dots, X^n)$ est libre et que l'intégrale de P^2 est un produit scalaire. La proposition 9 (ii), appliquée dans une base orthonormée de l'espace euclidien $\mathbb{R}_n[X]$, donne même $\det H_n > 0$; le calcul donne $\det H_1 = \dfrac{1}{12}$, $\det H_2 = \dfrac{1}{2160}$ et $\det H_3 = \dfrac{1}{6\,048\,000}$, valeurs très petites qui expliquent la réputation numérique de cette matrice. ✓

2. Les deux membres dépendent linéairement de Q, par linéarité de l'intégrale (chapitre 18, théorème 3, et proposition 8 pour les fonctions complexes) : il suffit de vérifier l'égalité pour $Q = X^k$, $k \in \mathbb{N}$. On intègre au moyen des primitives usuelles : celle de t^k à gauche, celle de $\theta \mapsto \mathrm{e}^{\mathrm{i}(k+1)\theta}$, qui est $\theta \mapsto \dfrac{\mathrm{e}^{\mathrm{i}(k+1)\theta}}{\mathrm{i}(k+1)}$, à droite.

$$\int_{-1}^1 t^k\,\mathrm{d}t = \frac{1-(-1)^{k+1}}{k+1} = \frac{1+(-1)^k}{k+1}, \qquad -\mathrm{i}\int_0^\pi \mathrm{e}^{\mathrm{i}(k+1)\theta}\,\mathrm{d}\theta = -\mathrm{i}\cdot\frac{(-1)^{k+1}-1}{\mathrm{i}(k+1)} = \frac{1+(-1)^k}{k+1}.$$

$$\boxed{\forall Q \in \mathbb{R}[X], \quad \int_{-1}^1 Q(t)\,\mathrm{d}t = -\mathrm{i}\int_0^\pi Q(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta}$$

Pour $Q = (1+X)^3 + 2X^4$, les deux membres valent $\dfrac{24}{5}$, d'après un calcul symbolique. ✓ On peut lire cette égalité comme le fait qu'on va de -1 à 1 « par le segment » ou « par le demi-cercle supérieur » avec la même intégrale : c'est une première rencontre avec l'intégration le long d'un chemin du plan complexe.

3. Avec $X = (a_0, \dots, a_n)$, la question 1 donne $\displaystyle\sum_{k,\ell} \frac{a_k a_\ell}{k+\ell+1} = \int_0^1 P^2$. Comme $P^2 \geqslant 0$ sur $[-1\,,0]$, la relation de Chasles et la positivité de l'intégrale donnent $\displaystyle\int_0^1 P^2 \leqslant \int_{-1}^1 P^2$. On applique la question 2 au polynôme $Q = P^2$; le premier membre étant un réel positif, il est égal à son module, et l'inégalité triangulaire complexe (chapitre 18, théorème 5), avec $|\mathrm{e}^{\mathrm{i}\theta}| = 1$, donne

$$\int_{-1}^1 P^2 = \Big|\int_0^\pi P(\mathrm{e}^{\mathrm{i}\theta})^2\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta\Big| \leqslant \int_0^\pi \left|P(\mathrm{e}^{\mathrm{i}\theta})\right|^2\mathrm{d}\theta.$$

Les coefficients a_k étant réels, $P(\mathrm{e}^{-\mathrm{i}\theta}) = \overline{P(\mathrm{e}^{\mathrm{i}\theta})}$: la fonction $h : \theta \mapsto |P(\mathrm{e}^{\mathrm{i}\theta})|^2$ est paire, et 2π-périodique. La proposition 6 du chapitre 6 donne alors

$$\int_0^\pi h = \frac{1}{2}\int_{-\pi}^{\pi} h = \frac{1}{2}\int_0^{2\pi} h = \frac{1}{2}\times 2\pi \sum_{k=0}^{n} a_k^2,$$

la dernière égalité étant celle de l'indication. En enchaînant les inégalités :

$$\boxed{\sum_{0\leqslant k,\ell\leqslant n} \frac{a_k a_\ell}{k+\ell+1} = \int_0^1 P^2 \leqslant \int_{-1}^1 P^2 \leqslant \int_0^\pi \left|P(\mathrm{e}^{\mathrm{i}\theta})\right|^2 \mathrm{d}\theta = \pi \sum_{k=0}^{n} a_k^2}$$

Pour $(a_0, a_1, a_2, a_3) = (1, -2, 3, 1)$, la chaîne s'écrit $\dfrac{83}{42} \leqslant \dfrac{230}{21} \leqslant 15\pi$, ce que confirme un calcul symbolique. ✓

À retenir. La constante π ne dépend pas de n : c'est ce qui fait la force de l'inégalité. On peut montrer, ce que l'on admet, qu'elle ne peut pas être remplacée par une constante plus petite valable pour tout n ; la plus grande valeur du quotient $\dfrac{X^T H_n X}{X^T X}$, calculée numériquement, vaut environ 1,27 pour $n = 1$, 1,92 pour $n = 20$ et 2,27 pour $n = 200$: elle s'approche de π très lentement.

Projection orthogonale sur un sous-espace de dimension finie

EXERCICE 28 ⋆

1. Par la proposition 12 (iii), le vecteur $n = (2, 1, -1)$ est normal au plan P, et $P = \mathrm{Vect}(n)^\perp$. On a $\langle u, n\rangle = 6 + 4 - 3 = 7$ et $\|n\|^2 = 6$; la proposition 13 donne la distance, et le projeté :

$$\boxed{d(u,P) = \frac{|\langle u,n\rangle|}{\|n\|} = \frac{7}{\sqrt{6}} = \frac{7\sqrt{6}}{6}}, \qquad \boxed{p_P(u) = u - \frac{7}{6}n = \left(\frac{2}{3}, \frac{17}{6}, \frac{25}{6}\right)}$$

On contrôle que $p_P(u)$ appartient à P : $\dfrac{4}{3} + \dfrac{17}{6} - \dfrac{25}{6} = \dfrac{8+17-25}{6} = 0$. ✓

2. Le vecteur $n = (1, 1, 2)$ est normal à H, avec $\|n\|^2 = 6$. Par la proposition 13, $p_H(x) = x - \dfrac{\langle x, n\rangle}{6} n$. En colonnes, $\langle x, n\rangle n = n(n^T X) = (nn^T)X$, de sorte que la matrice de p_H est $I_3 - \dfrac{1}{6}nn^T$, avec

$$nn^T = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{pmatrix}, \qquad \boxed{\mathrm{Mat}(p_H) = \frac{1}{6}\begin{pmatrix} 5 & -1 & -2 \\ -1 & 5 & -2 \\ -2 & -2 & 2 \end{pmatrix}}$$

On contrôle que $p_H(n) = 0$ et que $p_H(1, -1, 0) = (1, -1, 0)$, vecteur de H. ✓ La symétrie par rapport à H parallèlement à $H^\perp$ est $s = 2p_H - \mathrm{id}$ (chapitre 14, définition des symétries), de matrice $S = I_3 - \dfrac{1}{3}nn^T$:

$$\boxed{S = \mathrm{Mat}(s) = \frac{1}{3}\begin{pmatrix} 2 & -1 & -2 \\ -1 & 2 & -2 \\ -2 & -2 & -1 \end{pmatrix}}$$

Elle est symétrique, puisque $(nn^T)^T = nn^T$. Comme $n^T n = \|n\|^2 = 6$, on a

$$(nn^T)^2 = n(n^T n)n^T = 6\,nn^T, \qquad \text{donc} \qquad \boxed{S^2 = I_3 - \frac{2}{3}nn^T + \frac{1}{9}\times 6\,nn^T = I_3}$$

ce que confirme le produit des matrices. ✓

À retenir. Les deux propriétés vérifiées donnent $S^T S = S^2 = I_3$: les colonnes de S forment une base orthonormée (remarque qui suit la proposition 9), et s conserve la norme. Une symétrie par rapport à H parallèlement à une autre droite que $H^\perp$ vérifierait encore $S^2 = I_3$, mais sa matrice ne serait pas symétrique.

EXERCICE 29 ⋆

Les deux équations de F s'écrivent $\langle x, n_1\rangle = 0$ et $\langle x, n_2\rangle = 0$, où l'on a posé $n_1 = (1,1,0,1)$ et $n_2 = (1,1,2,-1)$, vecteurs non colinéaires. Ainsi $F = \text{Vect}(n_1, n_2)^\perp$, et, par la proposition 11, $F^\perp = \text{Vect}(n_1, n_2)$, de dimension 2, comme F. Il est plus court de projeter u sur $F^\perp$, puis d'utiliser $p_F = \text{id} - p_{F^\perp}$ (proposition 10 (iii)). Comme $\langle n_1, n_2\rangle = 1 \neq 0$, la base (n_1, n_2) n'est pas orthogonale : on écrit $p_{F^\perp}(u) = \alpha n_1 + \beta n_2$, et l'on résout les équations normales (proposition 10 (ii)), avec $\langle u, n_1\rangle = 10$, $\langle u, n_2\rangle = 10$, $\|n_1\|^2 = 3$ et $\|n_2\|^2 = 7$:

$$\begin{cases} 3\alpha + \beta = 10 \\ \alpha + 7\beta = 10 \end{cases} \qquad \text{d'où} \qquad \alpha = 3, \quad \beta = 1, \quad p_{F^\perp}(u) = 3n_1 + n_2 = (4,4,2,2).$$

$$\boxed{p_F(u) = u - (4,4,2,2) = (-3,4,-1,-1), \qquad d(u,F) = \|(4,4,2,2)\| = \sqrt{40} = 2\sqrt{10}}$$

On contrôle que $p_F(u)$ vérifie les deux équations : $-3+4-1=0$ et $-3+4-2+1=0$. ✓

Piège. L'erreur la plus répandue est d'additionner les projetés de u sur les deux droites $\text{Vect}(n_1)$ et $\text{Vect}(n_2)$, ce qui donnerait $\frac{10}{3}n_1 + \frac{10}{7}n_2$: la formule de la proposition 10 (i) exige une base orthogonale (première remarque qui suit cette proposition), et n_1, n_2 ne sont pas orthogonaux.

EXERCICE 30 ⋆

1. Les deux équations sont indépendantes, donc G est de dimension 2 (chapitre 15, théorème 13). Les vecteurs $(1,-1,0,0)$ et $(0,0,1,-1)$ appartiennent à G, et ils sont orthogonaux : il suffit de les normaliser.

$$\boxed{\varepsilon_1 = \frac{1}{\sqrt{2}}(1,-1,0,0), \qquad \varepsilon_2 = \frac{1}{\sqrt{2}}(0,0,1,-1)}$$

2. Par la proposition 10 (i), $p_G(x) = \langle x, \varepsilon_1\rangle\varepsilon_1 + \langle x, \varepsilon_2\rangle\varepsilon_2$, avec $\langle x, \varepsilon_1\rangle\varepsilon_1 = \dfrac{x_1 - x_2}{2}(1,-1,0,0)$ et $\langle x, \varepsilon_2\rangle\varepsilon_2 = \dfrac{x_3 - x_4}{2}(0,0,1,-1)$.

$$\boxed{\text{Mat}(p_G) = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}}$$

3. Par le théorème 7, $d(x,G) = \|x - p_G(x)\|$, et

$$x - p_G(x) = \frac{x_1 + x_2}{2}(1,1,0,0) + \frac{x_3 + x_4}{2}(0,0,1,1),$$

somme de deux vecteurs orthogonaux de normes au carré $\dfrac{(x_1+x_2)^2}{4} \times 2$ et $\dfrac{(x_3+x_4)^2}{4} \times 2$: le théorème de Pythagore conclut.

$$\boxed{d(x,G) = \sqrt{\frac{(x_1+x_2)^2 + (x_3+x_4)^2}{2}}}$$

Le résultat se lit directement : $x - p_G(x)$ est le projeté de x sur $G^\perp = \text{Vect}((1,1,0,0),(0,0,1,1))$, et la distance à G ne dépend que des premiers membres des équations de G, nuls exactement sur G.

EXERCICE 31 ⋆

1. L'application $\varphi : P \mapsto \displaystyle\int_0^1 P(t)\,\mathrm{d}t$ est une forme linéaire sur $\mathbb{R}_n[X]$ (chapitre 18, théorème 3), non nulle puisque $\varphi(1) = 1$. Son noyau H est un hyperplan de $\mathbb{R}_n[X]$, qui est de dimension $n+1$ (chapitre 15, théorème 12).

$$\boxed{H = \operatorname{Ker}\varphi \text{ est un sous-espace de } \mathbb{R}_n[X], \qquad \dim H = n}$$

2. La droite $\mathbb{R}_0[X]$ des polynômes constants n'est pas contenue dans H, puisque $1 \notin H$: par le théorème 7 du chapitre 14, $\mathbb{R}_n[X] = \mathbb{R}_0[X] \oplus H$. La décomposition de P s'écrit explicitement : la constante $c = \displaystyle\int_0^1 P$ vérifie $\displaystyle\int_0^1 (P-c) = c - c = 0$, donc $P = c + (P-c)$ avec $c \in \mathbb{R}_0[X]$ et $P - c \in H$.

$$\boxed{p(P) = \int_0^1 P(t)\,\mathrm{d}t \quad \text{(polynôme constant)}}$$

3. On munit $\mathbb{R}_n[X]$ du produit scalaire $\langle P, Q\rangle = \displaystyle\int_0^1 PQ$ (proposition 2 (ii), restreinte à $\mathbb{R}_n[X]$). Pour lui, $H = \{P \mid \langle P, 1\rangle = 0\} = \mathbb{R}_0[X]^\perp$: le projecteur sur $\mathbb{R}_0[X]$ parallèlement à H est le projecteur sur $\mathbb{R}_0[X]$ parallèlement à $\mathbb{R}_0[X]^\perp$, c'est-à-dire la projection orthogonale sur $\mathbb{R}_0[X]$.

$$\boxed{\langle P, Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t}$$

La proposition 10 (i) le confirme : $\|1\| = 1$, et le projeté de P sur $\mathrm{Vect}(1)$ est $\langle P, 1\rangle\, 1 = \displaystyle\int_0^1 P$. ✓ La réponse n'est pas unique : tout produit scalaire pour lequel 1 est orthogonal à H convient. Ainsi la projection orthogonale sur une droite calcule une moyenne, et le choix du produit scalaire décide de laquelle.

EXERCICE 32 ⋆

1. On a $F = \mathrm{Vect}(I_2, K)$ avec $K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, famille libre, donc $\dim F = 2$ et $\dim F^\perp = 2$ (proposition 11). Pour $M = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, le produit scalaire se calcule en multipliant les coefficients de même position (proposition 1 (ii)) : $\langle M, I_2\rangle = p + s$ et $\langle M, K\rangle = q - r$. Ainsi $M \in F^\perp$ si et seulement si $s = -p$ et $r = q$:

$$F^\perp = \left\{ \begin{pmatrix} c & d \\ d & -c \end{pmatrix} \;\middle|\; (c,d) \in \mathbb{R}^2 \right\}.$$

Les deux matrices obtenues pour $(c,d) = (1,0)$ et $(0,1)$ sont orthogonales, de norme $\sqrt{2}$:

$$\boxed{E_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}$$

2. On a $\langle J, E_1\rangle = \dfrac{1-1}{\sqrt{2}} = 0$ et $\langle J, E_2\rangle = \dfrac{2}{\sqrt{2}} = \sqrt{2}$; la proposition 10 (i) donne

$$\boxed{p_{F^\perp}(J) = \sqrt{2}\,E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}$$

3. Par la proposition 10 (iii), $J - p_{F^\perp}(J) = p_F(J)$, donc $J - p_F(J) = p_{F^\perp}(J)$, et le théorème 7 donne :

$$\boxed{d(J, F) = \|p_{F^\perp}(J)\| = \sqrt{2}, \qquad p_F(J) = J - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = I_2}$$

Un calcul direct le confirme : $\|J-(aI_2+bK)\|^2 = 2(1-a)^2+(1-b)^2+(1+b)^2$ est minimal pour $a=1$, $b=0$, et vaut alors 2. ✓ Le sous-espace F est formé des matrices qui représentent les nombres complexes $a+\mathrm{i}b$ agissant sur le plan par multiplication (au signe de b près) ; parmi elles, la plus proche de J est l'identité.

EXERCICE 33 ⋆⋆

1. Un vecteur y est orthogonal à tous les éléments de $F\cup G$ si et seulement s'il est orthogonal à tous ceux de F et à tous ceux de G : $(F\cup G)^\perp = F^\perp\cap G^\perp$. Par ailleurs, $F+G$ est le sous-espace engendré par $F\cup G$ (chapitre 14, proposition 11), et la proposition 5 (ii) donne $A^\perp = (\mathrm{Vect}A)^\perp$.

$$\boxed{(F\cup G)^\perp = (F+G)^\perp = F^\perp\cap G^\perp}$$

2. Comme $F\cap G\subset F$, la proposition 5 (ii) donne $F^\perp\subset(F\cap G)^\perp$, et de même $G^\perp\subset(F\cap G)^\perp$. Le sous-espace $(F\cap G)^\perp$ (proposition 5 (i)) contient donc la somme $F^\perp+G^\perp$.

$$\boxed{F^\perp+G^\perp\subset(F\cap G)^\perp}$$

3. On applique la question 1 aux sous-espaces $F^\perp$ et $G^\perp$, puis la proposition 11, qui donne $(F^\perp)^\perp = F$ en dimension finie :

$$(F^\perp+G^\perp)^\perp = (F^\perp)^\perp\cap(G^\perp)^\perp = F\cap G.$$

On prend l'orthogonal des deux membres, et l'on utilise encore la proposition 11 pour le sous-espace $F^\perp+G^\perp$. Pour une partie A quelconque, la proposition 5 (ii) et la proposition 11 donnent $(A^\perp)^\perp = ((\mathrm{Vect}A)^\perp)^\perp = \mathrm{Vect}A$.

$$\boxed{(F\cap G)^\perp = F^\perp+G^\perp, \qquad (A^\perp)^\perp = \mathrm{Vect}(A) \qquad (E \text{ euclidien})}$$

4. La remarque qui suit la proposition 11 du cours, faite avec ce produit scalaire, établit que $F^\perp=\{0\}$ pour l'hyperplan $F=\{P\mid P(1)=0\}$. Un polynôme constant c appartient à F si et seulement si $c=0$: $F\cap G=\{0\}$, d'où $(F\cap G)^\perp=\mathbb{R}[X]$. D'autre part, $G^\perp=\{1\}^\perp$ est l'ensemble des polynômes dont le coefficient constant $b_0=\langle Q,1\rangle$ est nul. Ainsi

$$\boxed{F^\perp+G^\perp=\{Q\in\mathbb{R}[X]\mid Q(0)=0\}\neq\mathbb{R}[X]=(F\cap G)^\perp}$$

puisque le polynôme 1 n'appartient pas au membre de gauche.
À retenir. Les deux formules de la question 3 reposent sur $(F^\perp)^\perp=F$, c'est-à-dire sur la proposition 11, qui suppose la dimension finie. En dimension infinie, seule l'inclusion de la question 2 subsiste, et la question 4 montre qu'elle peut être stricte. L'exercice 39 donne un autre exemple de sous-espace d'orthogonal nul.

EXERCICE 34 ⋆⋆

1. Par la proposition 1 (ii), la somme est $\|M-S\|^2$: il s'agit de calculer $d(M,\mathcal{S}_n(\mathbb{R}))^2$. Le sous-espace $\mathcal{S}_n(\mathbb{R})$ est de dimension finie, et l'exemple qui suit la proposition 11 en donne le projeté : $p_{\mathcal{S}_n(\mathbb{R})}(M)=\dfrac{1}{2}(M+M^T)$, avec $M-p_{\mathcal{S}_n(\mathbb{R})}(M)=\dfrac{1}{2}(M-M^T)$. Par le théorème 7 :

$$\boxed{\inf_{S\in\mathcal{S}_n(\mathbb{R})}\sum_{i,j}(m_{ij}-s_{ij})^2=\frac{1}{4}\sum_{i=1}^n\sum_{j=1}^n(m_{ij}-m_{ji})^2=\frac{1}{2}\sum_{i<j}(m_{ij}-m_{ji})^2}$$

La borne inférieure est un minimum, atteint pour la seule matrice $S=\dfrac{1}{2}(M+M^T)$. Elle est nulle si et seulement si M est symétrique, ce qui est cohérent.

2. On projette M sur $F = \mathrm{Vect}(I_n, J)$, qui est un plan puisque J n'est pas proportionnelle à I_n pour $n \geqslant 2$. On a $\langle I_n, J\rangle = \operatorname{tr} J = n = \|I_n\|^2$, de sorte que $J - I_n$ est orthogonale à I_n : la famille $(I_n, J - I_n)$ est une base orthogonale de F, avec $\|J - I_n\|^2 = n^2 - 2n + n = n(n-1)$. Par ailleurs, $\langle M, I_n\rangle = \operatorname{tr} M = 0$ et $\langle M, J - I_n\rangle = s - 0 = s$. La formule de la proposition 10 (i), adaptée aux bases orthogonales, donne

$$p_F(M) = \frac{\langle M, I_n\rangle}{n} I_n + \frac{\langle M, J - I_n\rangle}{n(n-1)}(J - I_n) = \frac{s}{n(n-1)}(J - I_n), \qquad \|p_F(M)\|^2 = \frac{s^2}{n(n-1)}.$$

Par le théorème 7, $d(M, F)^2 = \|M\|^2 - \|p_F(M)\|^2$:

$$\boxed{\inf_{(\alpha,\beta)\in\mathbb{R}^2} \|M - \alpha I_n - \beta J\| = \sqrt{\sum_{i,j} m_{ij}^2 - \frac{s^2}{n(n-1)}}}$$

atteinte pour $\beta = -\alpha = \dfrac{s}{n(n-1)}$ seulement. Une minimisation numérique, sur des matrices de trace nulle tirées au hasard pour n de 2 à 5, redonne cette valeur. ✓
Piège. Pour $n = 1$, on a $J = I_1$ et la famille (I_n, J) est liée ; c'est pourquoi l'énoncé suppose $n \geqslant 2$, et la formule, qui divise par $n(n-1)$, n'a pas de sens pour $n = 1$. Remplacer J par $J - I_n$, orthogonale à I_n, évite de résoudre le système des équations normales.

EXERCICE 35 ⋆⋆

(a) Toutes les intégrales à calculer sont celles de fonctions puissances $t \mapsto t^\alpha$, que l'on intègre par la primitive usuelle $\dfrac{t^{\alpha+1}}{\alpha+1}$. Dans $\mathcal{C}([0\,,1], \mathbb{R})$ muni de $\langle f, g\rangle = \displaystyle\int_0^1 fg$, l'intégrale vaut $\|t - (a\sqrt{t} + b)\|^2$: on cherche la distance de $t \mapsto t$ au plan $F = \mathrm{Vect}(\sqrt{t}, 1)$, la famille étant libre puisque $\sqrt{t}$ n'est pas constante. Par le théorème 7, la borne inférieure est atteinte pour un unique couple (a, b), celui du projeté, caractérisé par les équations normales (proposition 10 (ii)). Avec

$$\|\sqrt{t}\|^2 = \frac{1}{2}, \quad \langle\sqrt{t}, 1\rangle = \frac{2}{3}, \quad \|1\|^2 = 1, \quad \langle t, \sqrt{t}\rangle = \frac{2}{5}, \quad \langle t, 1\rangle = \frac{1}{2},$$

elles s'écrivent

$$\begin{cases} \dfrac{a}{2} + \dfrac{2b}{3} = \dfrac{2}{5} \\ \dfrac{2a}{3} + b = \dfrac{1}{2} \end{cases} \qquad \text{d'où} \qquad a = \frac{6}{5}, \quad b = -\frac{3}{10}.$$

Par le théorème 7, le carré de la distance vaut $\|t\|^2 - \|p_F(t)\|^2 = \langle t, t - p_F(t)\rangle$, puisque $t - p_F(t)$ est orthogonal à $p_F(t)$; or $\langle t, p_F(t)\rangle = a \cdot \dfrac{2}{5} + b \cdot \dfrac{1}{2} = \dfrac{12}{25} - \dfrac{3}{20} = \dfrac{33}{100}$.

$$\boxed{\inf_{(a,b)\in\mathbb{R}^2} \int_0^1 \left(t - a\sqrt{t} - b\right)^2 \mathrm{d}t = \frac{1}{3} - \frac{33}{100} = \frac{1}{300}, \quad \text{atteinte en } (a, b) = \left(\frac{6}{5}, -\frac{3}{10}\right) \text{ seulement}}$$

Le résultat est confirmé en minimisant directement le polynôme du second degré en (a, b) par un calcul symbolique. ✓ L'exercice 47 retrouvera cette valeur comme quotient de deux déterminants de Gram.

(b) On calcule les produits scalaires $\displaystyle\int_0^{2\pi} t\cos t\,\mathrm{d}t$ et $\displaystyle\int_0^{2\pi} t\sin t\,\mathrm{d}t$ par une intégration par parties (chapitre 6, proposition 4), en dérivant t : elle fait disparaître le facteur t. Dans $\mathcal{C}([0\,,2\pi], \mathbb{R})$ muni de $\langle f, g\rangle = \displaystyle\int_0^{2\pi} fg$, on cherche la distance de $t \mapsto t$ au plan $F = \mathrm{Vect}(\cos, \sin)$. L'exemple qui suit le théorème 3 donne une base orthogonale : $\langle\cos, \sin\rangle = 0$ et $\|\cos\|^2 = \|\sin\|^2 = \pi$. Les intégrations par parties donnent

$$\langle t, \cos\rangle = \big[t\sin t\big]_0^{2\pi} - \int_0^{2\pi} \sin t\,\mathrm{d}t = 0, \qquad \langle t, \sin\rangle = \big[-t\cos t\big]_0^{2\pi} + \int_0^{2\pi} \cos t\,\mathrm{d}t = -2\pi.$$

Les fonctions $t\sin t+\cos t$ et $\sin t - t\cos t$ ont bien pour dérivées $t\cos t$ et $t\sin t$. ✓ La proposition 10 (i), sous sa forme adaptée aux bases orthogonales, donne $p_F(t) = \dfrac{0}{\pi}\cos + \dfrac{-2\pi}{\pi}\sin = -2\sin$, d'où $a=0$ et $b=-2$. Avec $\|t\|^2 = \displaystyle\int_0^{2\pi} t^2\,\mathrm{d}t = \frac{8\pi^3}{3}$ et $\|p_F(t)\|^2 = 4\pi$, le théorème 7 donne

$$\boxed{\inf_{(a,b)\in\mathbb{R}^2} \int_0^{2\pi} \big(t - a\cos t - b\sin t\big)^2\,\mathrm{d}t = \frac{8\pi^3}{3} - 4\pi \approx 70{,}1, \quad \text{atteinte en } (a,b)=(0,-2) \text{ seulement}}$$

La distance reste grande, du même ordre que $\|t\|^2 \approx 82{,}7$: une fonction qui croît de 0 à 2π est mal approchée par des fonctions 2π-périodiques, qui reprennent en 2π leur valeur en 0.
Piège. Dans (a), la famille $(\sqrt{t}, 1)$ n'est pas orthogonale : appliquer la formule de la proposition 10 (i) sans résoudre le système est l'erreur la plus répandue. Dans (b), c'est au contraire l'orthogonalité de cos et sin qui dispense de tout système.

EXERCICE 36 ⋆⋆

Les intégrales à calculer sont celles de polynômes, que l'on intègre par les primitives usuelles. On munit $E = \mathcal{C}([0\,,1],\mathbb{R})$ du produit scalaire $\langle f,g\rangle = \displaystyle\int_0^1 fg$ (proposition 2), et l'on note $F = \mathrm{Vect}(1,x)$, de dimension 2. Les hypothèses s'écrivent $\langle f,1\rangle = \langle f,x\rangle = 1$.

Le projeté de f sur F. Par la proposition 10 (ii), $g = p_F(f)$ est l'unique élément de F tel que $f-g$ soit orthogonal à 1 et à x, c'est-à-dire tel que $\langle g,1\rangle = \langle g,x\rangle = 1$: il est le même pour toutes les fonctions f de l'énoncé. On cherche $g(x) = \alpha + \beta x$:

$$\begin{cases} \alpha + \dfrac{\beta}{2} = 1 \\ \dfrac{\alpha}{2} + \dfrac{\beta}{3} = 1 \end{cases} \qquad \text{d'où} \qquad \beta = 6, \quad \alpha = -2, \qquad \boxed{g(x) = 6x-2}$$

L'inégalité. On a $\|g\|^2 = \displaystyle\int_0^1 (36x^2 - 24x + 4)\,\mathrm{d}x = 12 - 12 + 4 = 4$. Le théorème de Pythagore, appliqué à $f = g + (f-g)$, donne $\|f\|^2 = 4 + \|f-g\|^2$:

$$\boxed{\int_0^1 f(x)^2\,\mathrm{d}x = 4 + \int_0^1 \big(f(x) - 6x + 2\big)^2\,\mathrm{d}x \geqslant 4}$$

Le cas d'égalité. Il y a égalité si et seulement si $\|f-g\| = 0$, c'est-à-dire $f = g$; et la fonction g vérifie bien les deux conditions.

$$\boxed{\text{égalité} \iff f(x) = 6x-2 \text{ pour tout } x \in [0\,,1]}$$

On peut aussi conclure par l'inégalité de Cauchy-Schwarz : $\langle f,g\rangle = 6\langle f,x\rangle - 2\langle f,1\rangle = 4$, donc $4 \leqslant \|f\|\,\|g\| = 2\|f\|$, avec égalité seulement si f est proportionnelle à g. Par exemple, $f = g + h$ avec $h(x) = x^2 - x + \dfrac{1}{6}$, orthogonale à 1 et à x, vérifie les hypothèses et donne $\|f\|^2 = 4 + \dfrac{1}{180}$. ✓
À retenir. Minimiser une norme sous des contraintes linéaires de la forme $\langle f, u_i\rangle = c_i$, c'est chercher le vecteur de plus petite norme d'un sous-espace affine : il appartient à $\mathrm{Vect}(u_i)$, et il est orthogonal à la direction du sous-espace affine.

EXERCICE 37 ⋆⋆

1. On intègre par parties (chapitre 6, proposition 4) le premier membre de la condition, en dérivant t^3-t et en primitivant P' : c'est le choix naturel, puisque t^3-t s'annule en 0 et en 1, et que l'on veut faire apparaître P seul. L'application

$$\psi : P \longmapsto \int_0^1 (t^3-t)P'(t)\,\mathrm{d}t - \int_0^1 P(t)\,\mathrm{d}t$$

est linéaire, puisque la dérivation et l'intégrale le sont : $F = \operatorname{Ker}\psi$ est un sous-espace de E.

$$\boxed{F = \operatorname{Ker}\psi \text{ est un sous-espace de } E}$$

2. L'intégration par parties annoncée donne, le crochet étant nul,

$$\int_0^1 (t^3-t)P'(t)\,\mathrm{d}t = \Big[(t^3-t)P(t)\Big]_0^1 - \int_0^1 (3t^2-1)P(t)\,\mathrm{d}t = -\int_0^1 (3t^2-1)P(t)\,\mathrm{d}t,$$

$$\text{d'où} \qquad \psi(P) = -\int_0^1 (3t^2-1)P(t)\,\mathrm{d}t - \int_0^1 P(t)\,\mathrm{d}t = -3\int_0^1 t^2P(t)\,\mathrm{d}t = -3\langle P, X^2\rangle.$$

Ainsi $F = \operatorname{Vect}(X^2)^\perp$ est un hyperplan de vecteur normal X^2, et la proposition 13 s'applique, avec $\|X^2\|^2 = \displaystyle\int_0^1 t^4\,\mathrm{d}t = \frac{1}{5}$ et

$$\langle Q, X^2\rangle = \int_0^1 (t^2+t^3+t^4+t^5)\,\mathrm{d}t = \frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6} = \frac{19}{20}.$$

$$\boxed{d(Q,F) = \frac{|\langle Q, X^2\rangle|}{\|X^2\|} = \frac{19}{20}\sqrt{5} = \frac{19\sqrt{5}}{20} \approx 2{,}12}$$

Un calcul symbolique confirme que $\psi(P) = -3\langle P, X^2\rangle$ pour un polynôme P de degré 3 quelconque. ✓

Piège. Le vecteur normal à F est X^2, et non X^3-X ou $3X^2-1$, qui apparaissent dans les calculs : il faut d'abord écrire la condition sous la forme $\langle P, N\rangle = 0$ pour lire le vecteur normal N.

EXERCICE 38 ⋆⋆

1. Quels que soient les a_k, la formule est bilinéaire, symétrique, et $\langle P, P\rangle = \displaystyle\sum_k P(a_k)^2 \geqslant 0$: seul le caractère défini est en jeu. Si les a_k sont deux à deux distincts et si $\langle P, P\rangle = 0$, alors $P(a_k) = 0$ pour tout k : le polynôme P, de degré au plus n, a $n+1$ racines distinctes, donc il est nul (chapitre 13, théorème 3). Si au contraire $a_i = a_j$ pour deux indices $i \neq j$, l'ensemble $\{a_0, \dots, a_n\}$ a au plus n éléments ; le produit P des $X-a$, pour a dans cet ensemble, est un polynôme non nul de degré au plus n, et $\langle P, P\rangle = 0$.

$$\boxed{\langle\cdot,\cdot\rangle \text{ est un produit scalaire sur } \mathbb{R}_n[X] \iff a_0, \dots, a_n \text{ sont deux à deux distincts}}$$

2. Les polynômes d'interpolation de Lagrange (chapitre 13, théorème 13),

$$L_j = \prod_{k\neq j} \frac{X-a_k}{a_j-a_k} \qquad (0 \leqslant j \leqslant n),$$

vérifient $L_j(a_k) = \delta_{jk}$, donc $\langle L_i, L_j\rangle = \displaystyle\sum_k \delta_{ik}\delta_{jk} = \delta_{ij}$. Cette famille orthonormée de $n+1$ vecteurs de l'espace $\mathbb{R}_n[X]$, de dimension $n+1$, en est une base orthonormée.

$$\boxed{(L_0, \dots, L_n) \text{ est une base orthonormée de } \mathbb{R}_n[X]}$$

C'est l'exemple de la section 19.5 du cours, pour $n = 2$; la proposition 8 y redonne la formule d'interpolation $P = \sum_j P(a_j)L_j$.

3. Pour tout P, $\sum_k P(a_k) = \sum_k P(a_k) \times 1 = \langle P, 1\rangle$: $F = \text{Vect}(1)^\perp$ est l'hyperplan de vecteur normal le polynôme constant 1, avec $\|1\|^2 = n+1$. La proposition 11 et la proposition 13 donnent :

$$\boxed{F^\perp = \mathbb{R}_0[X], \qquad d(Q,F) = \frac{|\langle Q,1\rangle|}{\|1\|} = \frac{|\sum_{k=0}^n Q(a_k)|}{\sqrt{n+1}}, \qquad d(X^n,F) = \frac{|\sum_{k=0}^n a_k^n|}{\sqrt{n+1}}}$$

Par exemple, pour $n = 2$ et $(a_0, a_1, a_2) = (0, 1, 2)$, on trouve $d(X^2, F) = \dfrac{0+1+4}{\sqrt{3}} = \dfrac{5}{\sqrt{3}}$. ✓

Piège. Le même produit n'est pas un produit scalaire sur $\mathbb{R}[X]$, ni sur $\mathbb{R}_m[X]$ pour $m > n$: le polynôme $(X-a_0)\cdots(X-a_n)$, non nul, y est de norme nulle. C'est la deuxième remarque qui suit la proposition 2 : l'espace sur lequel on travaille fait partie de la donnée.

EXERCICE 39 ⋆⋆

1. L'application $f \mapsto f(0)$ est une forme linéaire sur E, non nulle puisqu'elle vaut 1 en la fonction constante 1 : son noyau F est un hyperplan de E (chapitre 14). Soit $g \in F^\perp$. La fonction $f : t \mapsto tg(t)$ est continue sur $[0\,,1]$ et nulle en 0 : elle appartient à F, donc

$$0 = \langle f, g\rangle = \int_0^1 t\,g(t)^2\,\mathrm{d}t.$$

La fonction $t \mapsto tg(t)^2$ est continue, positive, d'intégrale nulle : elle est nulle sur $[0\,,1]$ (chapitre 18, théorème 4). Donc $g(t) = 0$ pour tout $t \in]0\,,1]$, et $g(0) = 0$ aussi, par continuité de g en 0.

$$\boxed{F \text{ est un hyperplan de } E, \qquad F^\perp = \{0\}}$$

2. Un supplémentaire orthogonal de F serait un supplémentaire G de F formé de vecteurs orthogonaux à F, c'est-à-dire contenu dans $F^\perp = \{0\}$. On aurait alors $E = F \oplus \{0\} = F$, ce qui est faux puisque la fonction constante 1 n'appartient pas à F.

$$\boxed{F \text{ n'admet pas de supplémentaire orthogonal : } F \oplus F^\perp = F \neq E}$$

Le théorème 6 suppose F *de dimension finie*, et ce n'est pas le cas : F contient la famille libre $(t, t^2, \ldots, t^k)$ pour tout k (chapitre 15). Ce n'est pas la dimension de E qui compte, mais celle du sous-espace sur lequel on projette.

À retenir. C'est le même phénomène que dans la remarque qui suit la proposition 11 et dans l'exercice 33 : en dimension infinie, un hyperplan peut avoir un orthogonal nul, et alors $(F^\perp)^\perp = E \neq F$. L'astuce de la fonction $t \mapsto tg(t)$, qui fabrique un élément de F « à partir » de g, resservira à l'exercice 40.

EXERCICE 40 ⋆⋆

Toutes les intégrales sont celles de polynômes sur $[0\,,1]$, calculées par les primitives usuelles, avec $\displaystyle\int_0^1 t^k\,\mathrm{d}t = \frac{1}{k+1}$.

1. *(a)* Le sous-espace $\text{Vect}(X, X^2)$ est de dimension 2, donc son orthogonal dans $\mathbb{R}_2[X]$ est de dimension 1 (proposition 11). Le polynôme $P = a + bX + cX^2$ lui appartient si et seulement si

$$\begin{cases} \langle P, X\rangle = \dfrac{a}{2} + \dfrac{b}{3} + \dfrac{c}{4} = 0 \\ \langle P, X^2\rangle = \dfrac{a}{3} + \dfrac{b}{4} + \dfrac{c}{5} = 0 \end{cases} \iff \begin{cases} 6a + 4b + 3c = 0 \\ 20a + 15b + 12c = 0 \end{cases} \iff b = -4a \text{ et } c = \frac{10a}{3}.$$

On choisit $a = 3$ pour éviter les fractions.

$$\boxed{\mathrm{Vect}(X, X^2)^\perp = \mathrm{Vect}(10X^2 - 12X + 3)}$$

(b) Pour $P \in \mathrm{Vect}(X, X^2)$, on a $P(0) = 0$: la condition impose $\langle A_2, X\rangle = \langle A_2, X^2\rangle = 0$, donc $A_2 = \lambda(10X^2 - 12X + 3)$ par (a). La condition pour $P = 1$ s'écrit

$$1 = \lambda \int_0^1 (10t^2 - 12t + 3)\,\mathrm{d}t = \lambda\Big(\frac{10}{3} - 6 + 3\Big) = \frac{\lambda}{3}, \qquad \text{d'où } \lambda = 3.$$

$$\boxed{A_2 = 30X^2 - 36X + 9}$$

Réciproquement, ce polynôme convient pour 1, X et X^2, donc pour tout $P \in \mathbb{R}_2[X]$ par linéarité. On contrôle : $\displaystyle\int_0^1 t(30t^2 - 36t + 9)\,\mathrm{d}t = \frac{15}{2} - 12 + \frac{9}{2} = 0$. ✓

2. L'espace $\mathbb{R}_n[X]$, muni de ce produit scalaire (proposition 2 (ii)), est euclidien, et $P \mapsto P(0)$ est une forme linéaire sur $\mathbb{R}_n[X]$. L'exercice 22 (question 2) fournit un unique $A_n \in \mathbb{R}_n[X]$ tel que $P(0) = \langle A_n, P\rangle$ pour tout $P \in \mathbb{R}_n[X]$. Pour le degré, on raisonne par l'absurde : si $n \geqslant 1$ et $\deg A_n \leqslant n - 1$, le polynôme $P = XA_n$ appartient à $\mathbb{R}_n[X]$, et

$$0 = P(0) = \langle A_n, XA_n\rangle = \int_0^1 t\,A_n(t)^2\,\mathrm{d}t.$$

La fonction $t \mapsto tA_n(t)^2$ est continue, positive, d'intégrale nulle : elle est nulle sur $[0\,,1]$ (chapitre 18, théorème 4), et A_n, qui a alors une infinité de racines, est le polynôme nul. Mais alors $1 = \langle A_n, 1\rangle = 0$ pour $P = 1$: c'est absurde. Pour $n = 0$, $A_0 = 1$, de degré 0.

$$\boxed{\exists!\, A_n \in \mathbb{R}_n[X],\ \forall\, P \in \mathbb{R}_n[X],\ P(0) = \int_0^1 A_n(t)P(t)\,\mathrm{d}t, \qquad \text{et } \deg A_n = n}$$

Un calcul symbolique, qui résout le système des $n + 1$ conditions obtenues pour $P = 1, X, \ldots, X^n$, donne

$$A_1 = 4 - 6X, \qquad A_2 = 30X^2 - 36X + 9, \qquad A_3 = -140X^3 + 240X^2 - 120X + 16. \qquad ✓$$

3. Non. Si un tel polynôme A existait, le raisonnement de la question 2, appliqué cette fois sans contrainte de degré à $P = XA \in \mathbb{R}[X]$, donnerait $\displaystyle\int_0^1 tA(t)^2\,\mathrm{d}t = 0$, donc $A = 0$, puis $1 = \langle A, 1\rangle = 0$ pour $P = 1$.

$$\boxed{\text{il n'existe aucun polynôme } A \in \mathbb{R}[X] \text{ tel que } P(0) = \int_0^1 A(t)P(t)\,\mathrm{d}t \text{ pour tout } P \in \mathbb{R}[X]}$$

À retenir. Sur l'espace $\mathbb{R}[X]$, de dimension infinie, la forme linéaire $P \mapsto P(0)$ n'est représentée par aucun vecteur : l'application Φ de l'exercice 22 n'y est pas surjective. Les A_n ne peuvent pas s'approcher d'un tel A, car leur norme tend vers l'infini. En effet, pour $P = (1 - X)^n$, on a $P(0) = 1$ et $\|P\|^2 = \displaystyle\int_0^1 (1 - t)^{2n}\,\mathrm{d}t = \frac{1}{2n + 1}$, et l'inégalité de Cauchy-Schwarz donne

$$1 = P(0) = \langle A_n, P\rangle \leqslant \|A_n\|\,\|P\| = \frac{\|A_n\|}{\sqrt{2n + 1}}, \qquad \text{donc} \qquad \|A_n\|^2 \geqslant 2n + 1.$$

Le calcul donne même $\|A_n\|^2 = \langle A_n, A_n\rangle = A_n(0) = (n + 1)^2$ pour $n \leqslant 5$.

EXERCICE 41 ⋆⋆

On rappelle que, p étant un projecteur ($p \circ p = p$), les sous-espaces $\operatorname{Ker} p$ et $\operatorname{Im} p$ sont supplémentaires dans E, et que p est le projecteur sur $\operatorname{Im} p$ parallèlement à $\operatorname{Ker} p$ (chapitre 14, théorème 5). Pour tout $x \in E$, on a donc la décomposition

$$x = p(x) + (x - p(x)), \qquad p(x) \in \operatorname{Im} p, \qquad x - p(x) \in \operatorname{Ker} p,$$

la seconde appartenance venant de $p(x - p(x)) = p(x) - p(p(x)) = 0$.

Sens direct. Supposons $\operatorname{Ker} p$ et $\operatorname{Im} p$ orthogonaux, et soit $x \in E$. Les deux vecteurs $p(x)$ et $x - p(x)$ sont orthogonaux, et le théorème de Pythagore (théorème 3) donne

$$\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 \geqslant \|p(x)\|^2, \qquad \text{d'où} \qquad \boxed{\|p(x)\| \leqslant \|x\|}$$

par croissance de la racine carrée. Il y a égalité si et seulement si $x - p(x) = 0$, c'est-à-dire si $x \in \operatorname{Im} p$.

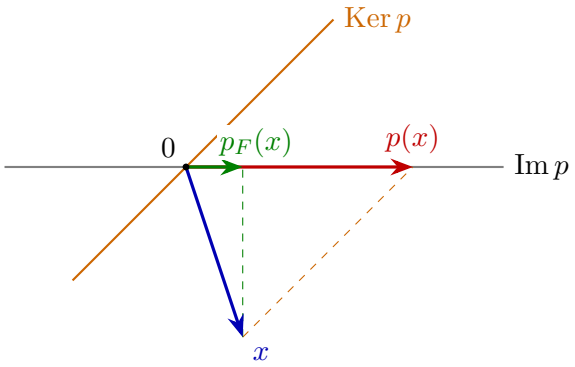
Réciproque. Supposons $\|p(z)\| \leqslant \|z\|$ pour tout $z \in E$, et soit $x \in \operatorname{Im} p$, $y \in \operatorname{Ker} p$ et $\lambda \in \mathbb{R}$. Comme $p(x) = x$ et $p(y) = 0$, on a $p(x + \lambda y) = x$, et l'hypothèse appliquée à $z = x + \lambda y$ donne, après élévation au carré et développement par la proposition 3 (ii),

$$\|x\|^2 \leqslant \|x + \lambda y\|^2 = \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2, \qquad \text{soit} \qquad 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2 \geqslant 0.$$

Pour $\lambda > 0$, on divise par λ : $2\langle x, y\rangle + \lambda\|y\|^2 \geqslant 0$, et l'on fait tendre λ vers 0 par valeurs positives, ce qui donne $\langle x, y\rangle \geqslant 0$. Pour $\lambda < 0$, la division renverse l'inégalité, et l'on obtient de même $\langle x, y\rangle \leqslant 0$. Ainsi $\langle x, y\rangle = 0$: c'est exactement la question 1 de l'exercice 13, puisque l'on a montré que $\|x + \lambda y\| \geqslant \|x\|$ pour tout réel λ. Tout vecteur de $\operatorname{Im} p$ est orthogonal à tout vecteur de $\operatorname{Ker} p$:

$$\boxed{p \text{ est un projecteur orthogonal } \iff \|p(x)\| \leqslant \|x\| \text{ pour tout } x \in E}$$

Piège. L'erreur la plus répandue est de croire qu'un projecteur, puisqu'il « écrase » l'espace sur un sous-espace, diminue toujours les longueurs. Un projecteur oblique peut au contraire les augmenter, comme le montre la figure ci-dessous.



Le projecteur p sur $F = \operatorname{Im} p = \operatorname{Vect}(e_1)$ parallèlement à $\operatorname{Ker} p = \operatorname{Vect}((1,1))$ s'écrit $p(a,b) = (a-b, 0)$. Pour $x = (\frac{1}{2}, -\frac{3}{2})$, on suit la direction de $\operatorname{Ker} p$ jusqu'à l'axe et l'on trouve $p(x) = (2, 0)$, de norme 2, alors que $\|x\| = \dfrac{\sqrt{10}}{2} \approx 1{,}58$. Le projeté orthogonal $p_F(x) = (\frac{1}{2}, 0)$, obtenu en descendant à angle droit, est lui plus court que x.

À retenir. Lorsque $F = \operatorname{Im} p$ est de dimension finie, un projecteur orthogonal sur F est exactement p_F : en effet $\operatorname{Ker} p \subset F^\perp$, et si $z \in F^\perp$ s'écrit $z = f + k$ avec $f \in F$ et $k \in \operatorname{Ker} p$, alors $\|f\|^2 = \langle z - k, f\rangle = 0$. La caractérisation obtenue ici, « ne jamais allonger », permet de parler de projecteur orthogonal sans supposer F de dimension finie ; les exercices 42 et 46 l'utilisent sous forme matricielle.

EXERCICE 42 ⋆⋆

On munit $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique $\langle X, Y\rangle = X^T Y$ (proposition 1).

1. Comme $A^2 = A$, l'endomorphisme $X \mapsto AX$ est un projecteur, d'image $\operatorname{Im} A$ et de noyau $\operatorname{Ker} A$.

Supposons A symétrique. Soit $X \in \operatorname{Ker} A$ et $Y = AZ \in \operatorname{Im} A$. Par la règle $(AX)^T = X^T A^T$ (chapitre 16, proposition 5),

$$\langle X, AZ\rangle = X^T AZ = X^T A^T Z = (AX)^T Z = 0,$$

donc $\operatorname{Ker} A$ et $\operatorname{Im} A$ sont orthogonaux.

Réciproquement, supposons $\operatorname{Ker} A$ et $\operatorname{Im} A$ orthogonaux. Pour toute colonne X, le vecteur $(I_n - A)X$ appartient à $\operatorname{Ker} A$, puisque $A(I_n - A) = A - A^2 = 0$. Pour toutes colonnes X et Y, on a donc

$$0 = \langle (I_n - A)X, AY\rangle = X^T (I_n - A)^T A\, Y.$$

En prenant pour X et Y les colonnes E_i et E_j de la base canonique, on lit le coefficient d'indice (i, j) de la matrice $(I_n - A)^T A$: elle est nulle. Ainsi $A = A^T A$, et cette matrice est symétrique, puisque $(A^T A)^T = A^T A$.

$$\boxed{\text{si } A^2 = A : \quad A \text{ est un projecteur orthogonal } \iff A^T = A}$$

2. Soit $F = \operatorname{Ker}(A - I_n)$ et $G = \operatorname{Ker}(A + I_n)$, supplémentaires par le théorème 6 du chapitre 14. On pose $P = \dfrac{1}{2}(A + I_n)$, de sorte que

$$P^2 = \frac{1}{4}(A^2 + 2A + I_n) = \frac{1}{4}(2A + 2I_n) = P.$$

C'est un projecteur, de noyau $\operatorname{Ker}(A + I_n) = G$ et d'image $\operatorname{Ker}(P - I_n) = \operatorname{Ker}\left(\frac{1}{2}(A - I_n)\right) = F$ (chapitre 14, théorème 5). La symétrie A est donc orthogonale si et seulement si le projecteur P l'est, c'est-à-dire, par la question 1, si et seulement si P est symétrique. Comme $A = 2P - I_n$ et que I_n est symétrique, cela équivaut à la symétrie de A.

$$\boxed{\text{si } A^2 = I_n : \quad A \text{ est une symétrie orthogonale } \iff A^T = A}$$

On contrôle sur deux exemples :

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La matrice A_1 vérifie $A_1^2 = I_2$ sans être symétrique, et la symétrie n'est pas orthogonale : $F = \operatorname{Vect}(E_1)$, $G = \operatorname{Vect}((-1, 2))$, et $\langle E_1, (-1, 2)\rangle = -1 \neq 0$ ✓. La matrice A_2, symétrique, est la symétrie orthogonale par rapport à la droite $\operatorname{Vect}((1, 1))$, parallèlement à la droite perpendiculaire $\operatorname{Vect}((1, -1))$.

Piège. Le mot « symétrique » a ici deux sens, qu'il ne faut pas confondre : une matrice symétrique vérifie $A^T = A$, une symétrie vérifie $A^2 = I_n$. L'exercice dit que, pour une symétrie, la propriété géométrique « être orthogonale » équivaut à la propriété matricielle « être symétrique ». Une symétrie orthogonale vérifie alors $A^T A = A^2 = I_n$: ses colonnes forment une base orthonormée, au sens de la remarque qui suit la proposition 9. Un projecteur orthogonal autre que I_n ne la vérifie pas, puisque $A^T A = A \neq I_n$.

EXERCICE 43 ⋆⋆

Soit $r = \dim F$, que l'on suppose non nul ($F = \{0\}$ donne une somme nulle), et $(f_1, \dots, f_r)$ une base orthonormée de F, qui existe par le théorème 5. La proposition 10 (iv) donne, pour chaque j,

$$\|p_F(e_j)\|^2 = \sum_{k=1}^{r} \langle e_j, f_k\rangle^2.$$

On somme sur j et l'on échange les deux sommes finies. Pour k fixé, les réels $\langle f_k, e_j\rangle$ sont les coordonnées de f_k dans la base orthonormée $(e_1, \dots, e_n)$, et la proposition 8 donne $\sum_j \langle f_k, e_j\rangle^2 = \|f_k\|^2 = 1$:

$$\sum_{j=1}^{n} \|p_F(e_j)\|^2 = \sum_{k=1}^{r} \sum_{j=1}^{n} \langle f_k, e_j\rangle^2 = \sum_{k=1}^{r} 1, \qquad \text{d'où} \qquad \boxed{\sum_{j=1}^{n} \|p_F(e_j)\|^2 = \dim F}$$

Le résultat ne dépend ni de la base orthonormée (e_j) choisie, ni de la position de F, mais seulement de sa dimension. Pour la droite $F = \mathrm{Vect}((1,1,1))$ de $\mathbb{R}^3$, chaque $p_F(e_j)$ vaut $\frac{1}{3}(1,1,1)$, de norme au carré $\frac{1}{3}$, et la somme vaut bien 1 ✓.

À retenir. On peut aussi passer par la trace. Comme $e_j - p_F(e_j)$ est orthogonal à $p_F(e_j)$, on a $\|p_F(e_j)\|^2 = \langle p_F(e_j), e_j\rangle$, qui est le coefficient diagonal d'indice j de la matrice de p_F dans la base orthonormée (e_j), par la proposition 8. La somme est donc $\mathrm{tr}(p_F)$, qui vaut le rang du projecteur p_F, c'est-à-dire $\dim F$ (chapitre 16, proposition 15).

EXERCICE 44 ⋆⋆

On munit $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique, et l'on identifie A à l'endomorphisme $X \mapsto AX$.

1. Soit $Y \in \mathcal{M}_{n,1}(\mathbb{R})$. Pour toute colonne X, on a $\langle AX, Y\rangle = X^T A^T Y = \langle X, A^T Y\rangle$. Si $A^T Y = 0$, le vecteur Y est donc orthogonal à tous les AX, c'est-à-dire à $\mathrm{Im}\, A$. Réciproquement, si $Y \in (\mathrm{Im}\, A)^\perp$, on choisit $X = A^T Y$, ce qui donne $\|A^T Y\|^2 = 0$, donc $A^T Y = 0$.

$$\boxed{\mathrm{Ker}\,(A^T) = (\mathrm{Im}\, A)^\perp}$$

2. Si $AX = 0$, alors $A^T AX = 0$: on a $\mathrm{Ker}\, A \subset \mathrm{Ker}\,(A^T A)$. Réciproquement, si $A^T AX = 0$, alors

$$0 = X^T A^T AX = (AX)^T (AX) = \|AX\|^2,$$

donc $AX = 0$. Ainsi

$$\boxed{\mathrm{Ker}\,(A^T A) = \mathrm{Ker}\, A}$$

On applique ce résultat à A^T, dont la transposée est A, et l'on obtient $\mathrm{Ker}\,(AA^T) = \mathrm{Ker}\,(A^T)$. Par le théorème du rang (chapitre 15, théorème 9) et l'égalité $\mathrm{rg}(A^T) = \mathrm{rg}(A)$ (chapitre 16, théorème 6 (iii)),

$$\mathrm{rg}(AA^T) = n - \dim \mathrm{Ker}\,(AA^T) = n - \dim \mathrm{Ker}\,(A^T) = \mathrm{rg}(A^T) = \mathrm{rg}(A).$$

Or $\mathrm{Im}\,(AA^T) \subset \mathrm{Im}\, A$, puisque $AA^T Y = A(A^T Y)$. Les deux sous-espaces ont la même dimension, et ils sont égaux (chapitre 15, théorème 5) : $\boxed{\mathrm{Im}\,(AA^T) = \mathrm{Im}\, A}$.

3. L'hypothèse $\mathrm{Im}\, A \subset \mathrm{Ker}\, A$ donne $A^2 = 0$. Soit X tel que $(A + A^T)X = 0$, c'est-à-dire $AX = -A^T X$. Alors

$$\|AX\|^2 = \langle AX, -A^T X\rangle = -X^T A^T A^T X = -X^T (A^2)^T X = 0.$$

Donc $AX = 0$, puis $A^T X = -AX = 0$. Le vecteur X appartient à $\mathrm{Ker}\, A \cap \mathrm{Ker}\,(A^T)$, c'est-à-dire, par l'hypothèse et la question 1, à $\mathrm{Im}\, A \cap (\mathrm{Im}\, A)^\perp$, qui est réduit à $\{0\}$ (proposition 5 (iv)). La matrice $A + A^T$ est donc injective, donc inversible (chapitre 16, théorème 3).

$$\boxed{\mathrm{Ker}\, A = \mathrm{Im}\, A \implies A + A^T \in GL_n(\mathbb{R})}$$

Pour la matrice suivante, on a $\operatorname{Ker} A = \operatorname{Im} A = \operatorname{Vect}(E_1)$, et $A + A^T$ est bien inversible ✓ :

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A + A^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Piège. L'hypothèse $A^2 = 0$ ne suffit pas : pour $n = 3$ et $A = E_{12}$, on a $A^2 = 0$, mais $A + A^T = E_{12} + E_{21}$ a une troisième colonne nulle. Ici $\operatorname{Im} A = \operatorname{Vect}(E_1)$ est strictement inclus dans $\operatorname{Ker} A = \operatorname{Vect}(E_1, E_3)$, et le vecteur E_3, qui appartient à $\operatorname{Ker} A \cap \operatorname{Ker}(A^T)$, n'est pas dans $\operatorname{Im} A$. Notons aussi que l'hypothèse impose, par le théorème du rang, $n = 2\operatorname{rg}(A)$: n est pair.

EXERCICE 45 ⋆⋆

On munit $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique.

1. L'ensemble $F = \operatorname{Im} A = \{AX \mid X \in \mathcal{M}_{p,1}(\mathbb{R})\}$ est un sous-espace de $\mathcal{M}_{n,1}(\mathbb{R})$, de dimension finie $\operatorname{rg} A = p$. Par le théorème 7, la borne inférieure des $\|Y - B\|$, pour $Y \in F$, est atteinte en un unique $Y_0 = p_F(B)$. Par le théorème du rang, $\dim \operatorname{Ker} A = p - \operatorname{rg} A = 0$: l'application $X \mapsto AX$ est injective, et il existe un unique X_0 tel que $AX_0 = Y_0$. Ce X_0 convient ; si X_1 convient aussi, AX_1 réalise le minimum sur F, donc $AX_1 = Y_0 = AX_0$, et $X_1 = X_0$ par injectivité.

$$\boxed{\text{il existe un unique } X_0 \text{ réalisant } \|AX_0 - B\| = d(B, \operatorname{Im} A)}$$

2. La démonstration de la question 1 de l'exercice 44 n'utilise pas que A est carrée : pour $A \in \mathcal{M}_{n,p}(\mathbb{R})$, on a encore $(\operatorname{Im} A)^\perp = \operatorname{Ker}(A^T)$. Par la proposition 10 (ii), $AX_0 = p_F(B)$ équivaut à $B - AX_0 \in F^\perp$, donc à

$$A^T(B - AX_0) = 0, \qquad \text{soit} \qquad \boxed{A^T A X_0 = A^T B}$$

Ce sont les équations normales de la méthode du cours. La matrice $A^T A \in \mathcal{M}_p(\mathbb{R})$ est inversible : la question 2 de l'exercice 44, dont la démonstration vaut aussi pour une matrice rectangulaire, donne $\operatorname{Ker}(A^T A) = \operatorname{Ker} A = \{0\}$. Le système a donc une unique solution, qui est X_0.

3. La quantité à minimiser est $\|AX - B\|^2$, avec

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Les deux colonnes de A ne sont pas colinéaires, donc $\operatorname{rg} A = 2$. On calcule

$$A^T A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}, \qquad A^T B = \begin{pmatrix} -3 \\ -1 \end{pmatrix}, \qquad \begin{cases} 6x + 2y = -3 \\ 2x + 3y = -1 \end{cases}$$

Le déterminant du système vaut 14, et l'on trouve $X_0 = (-\frac{1}{2}, 0)$. Le résidu $R = B - AX_0 = (\frac{3}{2}, \frac{1}{2}, -1)$ est orthogonal aux deux colonnes de A, puisque $\frac{3}{2} + \frac{1}{2} - 2 = 0$ et $\frac{3}{2} - \frac{1}{2} - 1 = 0$ ✓. Le minimum vaut $\|R\|^2$:

$$\boxed{\inf_{(x,y)\in\mathbb{R}^2} \left((x + y - 1)^2 + (x - y)^2 + (2x + y + 2)^2\right) = \frac{9}{4} + \frac{1}{4} + 1 = \frac{7}{2}, \quad \text{atteint en } \left(-\frac{1}{2}, 0\right)}$$

On retrouve la même valeur en remplaçant directement : $(-\frac{3}{2})^2 + (-\frac{1}{2})^2 + 1^2 = \frac{7}{2}$ ✓. Le système $AX = B$ de trois équations à deux inconnues n'a pas de solution ; X_0 est celle qui « s'en approche le plus » au sens des moindres carrés.

À retenir. La proposition 14 du cours est le cas particulier où A a pour colonnes (x_i) et $(1, \dots, 1)$, et où $X = (a, b)$. L'hypothèse « abscisses non toutes égales » y est exactement l'hypothèse $\operatorname{rg} A = 2$.

EXERCICE 46 ⋆⋆

1. Soit i fixé. On multiplie à gauche l'égalité $\sum_j A_j = I_n$ par ${A_i}^T$; tous les termes ${A_i}^T A_j$ avec $j \neq i$ sont nuls, et il reste

$$ {A_i}^T = {A_i}^T A_i. $$

Le second membre est une matrice symétrique, puisque $\left({A_i}^T A_i\right)^T = {A_i}^T A_i$. En transposant, on obtient $A_i = {A_i}^T A_i = {A_i}^T$: la matrice A_i est symétrique, et l'égalité devient $A_i = A_i^2$. Ainsi A_i est un projecteur, de matrice symétrique, donc un projecteur orthogonal par la question 1 de l'exercice 42.

$$ \boxed{A_i^2 = A_i = {A_i}^T : \ A_i \text{ est un projecteur orthogonal}} $$

2. La trace d'un projecteur est égale à son rang (chapitre 16, proposition 15), et la trace est linéaire :

$$ \sum_{k=1}^{p} \operatorname{rg}(A_k) = \sum_{k=1}^{p} \operatorname{tr}(A_k) = \operatorname{tr}\Big(\sum_{k=1}^{p} A_k\Big) = \operatorname{tr}(I_n), \qquad \text{d'où} \qquad \boxed{\sum_{k=1}^{p} \operatorname{rg}(A_k) = n} $$

Pour $n = 3$, $A_1 = \dfrac{1}{3}J$, où J est la matrice dont tous les coefficients valent 1, et $A_2 = I_3 - \dfrac{1}{3}J$, on a $J^2 = 3J$, donc ${A_1}^T A_2 = \dfrac{1}{3}J - \dfrac{1}{9}J^2 = 0$, et les rangs valent 1 et 2 ✓ : A_1 projette sur la droite $\operatorname{Vect}((1,1,1))$, A_2 sur le plan orthogonal.
Piège. Les transposées de l'hypothèse sont essentielles. Avec seulement $A_i A_j = 0$ pour $i \neq j$, on obtient encore des projecteurs, mais pas forcément orthogonaux : pour le projecteur oblique p de l'exercice 41 et $q = I_2 - p$, on a $pq = qp = 0$, mais

$$ p = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \qquad q = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \qquad p^T q = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \neq 0. $$

À retenir. L'hypothèse dit que $\langle A_i X, A_j Y\rangle = X^T {A_i}^T A_j Y = 0$: les images des A_k sont deux à deux orthogonales, et $X = \sum_k A_k X$ décompose tout vecteur sur elles. L'espace est ainsi coupé en sous-espaces deux à deux orthogonaux, et la question 2 en est la traduction dimensionnelle, comme le théorème 6 pour deux morceaux.

EXERCICE 47 ⋆⋆⋆

Si tous les u_i sont nuls, la matrice de Gram est nulle, la famille est liée, et les questions 1 à 3 sont immédiates ; on écarte ce cas.

1. Le sous-espace $F = \operatorname{Vect}(u_1, \dots, u_p)$ est de dimension finie $m \geqslant 1$. Muni de la restriction du produit scalaire, c'est un espace euclidien ; il possède une base orthonormée $\mathcal{B} = (e_1, \dots, e_m)$ par le théorème 5, et $\operatorname{Vect}(e_1, \dots, e_m) = F$ contient les u_j. Par la proposition 8 appliquée dans F, la coordonnée d'indice k de u_j est $\langle u_j, e_k\rangle$:

$$ \boxed{M = \left(\langle u_j, e_k\rangle\right)_{1 \leqslant k \leqslant m,\ 1 \leqslant j \leqslant p}} $$

Le coefficient d'indice (i,j) de $M^T M$ est le produit de la colonne i de M par sa colonne j, c'est-à-dire $\sum_k \langle u_i, e_k\rangle\langle u_j, e_k\rangle$, qui vaut $\langle u_i, u_j\rangle$ par la proposition 8 :

$$ \boxed{M^T M = \left(\langle u_i, u_j\rangle\right)_{1 \leqslant i,j \leqslant p}} $$

C'est la proposition 9 (i), dans le sous-espace F au lieu de l'espace entier.

2. On note $\Gamma = M^T M$. Pour une colonne $X = (x_j) \in \mathcal{M}_{p,1}(\mathbb{R})$, la colonne MX est celle des coordonnées dans $\mathcal{B}$ du vecteur $v = \sum_j x_j u_j$, et la proposition 8 donne

$$X^T \Gamma X = (MX)^T (MX) = \|v\|^2.$$

Si la famille est libre et si $\Gamma X = 0$, alors $\|v\|^2 = 0$, donc $v = 0$, et $X = 0$: la matrice Γ est injective, donc inversible, et $G(u_1, \dots, u_p) = \det \Gamma \neq 0$ (chapitre 17, théorème 6 (ii)). Si la famille est liée, il existe $X \neq 0$ tel que $v = 0$, donc $MX = 0$ et $\Gamma X = 0$: la matrice Γ n'est pas inversible, et son déterminant est nul.

$$\boxed{G(u_1, \dots, u_p) \neq 0 \iff (u_1, \dots, u_p) \text{ est libre}}$$

3. Si la famille est liée, $G = 0$. Si elle est libre, c'est une base de F, donc $m = p$: la matrice M est carrée, et le chapitre 17 (théorème 6 (i) et (iv)) donne

$$G(u_1, \dots, u_p) = \det(M^T) \det(M) = (\det M)^2, \qquad \text{d'où} \qquad \boxed{G(u_1, \dots, u_p) \geqslant 0}$$

avec inégalité stricte pour une famille libre. Pour $p = 2$, l'inégalité s'écrit $\|u_1\|^2 \|u_2\|^2 - \langle u_1, u_2 \rangle^2 \geqslant 0$, et la question 2 en donne le cas d'égalité : c'est l'inégalité de Cauchy-Schwarz (théorème 1), dont le déterminant de Gram est une généralisation.

4. On décompose $x = y + h$, avec $y = p_F(x) \in F$ et $h = x - p_F(x) \in F^\perp$, de sorte que $\|h\| = d(x, F)$ par le théorème 7. Comme h est orthogonal aux u_j, on a $\langle x, u_j \rangle = \langle y, u_j \rangle$ pour tout j, et, par Pythagore, $\langle x, x \rangle = \|y\|^2 + \|h\|^2$. Les matrices de Gram des familles $(x, u_1, \dots, u_p)$ et $(y, u_1, \dots, u_p)$ ne diffèrent donc que par leur coefficient d'indice $(1, 1)$, et

$$\big(\langle \cdot, \cdot \rangle\big)_{(x, u_1, \dots, u_p)} = \big(\langle \cdot, \cdot \rangle\big)_{(y, u_1, \dots, u_p)} + \|h\|^2 E_{11}.$$

Le déterminant est linéaire par rapport à la première colonne (chapitre 17, proposition 9 (i)). La matrice obtenue en remplaçant la première colonne par $\|h\|^2 E_1$ se développe selon cette colonne (chapitre 17, théorème 7) en $\|h\|^2 G(u_1, \dots, u_p)$, le mineur restant étant la matrice de Gram des u_j. D'autre part, $y \in F$, donc la famille $(y, u_1, \dots, u_p)$ est liée et son déterminant de Gram est nul par la question 2. Il reste

$$G(x, u_1, \dots, u_p) = 0 + d(x, F)^2 \, G(u_1, \dots, u_p), \qquad \text{d'où} \qquad \boxed{d(x, F)^2 = \frac{G(x, u_1, \dots, u_p)}{G(u_1, \dots, u_p)}}$$

puisque $G(u_1, \dots, u_p) \neq 0$ par la question 2. On a contrôlé la formule sur un exemple aléatoire dans $\mathbb{R}^4$, avec $p = 2$ ✓.

5. L'exercice 35 (a) demande la borne inférieure de $\displaystyle\int_0^1 \big(t - a\sqrt{t} - b\big)^2 \, \mathrm{d}t = \|t - a\sqrt{t} - b\|^2$ dans $\mathcal{C}([0\,,1], \mathbb{R})$, muni de $\langle f, g \rangle = \displaystyle\int_0^1 fg$: c'est $d(x, F)^2$, avec $x = t$ et $F = \mathrm{Vect}(\sqrt{t}, 1)$. Les produits scalaires sont des intégrales de fonctions puissances, que l'on calcule par leurs primitives (chapitre 6) : $\displaystyle\int_0^1 t^\alpha \, \mathrm{d}t = \frac{1}{\alpha + 1}$ pour $\alpha \geqslant 0$. On trouve $\|\sqrt{t}\|^2 = \frac{1}{2}$, $\langle \sqrt{t}, 1 \rangle = \frac{2}{3}$, $\|1\|^2 = 1$, $\|t\|^2 = \frac{1}{3}$, $\langle t, \sqrt{t} \rangle = \frac{2}{5}$ et $\langle t, 1 \rangle = \frac{1}{2}$, d'où

$$G(\sqrt{t}, 1) = \begin{vmatrix} \frac{1}{2} & \frac{2}{3} \\ \frac{2}{3} & 1 \end{vmatrix} = \frac{1}{18}, \qquad G(t, \sqrt{t}, 1) = \begin{vmatrix} \frac{1}{3} & \frac{2}{5} & \frac{1}{2} \\ \frac{2}{5} & \frac{1}{2} & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 1 \end{vmatrix} = \frac{1}{5400}.$$

La famille $(\sqrt{t}, 1)$ est libre, puisque $G(\sqrt{t}, 1) \neq 0$, et la question 4 donne

$$\boxed{\inf_{(a,b) \in \mathbb{R}^2} \int_0^1 \big(t - a\sqrt{t} - b\big)^2 \, \mathrm{d}t = \frac{1/5400}{1/18} = \frac{1}{300}}$$

C'est la valeur trouvée à l'exercice 35, sans avoir à résoudre les équations normales ✓. En revanche, la formule ne dit pas en quel point le minimum est atteint ; les équations normales donnent $a = \dfrac{6}{5}$ et $b = -\dfrac{3}{10}$.

À retenir. Le chapitre 17 interprétait un déterminant comme un volume, et la remarque qui suit la proposition 9 en déduit que $\sqrt{G(u_1, \dots, u_p)}$ est le volume du parallélotope construit sur les u_j. La formule de la question 4 s'écrit alors « volume = base × hauteur » : le volume construit sur $(x, u_1, \dots, u_p)$ est celui de la base construite sur les u_j, multiplié par la hauteur $d(x, F)$.

EXERCICE 48 ⋆⋆⋆

1. L'application est symétrique, et linéaire par rapport à f, par linéarité de la dérivation et de l'intégrale. Pour $f \in E$, la fonction $f^2 + f'^2$ est continue et positive, donc $\langle f, f\rangle \geqslant 0$. Si $\langle f, f\rangle = 0$, le théorème 4 du chapitre 18 montre que $f^2 + f'^2$ est nulle sur $[0\,,1]$, donc $f = 0$.

$$\boxed{\langle f, g\rangle = \int_0^1 (fg + f'g') \text{ est un produit scalaire sur } \mathcal{C}^1([0\,,1], \mathbb{R})}$$

2. L'équation $y'' - y = 0$ est linéaire, homogène, à coefficients constants, d'équation caractéristique $r^2 - 1 = 0$, de racines ± 1 (chapitre 6, théorème 5) : ses solutions sont les fonctions $\lambda \mathrm{e}^t + \mu \mathrm{e}^{-t}$, c'est-à-dire les combinaisons linéaires de ch et sh. Ainsi $F = \mathrm{Vect}(\mathrm{ch}, \mathrm{sh})$ est un sous-espace de E, et la famille $(\mathrm{ch}, \mathrm{sh})$ est libre : si $a\,\mathrm{ch} + b\,\mathrm{sh} = 0$, l'évaluation en 0 donne $a = 0$, puis celle en 1 donne $b = 0$. Donc $\boxed{\dim F = 2}$, ce que donnait aussi la proposition 1 (ii) du chapitre 15.

Soit $f \in E$ et $g \in F$. On intègre par parties $\displaystyle\int_0^1 f'g'$, en primitivant f' et en dérivant g', qui est de classe $\mathcal{C}^1$ puisque $g'' = g$ (chapitre 6, proposition 4) :

$$\int_0^1 f'g' = [fg']_0^1 - \int_0^1 fg'' = f(1)g'(1) - f(0)g'(0) - \int_0^1 fg, \qquad \text{d'où} \qquad \boxed{\langle f, g\rangle = f(1)g'(1) - f(0)g'(0)}$$

Si $f(0) = f(1) = 0$, ce produit est nul pour tout $g \in F$, et $f \in F^\perp$. Réciproquement, si $f \in F^\perp$, le choix $g = \mathrm{ch}$, avec $g' = \mathrm{sh}$ et $\mathrm{sh}\,0 = 0$, donne $f(1)\,\mathrm{sh}\,1 = 0$, donc $f(1) = 0$; le choix $g = \mathrm{sh}$, avec $g' = \mathrm{ch}$, donne alors $-f(0) = 0$.

$$\boxed{F^\perp = \{f \in E \mid f(0) = f(1) = 0\}}$$

3. Par la proposition 10 (ii), $p_F(f)$ est l'unique $g \in F$ tel que $f - g \in F^\perp$, c'est-à-dire tel que $g(0) = f(0)$ et $g(1) = f(1)$. La fonction $t \mapsto \mathrm{sh}(1-t) = \mathrm{sh}\,1\,\mathrm{ch}\,t - \mathrm{ch}\,1\,\mathrm{sh}\,t$ appartient à F, s'annule en 1 et vaut $\mathrm{sh}\,1$ en 0 ; la fonction sh s'annule en 0 et vaut $\mathrm{sh}\,1$ en 1. On en déduit

$$\boxed{p_F(f)(t) = \frac{f(0)\ \mathrm{sh}(1-t) + f(1)\ \mathrm{sh}\,t}{\mathrm{sh}\,1}}$$

La projection orthogonale sur F remplace f par l'unique solution de $y'' = y$ qui a les mêmes valeurs aux deux bords.

4. Soit $g_{\alpha,\beta} = \dfrac{\alpha\ \mathrm{sh}(1-t) + \beta\ \mathrm{sh}\,t}{\mathrm{sh}\,1}$, qui appartient à $F \cap E_{\alpha,\beta}$. Pour toute $f \in E_{\alpha,\beta}$, la question 3 donne $p_F(f) = g_{\alpha,\beta}$, et le théorème de Pythagore, appliqué à $g_{\alpha,\beta} \in F$ et $f - g_{\alpha,\beta} \in F^\perp$, donne

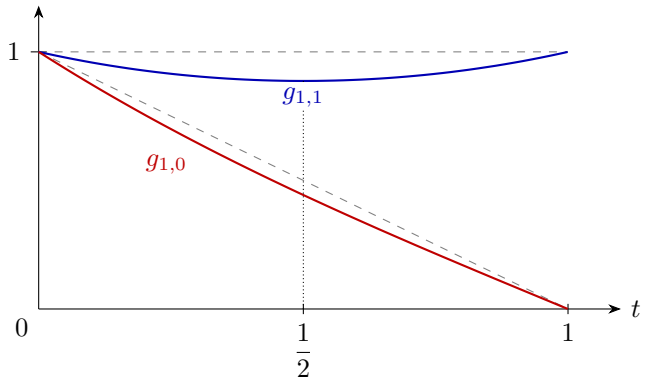
$$\int_0^1 (f^2 + f'^2) = \|f\|^2 = \|g_{\alpha,\beta}\|^2 + \|f - g_{\alpha,\beta}\|^2 \geqslant \|g_{\alpha,\beta}\|^2,$$

avec égalité si et seulement si $f = g_{\alpha,\beta}$. La borne inférieure est donc un minimum, atteint en la seule fonction $g_{\alpha,\beta}$. Pour le calculer, on applique la formule de la question 2 avec $f = g = g_{\alpha,\beta}$, après avoir dérivé :

$$g'_{\alpha,\beta}(t) = \frac{-\alpha\ \mathrm{ch}(1-t) + \beta\ \mathrm{ch}\,t}{\mathrm{sh}\,1}, \qquad g'_{\alpha,\beta}(0) = \frac{\beta - \alpha\,\mathrm{ch}\,1}{\mathrm{sh}\,1}, \qquad g'_{\alpha,\beta}(1) = \frac{\beta\,\mathrm{ch}\,1 - \alpha}{\mathrm{sh}\,1},$$

$$\|g_{\alpha,\beta}\|^2 = \beta\,\frac{\beta\operatorname{ch}1-\alpha}{\operatorname{sh}1} - \alpha\,\frac{\beta-\alpha\operatorname{ch}1}{\operatorname{sh}1}, \qquad \boxed{\inf_{f\in E_{\alpha,\beta}} \int_0^1 (f^2+f'^2) = \frac{(\alpha^2+\beta^2)\operatorname{ch}1 - 2\alpha\beta}{\operatorname{sh}1}}$$

atteinte seulement en $g_{\alpha,\beta}$. On a contrôlé cette valeur par un calcul direct de l'intégrale pour quatre couples (α,β) ✓.



Pour $\alpha=\beta=1$, en bleu, la fonction optimale est $g_{1,1}(t) = \dfrac{\operatorname{ch}(t-1/2)}{\operatorname{ch}(1/2)}$: elle descend jusqu'à $\dfrac{1}{\operatorname{ch}(1/2)} \approx 0{,}887$ au milieu, et le minimum vaut $\dfrac{2(\operatorname{ch}1-1)}{\operatorname{sh}1} = 2\operatorname{th}\dfrac{1}{2} \approx 0{,}924$, plus petit que la valeur 1 obtenue pour la fonction constante, en gris. En descendant un peu au milieu, elle gagne plus sur $\int f^2$ qu'elle ne perd sur $\int f'^2$, qui n'est plus nul mais reste petit. Pour $\alpha=1$ et $\beta=0$, en rouge, la fonction optimale $g_{1,0}(t) = \dfrac{\operatorname{sh}(1-t)}{\operatorname{sh}1}$ est très proche de la droite grise $1-t$, et les deux valeurs le sont aussi : $\dfrac{\operatorname{ch}1}{\operatorname{sh}1} \approx 1{,}313$ contre $\dfrac{4}{3} \approx 1{,}333$.

À retenir. L'ensemble $E_{\alpha,\beta} = g_{\alpha,\beta} + F^\perp$ est un sous-espace affine de direction $F^\perp$ (chapitre 14) ; son élément de plus petite norme est celui qui est orthogonal à la direction. C'est la même idée que la distance à un sous-espace, appliquée à 0 et à un sous-espace affine. Le produit scalaire de la question 1 n'est pas celui de la proposition 2 : il mesure aussi la dérivée, et c'est pourquoi la fonction optimale est une solution d'équation différentielle, et non une fonction affine.

Exercices de synthèse

EXERCICE 49 ⋆⋆

1. Faux. Dans $\mathbb{R}^2$ muni du produit canonique, les vecteurs $x=(1,0)$, $y=(0,1)$ et $z=(1,-1)$ ont pour somme $(2,0)$, et $\|(2,0)\|^2 = 4 = 1+1+2$, alors que $\langle x,z\rangle = 1 \neq 0$. C'est l'exemple de la remarque qui suit le théorème 3 : la somme des doubles produits, ici $0+2-2$, est nulle sans que chacun le soit. La réciproque du théorème de Pythagore ne vaut que pour deux vecteurs.

2. Vrai. Par bilinéarité et symétrie, $\langle x+y, x-y\rangle = \|x\|^2 - \|y\|^2 = 0$. Géométriquement, les diagonales d'un losange sont perpendiculaires.

3. Faux. La formule est bilinéaire, symétrique et positive, mais pas définie : l'indicatrice f de $\{0\}$ est continue par morceaux, non nulle, et $\displaystyle\int_0^1 f^2 = 0$ (chapitre 18, proposition 6). C'est la première remarque qui suit la proposition 2.

4. Vrai. L'égalité entraîne $|\langle x,y\rangle| = \|x\|\,\|y\|$, cas d'égalité de l'inégalité de Cauchy-Schwarz (théorème 1) : la famille (x,y) est liée, et comme $x\neq 0$, il existe $\lambda\in\mathbb{R}$ tel que $y=\lambda x$. Alors $\lambda\|x\|^2 = \langle x,\lambda x\rangle = \|x\|\,\|\lambda x\| = |\lambda|\|x\|^2$, d'où $\lambda = |\lambda| \geqslant 0$. C'est l'argument du cas d'égalité de l'inégalité triangulaire (théorème 2).

5. Vrai. Pour deux réels a, b, on a $ab \leqslant \dfrac{a^2+b^2}{2}$, puisque $(a-b)^2 \geqslant 0$. Avec $a = \|x+y\|$, $b = \|x-y\|$ et l'identité du parallélogramme (proposition 3 (iv)),

$$\|x+y\|\,\|x-y\| \leqslant \frac{\|x+y\|^2+\|x-y\|^2}{2} = \|x\|^2+\|y\|^2.$$

Il y a égalité si et seulement si $\|x+y\| = \|x-y\|$, c'est-à-dire si $x \perp y$: parmi les parallélogrammes de côtés donnés, le rectangle est celui dont le produit des diagonales est maximal.

6. Faux. Dans $\mathbb{R}^2$, soit $e_1 = (1,0)$ et $e_2 = (1,1)$. L'algorithme appliqué à (e_1, e_2) donne $u_1 = (1,0)$, puis $v_2 = e_2 - \langle e_2, u_1\rangle u_1 = (0,1)$ et $u_2 = (0,1)$. Appliqué à (e_2, e_1), il donne $u_1' = \dfrac{1}{\sqrt{2}}(1,1)$, puis $v_2' = e_1 - \dfrac{1}{\sqrt{2}}u_1' = \dfrac{1}{2}(1,-1)$ et $u_2' = \dfrac{1}{\sqrt{2}}(1,-1)$. Les deux bases orthonormées sont différentes, même à l'ordre près : le premier vecteur fourni par l'algorithme est toujours le premier vecteur de départ normalisé, et la famille obtenue dépend de l'ordre.

7. Vrai. Soit $x \in E$. Le vecteur $x - p_G(x)$ appartient à $G^\perp$, qui est inclus dans $F^\perp$ puisque $F \subset G$ (proposition 5 (ii)). Comme $F^\perp = \operatorname{Ker} p_F$, on a

$$p_F(x) = p_F(p_G(x)) + p_F(x - p_G(x)) = p_F(p_G(x)).$$

Projeter d'abord sur le grand sous-espace, puis sur le petit, revient à projeter directement sur le petit. On a aussi $p_G \circ p_F = p_F$, puisque $p_F(x) \in F \subset G$ est laissé fixe par p_G.

8. Faux. Dans $\mathbb{R}[X]$ muni du produit scalaire $\displaystyle\sum_k a_k b_k$ sur les coefficients, l'hyperplan H des polynômes P tels que $P(1) = 0$ vérifie $H^\perp = \{0\}$. Son double orthogonal est donc $\mathbb{R}[X]$ tout entier, qui n'est pas égal à H : c'est la remarque qui suit la proposition 11. L'exercice 39 en donne un autre exemple dans $\mathcal{C}([0\,,1],\mathbb{R})$. L'égalité est vraie si F est de dimension finie (proposition 10 (iii)), en particulier si E est euclidien.

9. Vrai. Soit H un hyperplan de E, de dimension n. Par la proposition 11, $\dim H^\perp = n-(n-1) = 1$. Deux vecteurs non nuls orthogonaux à H appartiennent à la droite $H^\perp$, donc sont colinéaires : c'est l'unicité du vecteur normal, à un facteur près, de la proposition 12.

10. Vrai. Soit A la famille génératrice. Par la proposition 5 (ii) et (iii),

$$x \in A^\perp = (\operatorname{Vect} A)^\perp = E^\perp = \{0\}.$$

Le caractère éventuellement infini de la famille ne gêne pas, puisque $\operatorname{Vect} A$ est formé de combinaisons linéaires finies.

Vrai : $2, 4, 5, 7, 9, 10$;	Faux : $1, 3, 6, 8$

Piège. Les assertions 1 et 8 sont fausses pour la même raison de fond : une propriété vraie dans un cadre restreint, deux vecteurs pour Pythagore, la dimension finie pour le double orthogonal, ne s'étend pas d'elle-même. L'assertion 3 rappelle que le caractère défini dépend de l'espace : la même formule est un produit scalaire sur les fonctions continues, et n'en est pas un sur les fonctions continues par morceaux.

EXERCICE 50 ⋆⋆

On note Σ la somme étendue aux 2^n éléments $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ de $\{-1,1\}^n$.

1. On développe chaque carré par bilinéarité, puis on échange les sommes finies :

$$\sum_\varepsilon \Big\| \sum_{i=1}^n \varepsilon_i u_i \Big\|^2 = \sum_\varepsilon \sum_{i,j} \varepsilon_i \varepsilon_j \langle u_i, u_j\rangle = \sum_{i,j} \langle u_i, u_j\rangle \sum_\varepsilon \varepsilon_i \varepsilon_j.$$

Pour $i = j$, on a $\varepsilon_i^2 = 1$, et la somme intérieure vaut 2^n. Pour $i \neq j$, l'application qui change le signe de ε_i sans toucher aux autres coordonnées est une bijection de $\{-1,1\}^n$ sur lui-même, égale à sa propre

réciproque, qui change $\varepsilon_i\varepsilon_j$ en son opposé : la somme $\sum_\varepsilon \varepsilon_i\varepsilon_j$ est égale à son opposée, donc nulle. Les doubles produits disparaissent, et

$$\boxed{\sum_{\varepsilon\in\{-1,1\}^n} \Big\|\sum_{i=1}^n \varepsilon_i u_i\Big\|^2 = 2^n \sum_{i=1}^n \|u_i\|^2}$$

Pour $n=2$, les termes associés à ε et $-\varepsilon$ sont égaux, et l'on retrouve deux fois l'identité du parallélogramme (proposition 3 (iv)) : $2(\|u_1+u_2\|^2+\|u_1-u_2\|^2) = 4(\|u_1\|^2+\|u_2\|^2)$.

2. Chacun des 2^n termes est au plus C^2, puisque $0 \leqslant \|\sum_i \varepsilon_i u_i\| \leqslant C$. Leur somme est donc au plus $2^n C^2$, et la question 1 donne, après division par 2^n,

$$\boxed{\sum_{i=1}^n \|u_i\|^2 \leqslant C^2}$$

Pour des vecteurs unitaires, on obtient $n \leqslant C^2$: si toutes les sommes signées de n vecteurs unitaires sont de norme au plus C, il y a au plus C^2 vecteurs. La borne est atteinte pour la base canonique de $\mathbb{R}^n$, dont toutes les sommes signées ont pour norme $\sqrt{n}$.

3. Soit $m = \sum_i \|u_i\|^2$. Par la question 1, m est la moyenne des 2^n réels $a_\varepsilon = \|\sum_i \varepsilon_i u_i\|^2$. Si tous ces réels étaient strictement inférieurs à m, leur moyenne le serait aussi ; il existe donc ε tel que $a_\varepsilon \geqslant m$. De même, ils ne peuvent pas être tous strictement supérieurs à m.

$$\boxed{\exists\,\varepsilon,\ \varepsilon' \in \{-1,1\}^n, \qquad \Big\|\sum_{i=1}^n \varepsilon_i u_i\Big\|^2 \geqslant \sum_{i=1}^n \|u_i\|^2 \geqslant \Big\|\sum_{i=1}^n \varepsilon'_i u_i\Big\|^2}$$

Si les u_i sont deux à deux orthogonaux, les $\varepsilon_i u_i$ le sont aussi, et le théorème de Pythagore (théorème 3 (ii)) donne $a_\varepsilon = \sum_i \|\varepsilon_i u_i\|^2 = m$ pour tout ε : les 2^n réels sont tous égaux à leur moyenne. Pour $u_1=(1,0)$ et $u_2=(1,1)$, non orthogonaux, on trouve au contraire les valeurs 5, 1, 1, 5, de moyenne $3 = 1+2$ ✓.

À retenir. L'argument « une moyenne est comprise entre le plus petit et le plus grand terme » démontre l'existence d'un choix de signes sans le construire. Le chapitre 21 relira ce calcul en termes de probabilités, en tirant les signes au hasard.

EXERCICE 51 ⋆⋆⋆

On munit $\mathbb{R}^n$ du produit scalaire canonique, et l'on note $N = n+1$ le nombre de vecteurs.

1. Soit $M \in \mathcal{M}_{n,n+1}(\mathbb{R})$ la matrice dont les colonnes sont les coordonnées des e_j dans la base canonique, qui est orthonormée. La proposition 9 (i) donne $A = M^T M$. Comme $n+1$ vecteurs de $\mathbb{R}^n$ forment une famille liée (chapitre 15, théorème 4), il existe une colonne $X \neq 0$ telle que $MX = 0$; alors $AX = M^T MX = 0$, et A n'est pas inversible.

$$\boxed{\det A = 0}$$

C'est aussi la question 2 de l'exercice 47 : la famille $(e_1,\dots,e_{n+1})$ est liée, donc son déterminant de Gram est nul.

2. La matrice A a ses coefficients diagonaux égaux à 1 et les autres égaux à α. On remplace la première colonne par la somme de toutes les colonnes, ce qui ne change pas le déterminant (chapitre 17, proposition 10) : tous ses coefficients valent $1+n\alpha$, que l'on met en facteur. On retranche ensuite la première ligne à chacune des autres : pour $i \geqslant 2$, la ligne i devient $(0,\dots,0,1-\alpha,0,\dots,0)$, le coefficient $1-\alpha$ étant en position i, puisque $a_{ij}-a_{1j}$ vaut $1-\alpha$ si $j=i$ et 0 sinon. La matrice obtenue est triangulaire supérieure, de diagonale $(1,1-\alpha,\dots,1-\alpha)$:

$$\boxed{\det A = (1+n\alpha)(1-\alpha)^n}$$

Ce résultat a été vérifié pour $1 \leqslant n \leqslant 6$ ✓. Par la question 1, $\boxed{\alpha = 1 \text{ ou } \alpha = -\frac{1}{n}}$.

— Si $\alpha = 1$, on a $\|e_i - e_j\|^2 = 2 - 2\alpha = 0$: tous les vecteurs sont égaux, et $e_1 + \cdots + e_{n+1} = (n+1)e_1$.

— Si $\alpha = -\frac{1}{n}$, on développe :

$$\Big\| \sum_{i=1}^{n+1} e_i \Big\|^2 = \sum_i \|e_i\|^2 + \sum_{i \neq j} \langle e_i, e_j \rangle = (n+1) - (n+1)n \cdot \frac{1}{n} = 0, \qquad \text{d'où} \qquad \boxed{e_1 + \cdots + e_{n+1} = 0}$$

Les $n+1$ vecteurs sont les sommets d'un simplexe régulier centré en 0 : un triangle équilatéral pour $n = 2$, avec $\alpha = -\frac{1}{2}$, c'est-à-dire des angles de 120°, et un tétraèdre régulier pour $n = 3$.

3. On commence par l'hyperplan H d'équation $y_0 + \cdots + y_n = 0$ de $\mathbb{R}^{n+1}$, de dimension n, de vecteur normal $w = (1, \ldots, 1)$, avec $\|w\|^2 = n+1$. Soit $(f_0, \ldots, f_n)$ la base canonique de $\mathbb{R}^{n+1}$, et $g_i = p_H(f_i) = f_i - \frac{1}{n+1} w$ (proposition 13, avec $\langle f_i, w \rangle = 1$). En développant,

$$\langle g_i, g_j \rangle = \delta_{ij} - \frac{2}{n+1} + \frac{n+1}{(n+1)^2} = \delta_{ij} - \frac{1}{n+1}, \qquad \text{donc} \qquad \|g_i\|^2 = \frac{n}{n+1}.$$

Les vecteurs $x_i = \sqrt{\frac{n+1}{n}}\, g_i$ de H sont unitaires, et $\langle x_i, x_j \rangle = \frac{n+1}{n} \cdot \big(-\frac{1}{n+1} \big) = -\frac{1}{n}$ pour $i \neq j$.

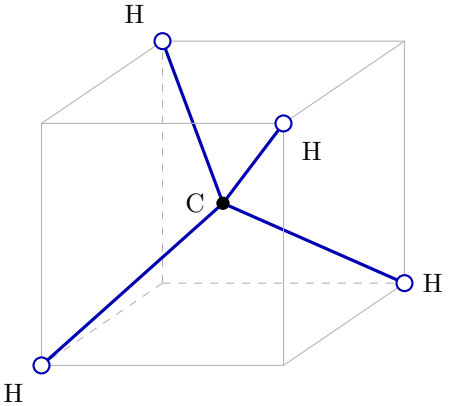
Soit maintenant E euclidien de dimension n, $(h_1, \ldots, h_n)$ une base orthonormée de H et $(\eta_1, \ldots, \eta_n)$ une base orthonormée de E, qui existent par le théorème 5. Soit $\varphi : H \to E$ l'application linéaire qui envoie h_k sur η_k (chapitre 14, théorème 4). Un vecteur de H et son image ont les mêmes coordonnées, l'un dans (h_k), l'autre dans (η_k), et la proposition 8 calcule leurs produits scalaires par la même formule $\sum_k y_k z_k$: φ conserve le produit scalaire. Les vecteurs $\varphi(x_0)$, …, $\varphi(x_n)$ conviennent.

$$\boxed{\text{il existe } x_0, \ldots, x_n \in E \text{ unitaires tels que } \langle x_i, x_j \rangle = -\frac{1}{n} \ (i \neq j)}$$

4. On place l'atome de carbone à l'origine de l'espace $\mathbb{R}^3$, et l'on note e_1, …, e_4 les vecteurs unitaires dirigés selon les quatre liaisons ; ce choix d'unité ne change pas les angles. Par hypothèse, $\langle e_i, e_j \rangle = \cos\theta$ pour $i \neq j$, où θ est l'angle commun. On est dans la situation de la question 2 avec $n = 3$: $\cos\theta = 1$ est exclu, puisque les quatre liaisons seraient confondues, et il reste $\cos\theta = -\frac{1}{3}$.

$$\boxed{\theta = \arccos\Big(-\frac{1}{3} \Big) \approx 109{,}47^\circ}$$

La question 2 montre de plus que $e_1 + e_2 + e_3 + e_4 = 0$: le carbone est au centre de gravité des quatre hydrogènes, qui forment un tétraèdre régulier.



Les quatre hydrogènes occupent un sommet sur deux du cube $[-1\,,1]^3$, aux points $(1,1,1)$, $(1,-1,-1)$, $(-1,1,-1)$ et $(-1,-1,1)$: ces vecteurs ont pour norme $\sqrt{3}$ et leurs produits scalaires deux à deux valent -1, d'où un cosinus égal à $-\dfrac{1}{3}$, comme le prévoit la question 2.

À retenir. Le résultat n'utilise que la géométrie : quatre directions issues d'un point, faisant deux à deux le même angle, forment nécessairement cet angle de 109,47° environ, appelé angle tétraédrique. L'hypothèse de modélisation est l'égalité des angles ; le déterminant de Gram fait le reste.

EXERCICE 52 ⋆⋆⋆

On munit $\mathbb{R}[X]$ du produit scalaire défini par $\langle P,Q\rangle = \displaystyle\int_0^1 PQ$ (proposition 2 (ii)). On a alors $I(P) = \|X^n - P\|^2$.

1. Le sous-espace $\mathbb{R}_{n-1}[X]$ est de dimension finie n. Par le théorème 7, la distance de X^n à $\mathbb{R}_{n-1}[X]$ est atteinte en un unique point, le projeté orthogonal :

$$\boxed{I \text{ atteint son minimum en un unique } P = p_{\mathbb{R}_{n-1}[X]}(X^n)}$$

2.(a) Pour $i \in [\![0\,;n-1]\!]$, les dénominateurs $i+j+1$ sont non nuls, et pour tout entier $k \geqslant 0$,

$$\frac{1}{i+k+1} = \int_0^1 t^{i+k}\,\mathrm{d}t = \langle X^i, X^k\rangle.$$

Donc

$$F(i) = \langle X^i, X^n\rangle - \sum_{j=0}^{n-1} a_j \langle X^i, X^j\rangle = \langle X^i, X^n - P\rangle = 0,$$

puisque $X^n - P$ est orthogonal à $\mathbb{R}_{n-1}[X]$ (proposition 10 (ii)) : $\boxed{F(i) = 0 \text{ pour } 0 \leqslant i \leqslant n-1}$.

(b) Soit $D = (X+1)(X+2)\cdots(X+n+1)$, de degré $n+1$. En réduisant au même dénominateur, $F = \dfrac{N}{D}$, où N est un polynôme de degré au plus n : chaque terme $\dfrac{1}{X+k}$ devient $\dfrac{D/(X+k)}{D}$, et $\dfrac{D}{X+k}$ est un produit de n facteurs de degré 1. Pour $i \in [\![0\,;n-1]\!]$, on a $D(i) \neq 0$ et $N(i) = D(i)F(i) = 0$. Le polynôme N a donc les n racines distinctes 0, 1, …, $n-1$; il est divisible par leur produit (chapitre 13, proposition 8), et, son degré étant au plus n, il existe une constante c telle que

$$N = c\,X(X-1)\cdots(X-n+1).$$

On détermine c par le pôle $-(n+1)$: la fraction $(X+n+1)F = 1 - \displaystyle\sum_j a_j \frac{X+n+1}{X+j+1}$ prend la valeur 1 en $-(n+1)$. D'autre part, $(X+n+1)F = \dfrac{N}{(X+1)\cdots(X+n)}$, dont la valeur en $-(n+1)$ est

$$c\,\frac{\prod_{i=0}^{n-1}\big(-(n+1+i)\big)}{\prod_{k=1}^{n}(k-n-1)} = c\,\frac{(-1)^n(n+1)(n+2)\cdots(2n)}{(-1)^n\,n!} = c\,\frac{(2n)!}{(n!)^2} = c\binom{2n}{n}.$$

Ainsi $c = \dbinom{2n}{n}^{-1}$, et

$$\boxed{F = \frac{X(X-1)\cdots(X-n+1)}{\binom{2n}{n}(X+1)(X+2)\cdots(X+n+1)}}$$

(c) En $X = n$, le numérateur vaut $n!$, et le produit $(n+1)(n+2)\cdots(2n+1)$ vaut $\dfrac{(2n+1)!}{n!}$, d'où

$$F(n) = \frac{(n!)^2}{\binom{2n}{n}(2n+1)!} = \frac{1}{\binom{2n}{n}(2n+1)\binom{2n}{n}}.$$

D'autre part, le calcul de la question (a), fait avec $i = n$, donne $F(n) = \langle X^n, X^n - P\rangle$. Comme P est orthogonal à $X^n - P$, on peut retrancher $\langle P, X^n - P\rangle = 0$:

$$F(n) = \langle X^n - P, X^n - P\rangle = I(P), \qquad \text{d'où} \qquad \boxed{\min_{P\in\mathbb{R}_{n-1}[X]} \int_0^1 \left(t^n - P(t)\right)^2 \mathrm{d}t = \frac{1}{(2n+1)\binom{2n}{n}^2}}$$

Pour $n = 1$, on trouve $\dfrac{1}{3\cdot 4} = \dfrac{1}{12}$, qui est bien $\displaystyle\int_0^1 (t - \frac{1}{2})^2\,\mathrm{d}t$ ✓ ; pour $n = 3$, $\dfrac{1}{7\cdot 400} = \dfrac{1}{2800}$. Toutes ces valeurs ont été contrôlées par la résolution directe des équations normales, pour $1 \leqslant n \leqslant 6$ ✓.

3. Pour $n = 2$, la formule donne $\dfrac{1}{5\cdot 36} = \dfrac{1}{180}$, qui est bien le carré de la distance de t^2 aux fonctions affines trouvé à la section 19.6 du cours ✓. Le cours donne aussi le projeté $t - \dfrac{1}{6}$, c'est-à-dire $a_0 = -\dfrac{1}{6}$ et $a_1 = 1$; on vérifie que

$$\frac{1}{X+3} + \frac{1}{6(X+1)} - \frac{1}{X+2} = \frac{X(X-1)}{6(X+1)(X+2)(X+3)},$$

en réduisant au même dénominateur, puisque le numérateur vaut

$$6(X+1)(X+2) + (X+2)(X+3) - 6(X+1)(X+3) = X^2 - X \quad ✓$$

À retenir. Par la formule de Stirling (chapitre 18, théorème 10), $\displaystyle\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}}$, et le minimum est équivalent à $\dfrac{\pi}{2\cdot 16^n}$: pour $n = 20$, le rapport des deux vaut 0,988 ✓. La distance de t^n aux polynômes de degré inférieur décroît comme 4^{-n} : sur $[0\,,1]$, la fonction t^n est très bien approchée par des polynômes de degré plus petit. Les coefficients a_j sont solution d'un système dont la matrice est la matrice de Hilbert de l'exercice 27 ; la fraction rationnelle évite de résoudre ce système, très sensible aux erreurs d'arrondi.

EXERCICE 53 ⋆⋆⋆

1. Les x_k étant distincts, on dispose des polynômes de Lagrange $\ell_1, \ldots, \ell_n \in \mathbb{R}_{n-1}[X]$, caractérisés par $\ell_k(x_j) = \delta_{kj}$, et de la formule d'interpolation $f = \displaystyle\sum_k f(x_k)\ell_k$ pour $f \in \mathbb{R}_{n-1}[X]$ (chapitre 13, théorème 13).

Si (λ_k) convient, l'égalité $(*)$ appliquée à $f = \ell_j$ donne $\displaystyle\int_{-1}^1 \ell_j = \sum_k \lambda_k\delta_{jk} = \lambda_j$: il y a au plus une solution. Réciproquement, avec $\lambda_k = \displaystyle\int_{-1}^1 \ell_k$, la linéarité de l'intégrale et la formule d'interpolation donnent, pour $f \in \mathbb{R}_{n-1}[X]$,

$$\int_{-1}^1 f = \sum_{k=1}^n f(x_k) \int_{-1}^1 \ell_k = \sum_{k=1}^n \lambda_k f(x_k).$$

$$\boxed{\text{il existe un unique } (\lambda_1, \ldots, \lambda_n), \text{ donné par } \lambda_k = \int_{-1}^1 \ell_k(t)\,\mathrm{d}t}$$

2. Soit $f \in \mathbb{R}_{2n-1}[X]$, et $f = QL_n + R$ sa division euclidienne par L_n, qui est de degré n (chapitre 13, théorème 2) : $\deg R \leqslant n-1$, et $\deg Q \leqslant n-1$, puisque $\deg f \leqslant 2n-1$. Comme L_n est orthogonal à $\mathbb{R}_{n-1}[X]$ (proposition 7 (i)), on a $\displaystyle\int_{-1}^1 QL_n = \langle Q, L_n\rangle = 0$. Comme $L_n(x_k) = 0$, on a $f(x_k) = R(x_k)$. Par la question 1 appliquée à R,

$$\int_{-1}^1 f = 0 + \int_{-1}^1 R = \sum_{k=1}^n \lambda_k R(x_k) = \sum_{k=1}^n \lambda_k f(x_k).$$

$$\boxed{(*) \text{ est vraie pour tout } f \in \mathbb{R}_{2n-1}[X]}$$

Avec n points, on intègre exactement les polynômes de degré au plus $2n-1$: c'est le double de ce que donne l'interpolation seule. Ce n'est plus vrai au degré $2n$: pour $f = L_n^2$, le second membre est nul, et le premier vaut $\|L_n\|^2 > 0$.

3. Soit $W = (X-y_1)\cdots(X-y_n)$, de degré n. Pour tout $Q \in \mathbb{R}_{n-1}[X]$, le polynôme QW est de degré au plus $2n-1$, et il s'annule en chaque y_k :

$$\langle Q, W\rangle = \int_{-1}^{1} QW = \sum_{k=1}^{n} \mu_k Q(y_k)W(y_k) = 0.$$

Ainsi W appartient à l'orthogonal de $\mathbb{R}_{n-1}[X]$ dans l'espace euclidien $\mathbb{R}_n[X]$. Par la proposition 11, cet orthogonal est de dimension $(n+1)-n = 1$; il contient $L_n \neq 0$, donc c'est la droite $\mathrm{Vect}(L_n)$, et $W = cL_n$ avec $c \neq 0$. Les deux polynômes ont les mêmes racines :

$$\boxed{\{y_1,\dots,y_n\} = \{x_1,\dots,x_n\}, \quad \text{donc } y_k = x_k \text{ pour tout } k}$$

puisque les deux listes sont rangées dans l'ordre croissant. Par l'unicité de la question 1, on a aussi $\mu_k = \lambda_k$: la formule de Gauss est la seule formule à n points exacte au degré $2n-1$.

4. Le polynôme ℓ_k^2 est de degré $2n-2 \leqslant 2n-1$; la question 2 lui applique $(*)$:

$$\int_{-1}^{1} \ell_k^2 = \sum_{j=1}^{n} \lambda_j \ell_k(x_j)^2 = \lambda_k,$$

et cette intégrale est strictement positive, puisque $\ell_k \neq 0$ et que $\displaystyle\int_{-1}^{1} PQ$ est un produit scalaire. Avec $f = 1$, on obtient $\displaystyle\int_{-1}^{1} 1 = 2$:

$$\boxed{\lambda_k = \int_{-1}^{1} \ell_k^2 > 0, \qquad \sum_{k=1}^{n} \lambda_k = 2}$$

5. La section 19.4 du cours donne $L_2 = X^2 - \dfrac{1}{3}$, de racines $\boxed{x_1 = -\dfrac{1}{\sqrt{3}},\ x_2 = \dfrac{1}{\sqrt{3}}}$. On a $\lambda_1+\lambda_2 = 2$, et $(*)$ appliquée à $f = X$ donne $0 = \dfrac{\lambda_2 - \lambda_1}{\sqrt{3}}$, d'où $\boxed{\lambda_1 = \lambda_2 = 1}$. La formule

$$\int_{-1}^{1} f(t)\,\mathrm{d}t \approx f\Big(-\frac{1}{\sqrt{3}}\Big) + f\Big(\frac{1}{\sqrt{3}}\Big)$$

est exacte pour les polynômes de degré au plus 3, mais pas pour X^4 : elle donne $\dfrac{2}{9}$ au lieu de $\dfrac{2}{5}$ ✓. Pour $f = \exp$, elle donne $2\,\mathrm{ch}\,\dfrac{1}{\sqrt{3}} \approx 2{,}3427$ au lieu de $2\,\mathrm{sh}\,1 \approx 2{,}3504$, soit une erreur de 0,0077 avec deux évaluations seulement ; la méthode des trapèzes à trois points, -1, 0, 1, se trompe de 0,19. Pour $n = 3$, on trouve les nœuds 0 et $\pm\sqrt{3/5}$, de poids $\dfrac{8}{9}$ et $\dfrac{5}{9}$ ✓.

6. Soit $\varepsilon > 0$, et Q un polynôme tel que $|f-Q| \leqslant \varepsilon$ sur $[-1\,,1]$. Pour n assez grand, $2n-1 \geqslant \deg Q$, et la question 2 donne $I_n(Q) = \displaystyle\int_{-1}^{1} Q$. On écrit

$$\Big|I_n(f) - \int_{-1}^{1} f\Big| \leqslant |I_n(f) - I_n(Q)| + \Big|\int_{-1}^{1} Q - \int_{-1}^{1} f\Big|.$$

Par l'inégalité triangulaire et la positivité des poids (question 4), le premier terme est au plus $\displaystyle\sum_k \lambda_k |f(x_k) - Q(x_k)| \leqslant \varepsilon \sum_k \lambda_k = 2\varepsilon$. Par l'inégalité de la moyenne (chapitre 18, proposition 7),

le second est au plus 2ε. Ainsi $|I_n(f) - \int_{-1}^{1} f| \leqslant 4\varepsilon$ pour n assez grand :

$$\boxed{I_n(f) \xrightarrow[n\to+\infty]{} \int_{-1}^{1} f(t)\,\mathrm{d}t}$$

Piège. La positivité des poids est le point décisif : c'est elle qui transforme $\sum_k \lambda_k |f(x_k) - Q(x_k)|$ en au plus 2ε. Avec des nœuds équidistants et des poids définis comme à la question 1, certains poids deviennent négatifs dès neuf points (le plus petit vaut alors $-\frac{908}{2835}$), la somme des $|\lambda_k|$ dépasse alors 2, et la majoration par 2ε tombe. Placer les nœuds aux racines d'un polynôme orthogonal n'est donc pas un raffinement : c'est ce qui garantit la convergence.

Les quatre derniers exercices sont des problèmes en autonomie ; le corrigé commence par l'idée directrice.

EXERCICE 54 ⋆⋆⋆⋆

Idée. La condition $\langle x_i, x_j\rangle = -1$ s'écrit $\langle x_i, x_j\rangle + 1 \cdot 1 = 0$: en ajoutant aux x_i une coordonnée supplémentaire égale à 1, on obtient une famille orthogonale. L'inégalité est alors l'inégalité de Bessel pour le vecteur « vertical » $(0, 1)$, et le cas d'égalité est celui de Bessel.

Soit $n = \dim E$. On munit $E' = E \times \mathbb{R}$ de

$$\langle (x, s), (y, u)\rangle' = \langle x, y\rangle + su.$$

C'est un produit scalaire : l'application est bilinéaire et symétrique, et $\langle (x, s), (x, s)\rangle' = \|x\|^2 + s^2$ est positif, nul seulement si $x = 0$ et $s = 0$. L'espace E' est euclidien, de dimension $n + 1$. On pose

$$y_i = (x_i, 1) \quad (1 \leqslant i \leqslant p) \qquad \text{et} \qquad z = (0, 1),$$

de sorte que, pour tous $i \neq j$,

$$\langle y_i, y_j\rangle' = -1 + 1 = 0, \qquad \|y_i\|'^2 = 1 + \|x_i\|^2, \qquad \langle z, y_i\rangle' = 1.$$

La famille $(y_1, \dots, y_p)$ est orthogonale et formée de vecteurs non nuls ; les vecteurs $\frac{y_i}{\|y_i\|'}$ forment donc une base orthonormée du sous-espace $F = \mathrm{Vect}(y_1, \dots, y_p)$ (proposition 6). L'inégalité de Bessel (proposition 10 (iv)) appliquée à z, unitaire, donne

$$\sum_{i=1}^{p} \frac{\langle z, y_i\rangle'^2}{\|y_i\|'^2} \leqslant \|z\|'^2, \qquad \text{soit} \qquad \boxed{\sum_{i=1}^{p} \frac{1}{1 + \|x_i\|^2} \leqslant 1}$$

Cas d'égalité. On note $\lambda_i = \dfrac{1}{1 + \|x_i\|^2}$. Par la proposition 10, sous la forme adaptée aux bases orthogonales, $p_F(z) = \sum_i \dfrac{\langle z, y_i\rangle'}{\|y_i\|'^2}\, y_i = \sum_i \lambda_i y_i = (\sum_i \lambda_i x_i, \sum_i \lambda_i)$. Il y a égalité dans Bessel si et seulement si $z \in F$, c'est-à-dire $z = p_F(z)$. Si c'est le cas, la première composante donne $\sum_i \lambda_i x_i = 0$. Réciproquement, si $\sum_i \lambda_i x_i = 0$, alors $z - p_F(z) = (0, 1 - s)$ avec $s = \sum_i \lambda_i$; ce vecteur est orthogonal à y_1, ce qui s'écrit $1 - s = 0$, donc $z = p_F(z)$.

$$\boxed{\sum_{i=1}^{p} \frac{1}{1 + \|x_i\|^2} = 1 \iff \sum_{i=1}^{p} \frac{x_i}{1 + \|x_i\|^2} = 0}$$

Voici trois situations qui éclairent ce cas d'égalité.

— Pour $p = 1$, la condition sur les produits scalaires est vide, et l'égalité équivaut à $x_1 = 0$.

— Pour $p = 2$, avec $x_1 = au$ et $x_2 = -\frac{1}{a}u$, où u est unitaire et $a > 0$, on trouve $\frac{1}{1+a^2} + \frac{a^2}{a^2+1} = 1$: deux vecteurs colinéaires de sens contraires réalisent l'égalité. Pour $x_1 = (2, 0)$ et $x_2 = (-\frac{1}{2}, 1)$ dans $\mathbb{R}^2$, on a bien $\langle x_1, x_2\rangle = -1$, mais la somme vaut $\frac{1}{5} + \frac{4}{9} = \frac{29}{45} < 1$.

— Pour $p = n+1$, la famille orthogonale (y_i) de vecteurs non nuls est une base de E', donc $z \in F$: il y a toujours égalité. C'est le cas du simplexe régulier de l'exercice 51, avec $x_i = \sqrt{n}\, e_i$: on a $\langle x_i, x_j\rangle = n \cdot (-\frac{1}{n}) = -1$, et chaque terme vaut $\frac{1}{n+1}$.

Les égalités et inégalités ont été testées sur des familles aléatoires construites de cette façon, pour $n \leqslant 4$ ✓.

À retenir. La construction montre en outre que $p \leqslant n+1$, puisqu'une famille orthogonale de vecteurs non nuls de E' est libre (proposition 6) : on ne peut pas trouver plus de $n+1$ vecteurs de E dont les produits scalaires deux à deux valent -1. Ajouter une dimension pour rendre une famille orthogonale est une idée qui resservira.

EXERCICE 55 ⋆⋆⋆⋆

Idée. Une matrice de rang 1 s'écrit XY^T, avec deux colonnes non nulles. Sa norme vaut $\|X\|\,\|Y\|$ et sa trace $\langle X, Y\rangle$; en développant $\|M - I_n\|^2$, on fait apparaître l'inégalité de Cauchy-Schwarz, puis un trinôme en $\|X\|\,\|Y\|$.

On note $\langle X, Y\rangle = X^TY$ le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ et $\|\cdot\|$ les deux normes, sur les colonnes et sur les matrices, ce qui ne crée pas d'ambiguïté.

Les matrices de rang 1. Si $\mathrm{rg} M = 1$, l'image de M est une droite $\mathrm{Vect}(X)$, avec $X \neq 0$, et chaque colonne de M s'écrit y_jX : ainsi $M = XY^T$, avec $Y = (y_j) \neq 0$ puisque $M \neq 0$. Réciproquement, pour X et Y non nuls, la matrice XY^T, de colonnes y_jX non toutes nulles, est de rang 1.

Le calcul. Les coefficients de $M = XY^T$ sont les x_iy_j, donc, par la proposition 1 (ii),

$$\|M\|^2 = \sum_{i,j} x_i^2 y_j^2 = \|X\|^2\|Y\|^2, \qquad \langle M, I_n\rangle = \mathrm{tr}(M) = \sum_i x_iy_i = \langle X, Y\rangle.$$

On note $s = \|X\|\,\|Y\|$, qui est strictement positif. Comme $\|I_n\|^2 = n$, la proposition 3 (ii) et l'inégalité de Cauchy-Schwarz donnent

$$\|M - I_n\|^2 = s^2 - 2\langle X, Y\rangle + n \geqslant s^2 - 2s + n = (s-1)^2 + n - 1 \geqslant n - 1.$$

Le cas d'égalité. Il y a égalité si et seulement si $s = 1$ et $\langle X, Y\rangle = \|X\|\,\|Y\|$. La seconde condition équivaut, par le cas d'égalité de l'inégalité triangulaire (théorème 2), comme dans l'assertion 4 de l'exercice 49, à $Y = \lambda X$ avec $\lambda \geqslant 0$, et même $\lambda > 0$ puisque $Y \neq 0$. La première s'écrit alors $\lambda\|X\|^2 = 1$, et $M = \lambda XX^T = \dfrac{XX^T}{\|X\|^2}$. Réciproquement, pour une telle matrice, $s = 1$ et $\langle X, Y\rangle = 1$, et $\|M - I_n\|^2 = 1 - 2 + n = n - 1$.

$$\boxed{\inf_{M\in\mathcal{R}} \|M - I_n\| = \sqrt{n-1}, \quad \text{atteinte exactement en } M = \frac{XX^T}{\|X\|^2},\ X \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0\}}$$

Une recherche aléatoire sur 20 000 matrices de rang 1, pour $n = 2$, 3 et 4, ne descend jamais sous $\sqrt{n-1}$ ✓. La matrice $2XX^T$, avec X unitaire, donne $\|M - I_n\|^2 = n$: il ne suffit pas d'avoir la bonne direction, il faut aussi la bonne « taille ».

À retenir. Pour $X = u$ unitaire, la matrice uu^T est celle de la projection orthogonale sur la droite $\mathrm{Vect}(u)$, puisque $uu^TZ = \langle Z, u\rangle\, u$ (proposition 10 (i)). Les matrices de rang 1 les plus proches de l'identité sont donc les projections orthogonales sur une droite, et il y en a une infinité : la borne inférieure est atteinte, mais pas en un point unique. La valeur $n - 1$ s'interprète aussi : $I_n - uu^T$ est la projection orthogonale sur l'hyperplan $u^\perp$, et pour une matrice P symétrique telle que $P^2 = P$, on a $\|P\|^2 = \mathrm{tr}(P^TP) = \mathrm{tr}(P) = \mathrm{rg}(P)$.

EXERCICE 56 ⋆⋆⋆⋆

Idée. Pour une famille orthonormée $(f_1,\dots,f_m)$ de V et un point x fixé, on applique l'hypothèse à la fonction $\sum_i f_i(x)f_i$, choisie pour que sa valeur en x soit $\sum_i f_i(x)^2$. On obtient $\sum_i f_i(x)^2 \leqslant A^2$, et l'on intègre en x : le membre de gauche donne m.

On munit V du produit scalaire $\langle f,g\rangle = \int_0^1 fg$ de la proposition 2.

Une famille orthonormée a au plus A^2 éléments. Soit $(f_1,\dots,f_m)$ une famille orthonormée de V, $x\in[0\,,1]$, et $g=\sum_{i=1}^m f_i(x)f_i$, qui appartient à V. Par le théorème de Pythagore, $\|g\|^2=\sum_i f_i(x)^2$, et l'on a $g(x)=\sum_i f_i(x)^2$. En notant $S(x)=\sum_i f_i(x)^2$, l'hypothèse appliquée à g donne

$$S(x)=g(x)\leqslant \sup_{[0,1]}|g|\leqslant A\|g\|=A\sqrt{S(x)}.$$

Si $S(x)>0$, on divise par $\sqrt{S(x)}$ et l'on obtient $S(x)\leqslant A^2$; c'est encore vrai si $S(x)=0$. La fonction S est continue et majorée par A^2 sur $[0\,,1]$, et son intégrale vaut $\sum_i \|f_i\|^2 = m$. Par croissance de l'intégrale,

$$\boxed{m=\int_0^1 \sum_{i=1}^m f_i(x)^2\,\mathrm{d}x\leqslant A^2}$$

La dimension de V. Si V n'était pas de dimension finie, il contiendrait des familles libres de toute longueur : partant d'une famille libre $(v_1,\dots,v_k)$, qui n'engendre pas V, on lui ajoute un vecteur $v_{k+1}\notin \mathrm{Vect}(v_1,\dots,v_k)$, et la famille reste libre (chapitre 14, proposition 8). L'algorithme de Gram-Schmidt (théorème 4) transformerait une famille libre de $m>A^2$ éléments en une famille orthonormée de m éléments, ce qui contredit ce qui précède. Donc V est de dimension finie, et une base orthonormée de V (théorème 5) a $\dim V$ éléments :

$$\boxed{\dim V\leqslant A^2}$$

Optimalité. Pour $n\geqslant 1$ et $k\in[\![0\,;n-1]\!]$, soit φ_k la fonction égale à 1 sur $[\frac{k}{n}\,,\frac{k+1}{n}[$ et à 0 ailleurs, avec de plus $\varphi_{n-1}(1)=1$. Ces fonctions en escalier, à supports disjoints non vides, forment une famille libre, et $V=\mathrm{Vect}(\varphi_0,\dots,\varphi_{n-1})$ est de dimension n. Pour $f=\sum_k c_k\varphi_k$, on a

$$\sup_{[0,1]}|f|=\max_k|c_k|,\qquad \int_0^1 f^2=\frac{1}{n}\sum_k c_k^2,\qquad \text{donc}\qquad \Big(\sup_{[0,1]}|f|\Big)^2\leqslant\sum_k c_k^2=n\int_0^1 f^2.$$

$$\boxed{\sup_{[0,1]}|f|\leqslant\sqrt{n}\Big(\int_0^1 f^2\Big)^{1/2}\text{ pour toute }f\in V,\qquad \dim V=n=A^2}$$

La majoration $\dim V\leqslant A^2$ ne peut donc pas être améliorée pour les fonctions continues par morceaux, puisqu'on y a l'égalité. La démonstration de la première partie s'y applique d'ailleurs telle quelle, à une réserve près : sur $\mathcal{C}_m([0\,,1],\mathbb{R})$, la formule $\int_0^1 fg$ n'est pas définie, mais l'hypothèse elle-même la rend définie sur V, puisque $\int_0^1 f^2=0$ y entraîne $\sup|f|=0$.

À retenir. La démonstration dit plus : la meilleure constante A^2 pour un espace V de dimension m est le maximum de $S=\sum_i f_i^2$, puisqu'on a égalité dans Cauchy-Schwarz pour g, et ce maximum est au moins la moyenne $\int_0^1 S=m$. Pour $V=\mathbb{R}_1[X]$, de base orthonormée $(1,\sqrt{3}(2t-1))$, on trouve $S(x)=1+3(2x-1)^2$, de moyenne 2 et de maximum 4, atteint aux bords : la meilleure constante est $A=2$, et l'on a $\dim V=2<4=A^2$ ✓.

EXERCICE 57 ⋆⋆⋆⋆

Idée. Pour X fixé, les deux formes linéaires $Y \mapsto X^TAY$ et $Y \mapsto Y^TAX$ ont le même noyau. Ce sont les orthogonaux des vecteurs A^TX et AX, qui sont donc colinéaires. Un coefficient de proportionnalité qui ne dépend pas de X donne $A = cA^T$, et l'on conclut en transposant.

On munit $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique. Le réel Y^TAX, matrice 1×1, est égal à son transposé X^TA^TY, donc

$$X^TAY = \langle A^TX, Y\rangle, \qquad Y^TAX = \langle AX, Y\rangle.$$

L'hypothèse, appliquée au couple (X, Y) puis au couple (Y, X), donne une équivalence, et pour toute colonne X,

$$X^TAY = 0 \iff Y^TAX = 0 \quad \text{pour tout } Y, \qquad \text{c'est-à-dire} \qquad \mathrm{Vect}(A^TX)^\perp = \mathrm{Vect}(AX)^\perp.$$

En prenant les orthogonaux, par la proposition 11, on obtient

$$\mathrm{Vect}(A^TX) = \mathrm{Vect}(AX) \qquad \text{pour toute colonne } X.$$

En particulier, $AX = 0 \iff A^TX = 0$, et, si $A^TX \neq 0$, il existe un unique réel $c(X) \neq 0$ tel que $AX = c(X)\,A^TX$.

Le coefficient est constant. Supposons $A \neq 0$, et soit X, Y deux colonnes telles que $A^TX \neq 0$ et $A^TY \neq 0$.

— Si A^TX et A^TY ne sont pas colinéaires, alors $A^T(X+Y) \neq 0$, et

$$c(X)\,A^TX + c(Y)\,A^TY = A(X+Y) = c(X+Y)\,A^TX + c(X+Y)\,A^TY.$$

La famille (A^TX, A^TY) étant libre, on identifie : $c(X) = c(X+Y) = c(Y)$.

— Si $A^TY = \mu\,A^TX$, avec $\mu \neq 0$, alors $A^T(Y - \mu X) = 0$, donc $A(Y - \mu X) = 0$, et

$$AY = \mu AX = \mu c(X)\,A^TX = c(X)\,A^TY, \qquad \text{d'où} \qquad c(Y) = c(X).$$

Soit c la valeur commune. On a $AX = c\,A^TX$ pour tout X tel que $A^TX \neq 0$, et aussi lorsque $A^TX = 0$, puisqu'alors $AX = 0$. Ainsi $A = c\,A^T$, et en transposant, $A^T = cA$, d'où $A = c^2A$. Comme $A \neq 0$, $c^2 = 1$:

$$\boxed{A^T = A \quad \text{ou} \quad A^T = -A}$$

Réciproquement, une matrice symétrique ou antisymétrique vérifie $Y^TAX = \pm X^TAY$, et donc l'hypothèse.

Piège. L'hypothèse ne porte que sur l'annulation, et non sur les valeurs : l'erreur la plus répandue est d'en déduire directement $X^TAY = Y^TAX$. La matrice suivante, ni symétrique ni antisymétrique, montre que l'hypothèse n'est pas automatique : pour $X = E_1$ et $Y = E_2 - E_1$, on a $X^TAY = 0$ mais $Y^TAX = -1$ ✓.

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

À retenir. L'application $(X, Y) \mapsto X^TAY$ est bilinéaire, et l'on peut définir une « orthogonalité » par $X^TAY = 0$. L'exercice montre que cette relation n'est symétrique que dans deux cas : pour une forme symétrique, comme un produit scalaire (section 19.2), ou pour une forme antisymétrique, comme le déterminant de deux colonnes de $\mathcal{M}_{2,1}(\mathbb{R})$ (chapitre 17), pour lequel « orthogonal » signifie « colinéaire ».