CP2i. Mathématiques 3ème semestre. 2021-2022.
Notes de cours

Thomas Duyckaerts

CHAPITRE 1

Rappels d'algèbre linéaire

Ce cours chapitre est un chapitre de rappel sur les matrices et les applications linéaires dans un espace vectoriel de dimension finie. Les notions d'espace vectoriel, de base, de dimension sont supposées connues. Le lecteur est invité à relire son cours de l'année dernière pour réviser ces notions et pour les démonstrations des résultats présentés ici.

Dans tout ce chapitre $\mathbb{K}$ sera un corps commutatif (on peut choisir $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ pour fixer les idées).

1. Matrices

1.a. Définition. Une matrice $n \times p$ sur $\mathbb{K}$ est un élement M de $\mathbb{K}^{n\times p}$, notée comme un carré à n lignes et p colonnes. Les coordonnées m_{ij}, $1 \leq i \leq n$, $1 \leq j \leq p$ de la matrices sont appelés *éléments* ou coefficients. On note $M = [m_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$. Par convention le premier indice (ici i) désigne toujours le numéro de colonne et le deuxième (ici j) désigne la colonne.

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices $n\times p$ sur $\mathbb{K}$. L'ensemble des matrices *carrées* $\mathcal{M}_{n,n}(\mathbb{K})$ est aussi noté $\mathcal{M}_n(\mathbb{K})$

Exemple 1.1.

$$M = \begin{pmatrix} -1 & 2 \\ 4 & -3 \\ 5 & 6 \end{pmatrix} = [m_{i,j}]_{\substack{1\leq i\leq 3\\ 1\leq j\leq 2}}$$

est une matrice 3×2 sur $\mathbb{R}$ (un élément de $\mathcal{M}_{3,2}(\mathbb{R})$). On a $m_{2,1} = 4$, $m_{3,2} = 6$.

1.b. Opérations. Soit $A = [a_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$, $B = [b_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ deux matrices *de même taille* et $\lambda \in \mathbb{K}$. Le produit de A par le scalaire λ et la somme $A+B$ sont des matrices de même taille, définies en faisant les opérations coefficients par coefficients:

$$\lambda A = [\lambda a_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}, \quad A + B = [a_{ij} + b_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}.$$

Ces opérations confèrent à $\mathcal{M}_{n,p}(\mathbb{K})$ une structure d'espace vectoriel (en particulier, l'addition est commutative et associative). L'élément neutre pour l'addition est la matrice nulle, noté 0 ou $0_{n,p}$, dont tous les coefficients sont nuls.

Soit maintenant $C = [c_{i,j}]_{\substack{1\leq i\leq p\\ 1\leq j\leq q}} \in \mathcal{M}_{p,q}$, une matrice dont *le nombre de lignes est égal au nombre de colonnes de A*. Par définition le produit AC de A et C est la matrice $n \times q$ définie par

$$AC = [d_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq q}}, \quad d_{ij} = \sum_{k=1}^{p} a_{ik}c_{kj}, \quad 1 \leq i \leq n,\ 1 \leq j \leq q.$$

EXERCICE 1.2. Calculer le produit $\left(\begin{smallmatrix} 1 & 3 & 4 \\ 2 & 4 & 5 \end{smallmatrix}\right)\begin{pmatrix} -1 & 0 & 0 & 7 \\ 2 & 0 & 0 & 1 \\ 4 & 0 & 1 & 2 \end{pmatrix}$.

EXERCICE 1.3. A quel condition sur la matrice A le carré $A^2 = A \times A$ est-il bien défini?

On rappelle que la multiplication est associative et distributive par rapport à l'addition, mais n'est pas commutative: calculer par exemple les produits AC et CA pour $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

La multiplication admet un élément neutre, appelé *matrice identité* et noté I_n. Par définition, $I_n = [\delta_{ij}]_{1 \leq {i \atop j} \leq n}$, où δ_{ij} est le symbole de Kronecker, $\delta_{ij} = 0$ si $i \neq j$ et $\delta_{ii} = 1$. Ainsi, si A est une matrice $n \times p$, on a

$$A = I_n A = A I_p.$$

1.c. Matrices inversibles. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On dit que A est inversible lorsqu'il existe $B, C \in \mathcal{M}_n(\mathbb{K})$ tels que $BA = AC = I_n$. Dans ce cas $C = B$ est unique, et noté A^{-1}.

On peut vérifier l'inversibilité d'une matrice et calculer A^{-1} par la méthode du pivot de Gauss.

EXEMPLE 1.4. Lorsque A est une matrice 2×2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ est inversible } \iff ad - bc \neq 0.$$

Dans ce cas,

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

$ad - bc$ est appelé le *déterminant* de A.

Le déterminant, et la formule de l'exemple précédent se généralisent à la dimension supérieure.

Ainsi

$$\det[a_{ij}]_{1 \leq {i \atop j} \leq n} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon_\sigma \prod_{j=1}^{n} a_{\sigma(j)j},$$

où ε_σ est la signature de la permutation σ, et on a

$$A \text{ inversible } \iff \det(A) \neq 0.$$

Le lecteur est invité à réviser le chapitre sur le déterminant du cours de l'année dernière.

EXEMPLE 1.5. $\begin{pmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \\ 0 & 7 & 13 \end{pmatrix}$ n'est pas inversible. (pourquoi?). $\left(\begin{smallmatrix} 3 & 4 \\ 2 & 3 \end{smallmatrix}\right)$ est inversible (calculer son inverse). I_n est inversible, d'inverse I_n.

EXERCICE 1.6. Trouver d'autres matrices carrées qui sont leur propre inverse.

2. Applications linéaires

2.a. Définitions. Soit E et F des $\mathbb{K}$-espaces vectoriels. Une *application linéaire* de E dans F est une application de E dans F telle que

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

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F. Lorsque $E = F$, on dit que f est un endomorphisme de E. On note $\mathcal{L}(E) = \mathcal{L}(E, E)$ l'ensemble des endomorphismes de E.

Exemple 2.1. La formule $f\left(\left(\begin{smallmatrix}x_1\\x_2\\x_3\end{smallmatrix}\right)\right)=\left(\begin{smallmatrix}x_1-3x_2\\2x_1+4x_2-x_3\end{smallmatrix}\right)$ définit une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^2$. L'identité de E (que l'on notera Id_E) est un endomorphisme de E.

Si f, g sont des éléments de $\mathcal{L}(E,F)$ et $\lambda\in\mathbb{K}$, on définit les éléments $f+g$ et λf de $\mathcal{L}(E,F)$ par

$$\forall x\in E,\quad (f+g)(x)=f(x)+g(x),\quad (\lambda f)(x)=\lambda f(x).$$

(On utilise donc la structure de $\mathbb{K}$-espace vectoriel sur F). Ceci confère à $\mathcal{L}(E,F)$ une structure d'espace vectoriel.

2.b. Matrice de représentation d'une application linéaire. On suppose E et F de dimensions finies respectivement p et n. Soit $\mathcal{B}=(b_1,\dots,b_p)$ une base de E et $\mathcal{C}=(c_1,\dots,c_n)$ une base de F. Soit $f\in\mathcal{L}(E,F)$. Par définition, la *matrice de représentation* de f dans les bases $\mathcal{B}$ et $\mathcal{C}$, noté $\mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$ est la matrice $n\times p$ dont la j-ième colonne est formé des coordonnées de $f(b_j)$ dans la base $\mathcal{C}$.

Exercice 2.2. Soit $\mathcal{B}=\left(\left(\begin{smallmatrix}3\\1\\10\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right)\right)$ et $\mathcal{C}=\left(\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\right)$, et f donné par l'exemple 2.1. Justifier que $\mathcal{B}$ et $\mathcal{C}$ sont des bases et calculer $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$.

Un cas particulier important est celui de la base canonique de l'espace $\mathbb{R}^n$. On rappelle que cette base canonique $\mathcal{B}_n=(e_1,\dots,e_n)$ est donnée par $e_j=(\delta_{ij})_{1\le i\le n}$. Par exemple,

$$\mathcal{B}_3=\left(\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right).$$

Si $f\in\mathcal{L}(\mathbb{R}^p,\mathbb{R}^n)$ et $A=[a_{ij}]_{\substack{1\le i\le n\\1\le j\le p}}$ est la matrice de f dans les bases $\mathcal{B}_p$ et $\mathcal{B}_n$, on a

$$f\big((x_i)_{1\le i\le p}\big)=\left(\sum_{j=1}^p a_{i,j}x_j\right)_{1\le i\le n}.$$

La matrice A se "lit" donc sur la formule définissant f.

Exemple 2.3. Soit f donné par l'exemple 2.1. La matrice de f dans les bases canoniques de $\mathbb{R}^3$ et $\mathbb{R}^2$ est donnée par

$$\begin{pmatrix}1&-3&0\\2&4&-1\end{pmatrix}.$$

Comparer avec le résulat de l'exercice 2.2.

Lorsque f est un endomorphisme de E, on représente généralement f en choisissant la même base de départ et d'arrivée. Dans ce contexte, on note $\mathrm{Mat}_{\mathcal{B}}(f)=\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f)$.

Exercice 2.4. Calculer $\mathrm{Mat}_{\mathcal{B}}(\mathrm{Id}_E)$, où $\mathcal{B}$ est une base de E. Le résultat dépend-il du choix de $\mathcal{B}$? Persiste-t-il pour $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{Id}_E)$ où $\mathcal{B}$ et $\mathcal{C}$ sont deux bases *distinctes* de E?

2.c. Effet des opérations sur les matrices de représentations. On se donne deux $\mathbb{K}$-espaces vectoriels E et F de dimensions finies p et n respectivement. Soit $\mathcal{B}$ une base de E et $\mathcal{C}$ une base de F.

- Soit $f \in \mathcal{L}(E,F)$ et $x \in E$. Soit $X \in \mathcal{M}_{p,1}(\mathbb{K})$ le vecteur colonne des coordonnées de x dans $\mathcal{B}$, et $Y \in \mathcal{M}_{n,1}(\mathbb{K})$ le vecteur colonne des coordonnées de $f(x)$ dans $\mathcal{C}$. Alors
$$Y = \mathcal{M}_{\mathcal{B},\mathcal{C}}(f)X.$$
- Soit $f, g \in \mathcal{L}(E,F)$, et $\lambda \in \mathbb{K}$. Alors
$$\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f+g) = \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f) + \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(g), \quad \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(\lambda f) = \lambda \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f).$$
- Soit G un $\mathbb{K}$-espace vectoriel, $\mathcal{D}$ une base de G et $g \in \mathcal{L}(F,G)$. Alors
$$\operatorname{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f) = \operatorname{Mat}_{\mathcal{C},\mathcal{D}}(g) \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f).$$
Le terme de droite de l'égalité est un produit matriciel: le produit des matrices de représentations correspond à la composition des applications.

2.d. Changement de bases. Soit $f \in \mathcal{L}(E,F)$, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, $\mathcal{C}$ et $\mathcal{C}'$ deux bases de F. On voudrait exprimer $\operatorname{Mat}_{\mathcal{B}',\mathcal{C}'}(f)$ à partir de $\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f)$. On utilise pour cela la notion de *matrice de passage.*

DÉFINITION 2.5. La matrice de passage de $\mathcal{B} = (b_1, \dots, b_p)$ à $\mathcal{B}' = (b'_1, \dots, b'_p)$, notée $P_{\mathcal{B},\mathcal{B}'}$ est la matrice $p \times p$ (où $p = \dim E$) dont les vecteurs colonnes sont les coordonnées des vecteurs de la base $\mathcal{B}'$ dans la base $\mathcal{B}$.

Plus précisément, la j-ième colonne de $P_{\mathcal{B},\mathcal{B}'}$ est formée des coordonnées de b'_j dans la base $\mathcal{B}$.

On vérifie facilement, à partir de sa définition que la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice, dans la base $\mathcal{B}$, de l'endomorphisme u de E défini par $u(b_j) = b'_j$, $j = 1, \dots p$.

EXEMPLE 2.6. Soit $\mathcal{B}$ la base canonique de $\mathbb{R}^2$ et $\mathcal{B}' = \left(\left(\begin{smallmatrix}2\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\right)$. Alors
$$P_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

PROPOSITION 2.7. *La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est inversible. Son inverse est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.*

La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ transforme les coordonnées d'un vecteur de E dans $\mathcal{B}'$ en ses coordonnées dans $\mathcal{B}$ (**attention à l'inversion des ordres des bases!**):

PROPOSITION 2.8. *Soit $x \in E$, X le vecteur colonne de ses coordonnées dans $\mathcal{B}$ et X' le vecteur colonne de ses coordonnées dans $\mathcal{B}'$. Alors $X = P_{\mathcal{B},\mathcal{B}'}X'$.*

On déduit de la proposition 2.8 la formule de changement de base pour les applications linéaires:
$$\operatorname{Mat}_{\mathcal{B}',\mathcal{C}'}(f) = (P_{\mathcal{C},\mathcal{C}'})^{-1} \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f) P_{\mathcal{B},\mathcal{B}'}$$

On résume la formule par le schéma suivant On en déduit notamment que deux matrices $n \times n$, A et B, représentent le même endomorphisme de E (avec, à chaque fois, la même base de départ et d'arrivée) si et seulement si existe une matrice $n \times n$ inversible P, telle que $A = P^{-1}BP$.

CHAPITRE 2

Réduction des endomorphismes

Ce chapitre est une version étendue du chapitre correspondant du polycopié écrit par Benoît Rittaud pour le cours 2019-2020.

À la fin de ce chapitre, vous devrez notamment savoir :

- Déterminer les valeurs propres et les sous-espaces propres d'un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie.
- Savoir si une matrice est diagonalisable ou trigonalisable, dans les cas où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et la diagonaliser ou la trigonaliser selon le cas.
- Appliquer la réduction des endomorphismes à des problèmes tels que l'étude des suites à récurrence linéaire ou les équations différentielles linéaires homogènes à coefficients constants.

Dans toute ce chapitre, $\mathbb{K}$ désigne un corps commutatif (on pourra prendre $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ ou $\mathbb{Q}$), E désigne un $\mathbb{K}$-espace vectoriel de dimension finie n, et f un endomorphisme de E.

1. Éléments propres d'un endomorphisme

1.a. Valeurs propres et vecteurs propres.

Définition 1.1. Une *valeur propre* de f est un scalaire $\lambda \in \mathbb{K}$ pour lequel il existe un élément v de E, *non nul*, tel que $f(v) = \lambda v$. Un tel vecteur est appelé *vecteur propre* de f pour la valeur propre λ. L'ensemble des valeurs propres de f est appelé *spectre* de f et noté $\mathrm{Sp}(f)$.

Remarque 1.2. Le spectre de f est parfois noté aussi $\sigma(f)$. En anglais, les valeurs propres sont appelées *eigenvalues* et les vecteurs propres *eigenvectors*.

Exemple 1.3.

- Soit $a \in \mathbb{K}$ et h_a l'homothétie de rapport a, définie par $h_a(u) = au$ pour tout u de E. Alors $\mathrm{Sp}(h_a) = \{a\}$. Les vecteurs propres de h_a pour la valeur propre a sont tous les vecteurs non nuls de E.
- Soit $\theta \in \mathbb{R}$ un angle tel que $\theta \not\equiv 0\ [2\pi]$ (cf plus bas pour une explication sur cette notation). On considère, sur $\mathbb{R}^2$, la rotation r_θ, de centre $(0,0)$ et d'angle θ. Alors si $\theta \not\equiv \pi\ [2\pi]$, r_θ n'a pas de valeur propre, puisque si v est un élément non nul de $\mathbb{R}^2$, $r_\theta(v)$ et v ne sont jamais colinéaires. En revanche, $r_\pi = h_{-1}$, et donc $\mathrm{Sp}(r_\pi) = \{-1\}$ et tout vecteur non nul est vecteur propre de r_π.

Notation 1.4. Dans le deuxième exemple, on a utilisé les notations suivantes: $\theta \equiv \theta'\ [2\pi]$ si et seulement si il existe $k \in \mathbb{Z}$ tel que $\theta = \theta' + 2\pi k$ (on dit alors que θ est congru à θ' modulo 2π). $\theta \not\equiv \theta'\ [2\pi]$ si et seulement si θ n'est pas congru à θ' modulo 2π.

1.b. Polynôme caractéristique. On rappelle que le déterminant d'un endomorphisme $f \in \mathcal{L}(E)$ est donné par $\det(f) = \det(\mathrm{Mat}_{\mathcal{B}}(f))$, où $\mathcal{B}$ est n'importe quelle base de E. Par la formule de changement de bases rappelée au chapitre précédent, ce déterminant ne dépend pas du choix de $\mathcal{B}$. En effet, si $\mathcal{B}'$ est une autre base de E,

$$\det(\mathrm{Mat}_{\mathcal{B}'}(f)) = \det(P^{-1}\,\mathrm{Mat}_{\mathcal{B}}(f)P) = \det(\mathrm{Mat}_{\mathcal{B}}(f)PP^{-1}) = \det(\mathrm{Mat}_{\mathcal{B}}(f)),$$

où $P = P_{\mathcal{B},\mathcal{B}'}$. Par le cours de S2 de CP2i, f est injectif si et seulement si il est bijectif si et seulement si son déterminant est non nul.

Définition 1.5. La formule

$$\chi_f(\lambda) = \det(f - \lambda\,\mathrm{Id}_E)$$

définit une fonction polynôme en λ. Le polynôme correspondant, noté aussi χ_f, est appelé *polynôme caractéristique de f*. C'est un un polynôme de degré n, de terme dominant $(-1)^n X^n$.

Démonstration. On choisit une base $\mathcal{B}$ de E. Soit $A = \mathrm{Mat}_{\mathcal{B}}(f)$. Alors

$$\chi_f(\lambda) = \det(f - \lambda\,\mathrm{Id}_E) = \det(A - \lambda I_n).$$

En notant $A = [a_{ij}]_{1 \leq {i \atop j} \leq n}$ et $A - \lambda I_n = [b_{ij}]_{1 \leq {i \atop j} \leq n}$, on a

$$b_{ij} = \begin{cases} a_{ij} & \text{si } i \neq j \\ a_{ij} - \lambda & \text{si } i = j. \end{cases}, \tag{1}$$

et $\chi_f(\lambda) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n b_{\sigma(j)j}$. D'après (1), le produit $\prod_{j=1}^n b_{\sigma(j)j}$ est un polynôme de degré au plus n en λ. Il est de degré exactement n quand $\sigma(j) = j$ pour tout j, c'est à dire quand σ est la permutation identité. On en déduit que χ_f est une fonction polynôme de degré n, avec

$$\chi_f(X) = \prod_{j=1}^n (a_{jj} - X)^n + P_{n-1} = (-X)^n + \widetilde{P}_{n-1},$$

où P_{n-1} et $\widetilde{P}_{n-1}$ sont des polynômes de degré au plus $n-1$. La deuxième égalité a été obtenue en développant le produit $\prod_{j=1}^n (a_{jj} - X)^n$. □

Exercice 1.6. Calculer les polynômes caractéristiques des endomorphismes h_a et r_θ de l'exemple 1.3.

Proposition 1.7. *Les valeurs propres de $f \in \mathcal{L}(E)$ sont les racines, dans $\mathbb{K}$, de son polynôme caractéristique.*

Démonstration.

$$\begin{aligned} \lambda \in \mathrm{Sp}(f) &\iff \exists v \in E \setminus \{0_E\},\ f(v) = \lambda v \\ &\iff f - \lambda\,\mathrm{Id}_E \ \text{ n'est pas injectif} \\ &\iff \chi_f(\lambda) = \det(f - \lambda\,\mathrm{Id}_E) = 0. \end{aligned}$$

De la proposition 1.7 et des propriétés connues des racines des polynômes on déduit:

Corollaire 1.8.
(i) *Un endomorphisme de E a au plus $n = \dim E$ valeurs propres.*
(ii) *Un endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension ≥ 1 a au moins une valeur propre.*

(iii) *Un endomorphisme d'un $\mathbb{R}$-espace vectoriel de dimension impair a au moins une valeur propre.*

L'exemple de la rotation de $\mathbb{R}^2$ (exemple 1.3) montre qu'un endomorphisme d'un $\mathbb{R}$-espace vectoriel de dimension paire peut ne pas avoir de valeur propre. □

1.c. Point de vue matriciel.

DÉFINITION 1.9. Soit A une matrice carrée $n \times n$ sur $\mathbb{K}$. Les valeurs propres de A sont par définition les valeurs propres de n'importe quel endomorphisme f dans un espace vectoriel E de dimension n tel que $A = \mathrm{Mat}_{\mathcal{B}}(f)$ pour une certaine base $\mathcal{B}$ de E.

La proposition 1.7 justifie la définition 1.9. Par cette proposition, les valeurs propres de f sont données par les racines de son polynôme caractéristiques

$$\det(f - \lambda\,\mathrm{Id}_E) = \det(A - \lambda I_n),$$

qui ne dépend pas du choix de f.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les valeurs propres de A sont les valeurs propres de l'endomorphisme de $\mathbb{K}^n$ (identifié à $\mathcal{M}_{n,1}(\mathbb{K})$) défini par $f_A : X \mapsto AX$. La matrice A est exactement la matrice de représentation de f_A dans la base canonique de $\mathbb{K}^n$. Les vecteurs propres de f_A sont appelés *vecteurs propres de A*.

EXEMPLE 1.10. (i) Soit $A = [a_{ij}]_{1 \leq {i \atop j} \leq n}$ une matrice triangulaire supérieure. Calculer les valeurs propres de A.

(ii) Déterminer les valeurs propres et les vecteurs propres des matrices

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

où les $(\lambda_j)_{j=1,2,3}$ sont deux à deux distinctes.

AVERTISSEMENT 1.11. Les valeurs propres d'une matrice de $\mathcal{M}_n(\mathbb{R})$, et de la même matrice, considérée comme un élément de $\mathcal{M}_n(\mathbb{C})$ peuvent différer. Considérer par exemple $R_\theta = \left(\begin{smallmatrix} \cos & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$.

1.d. Le théorème de Cayley-Hamilton. Soit $P(X) = \sum_{j=0}^{d} a_j X^j$ un polynôme, et f un endomorphisme, on note $P(f) = \sum_{j=0}^{d} a_j f^j$ (où f^j désigne

$$f^j = \underbrace{f \circ f \circ \ldots \circ f}_{j \text{ fois.}}$$

De même, si A est une matrice carrée, on note $P(A) = \sum_{j=0}^{d} a_j A^j$. On a alors le théorème de Cayley-Hamilton (admis):

THÉORÈME 1.12. *Soit f un endomorphisme et χ_f son polynôme caractéristique. Alors $\chi_f(f) = 0$.*

La version matricielle du théorème de Cayley-Hamilton est bien sûr valable: si A est une matrice carrée, $\chi_A(A)$ est la matrice nulle.

EXERCICE 1.13. Vérifier le théorème de Cayley-Hamilton sur la matrice $R_\theta = \left(\begin{smallmatrix} \cos & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$.

1.e. Sous-espaces propres.

Définition 1.14. On rappelle que le noyau d'un endomorphisme $f \in \mathcal{L}(E)$ est l'ensemble $\ker f = \{x \in E\ ,\ f(x) = O_E\}$. L'image de f est $\operatorname{Im}(f) = f(E)$. Ce sont deux sous-espaces vectoriels de E. De plus, on a le théorème du rang:

$$\dim \operatorname{Ker}(f) + \dim \operatorname{Im}(f) = \dim E.$$

($\dim \operatorname{Im}(f)$ est aussi appelé *rang* de f). En particulier,

$$f \text{ est injectif} \iff f \text{ est bijectif} \iff f \text{ est surjectif} \iff \operatorname{Ker}(f) = \{0_E\}.$$

Notation 1.15. Pour tout $\lambda \in \mathbb{K}$ et $f \in \mathcal{L}(E)$, on note $E_\lambda(f) = \operatorname{Ker}(f - \lambda \operatorname{Id}_E) = \{x \in E,\ f(x) = \lambda x$.

Par ce qui précède,

$$\lambda \in \operatorname{Sp}(f) \iff E_\lambda(f) \neq \{0_E\}.$$

Définition 1.16. Lorsque λ est valeur propre de f, $E_\lambda(f)$ est appelé *sous-espace propre* de f associé à la valeur propre λ. Il est formé des vecteurs propres de f associés à λ, et de O_E. Si A est une matrice, les sous-espaces propres de A sont les sous-espaces propres de l'endomorphisme de $\mathbb{K}^n$ $X \mapsto AX$, représenté par A dans la base canonique de $\mathbb{K}^n$.

Exercice 1.17. Déterminer les sous-espaces propres des éléments suivants de $\operatorname{Mat}_n(\mathbb{C})$:

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n), \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

où $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ est la matrice $n \times n$ diagonale dont les éléments diagonaux sont, de haut en bas, $\lambda_1, \dots, \lambda_n$.

2. Diagonalisation

2.a. Définition.

Définition 2.1. L'endomorphisme f de E est *diagonalisable* quand il existe une base $\mathcal{B}$ de E telle que $\operatorname{Mat}_{\mathcal{B}}(f)$ est diagonale. La base $\mathcal{B}$ est alors appelée *base de diagonalisation*.

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une telle base. On a donc $\operatorname{Mat}_{\mathcal{B}}(f) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, pour certains scalaires $\lambda_1, \dots, \lambda_n$. Par définition de $\operatorname{Mat}_{\mathcal{B}}(f)$, on a $f(b_j) = \lambda_j b_j$: les b_j sont donc des vecteurs propres de f et les λ_j les valeurs propres correspondantes. Réciproquement, si $\mathcal{B} = (b_1, \dots, b_n)$ est une base formée de vecteurs propres de f, la matrice de représentation de f dans $\mathcal{B}$ est diagonale, et ses éléments diagonaux sont les valeurs propres correspondants à $v_1,\dots,v_n$. On a montré:

Proposition 2.2. *Un endomorphisme f de E est diagonalisable si et seulement si E admet une base de vecteurs propres de f.*

Exemple 2.3.
(i) L'homothétie $h_a = a \operatorname{Id}_E$, où $a \in \mathbb{K}$, est trivialement diagonalisable. N'importe quelle base de E est une base de diagonalisation de h_a.
(ii) La rotation r_θ de $\mathbb{R}^2$ de centre 0 et d'angle θ, où $\theta \not\equiv 0\ [\pi]$ n'est pas diagonalisable: elle n'a pas de valeur propre.

(iii) Soit s la symétrie de $\mathbb{R}^2$ par rapport à l'axe Ox. Alors s est diagonalisable. La base canonique est une base de diagonalisation, et la matrice de s dans cette base est $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

Définition 2.4. Une matrice A est *diagonalisable* quand l'endomorphisme f_A de $\mathbb{K}^n$ représenté par A dans la base canonique de $\mathbb{K}^n$, $X \mapsto AX$ est diagonalisable.

Ceci signifie que dans une certaine base $\mathcal{B}$ de $\mathbb{K}^n$, la matrice de f_A est diagonalisale. Soit $\mathcal{B}_n$ la base canonique de $\mathbb{K}^n$. Alors la matrice de f_A dans la base $\mathcal{B}$ est donnée par $P^{-1}_{\mathcal{B}_n,\mathcal{B}} A \mathcal{P}_{\mathcal{B}_n,\mathcal{B}}$. On en déduit:

Proposition 2.5. *La matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable si et seulement si il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $P^{-1}AP$ est diagonale. La matrice P est alors la matrice de passage de la base canonique $\mathcal{B}_n$ à la base $\mathcal{B}$.*

Exercice 2.6. Soit $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right) \in \mathcal{M}_2(\mathbb{R})$. Calculer les valeurs propres et vecteurs propres de M. Montrer que M est diagonalisable et donner une matrice inversible P telle que $P^{-1}MP$ est diagonale.

Même question pour la matrice $R_\theta = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right) \in \mathcal{M}_2(\mathbb{C})$.

Dans la suite de cette section, tous les résultats sur les endomorphismes peuvent être bien sûr réinterprétés en terme matriciel.

2.b. Diagonalisation et polynôme caractéristique. On rappelle que le polynôme $Q \in \mathbb{K}[X]$ est scindé sur $\mathbb{K}$ si il a n racines sur $\mathbb{K}$ comptées avec leurs ordres de multiplicité. En d'autres termes, on a

$$Q(X) = a \prod_{1\leq j\leq d} (X - z_j),$$

où $a \neq 0$, et $z_1,\ldots,z_d$ sont des éléments de $\mathbb{K}$.

Lemme 2.7. *Soit $f \in \mathcal{L}(E)$ diagonalisable. Alors χ_f est scindé sur $\mathbb{K}$.*

Démonstration. En effet, dans une base $\mathcal{B}$ de diagonalisation, on a $\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) =$. Le polynôme caractéristique de f est donc

$$\prod_{j=1}^{n} (\lambda_j - X).$$

□

Exercice 2.8. Soit $A = \left(\begin{smallmatrix} 2 & -2 & 1 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Montrer que A n'est pas diagonalisable.

Le fait que χ_f est scindé n'est pas suffisant pour que f soit diagonalisable. Par exemple $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ n'est pas diagonalisable: sa seule valeur propre est 1, et on ne peut pas avoir $A = P^{-1}I_2P$ puisque ceci impliquerait $A = I_2$. Il existe toutefois une condition suffisante simple pour qu'une matrice soit diagonalisable en terme de polynôme caractéristique:

Théorème 2.9. *Soit f un endomorphisme dont le polynôme caractéristique est scindé à racines simples. Alors f est diagonalisable.*

Idée de la preuve. Par hypothèse, le polynôme caractéristique de f est de la forme $\prod_{j=1}^{n}(z_j - X)$, où les z_j sont deux à deux distincts. Les z_j sont donc des valeurs propres de f. Soit $(e_j)_{j=1,\ldots n}$ des vecteurs propres correspondants.

Montrons que e_j est une base de E: cela donnerait une base de E de vecteurs propres de f, impliquant que f est diagonalisable. Puisque $\mathcal{B}$ est de cardinal $n = \dim E$, il suffit pour cela de montrer que $\mathcal{B}$ est libre.

On se place dans le cas $n = 2$ pour donner l'idée de la preuve. On suppose donc

$$x_1 e_1 + x_2 e_2 = 0 \tag{2}$$

On doit montrer $x_1 = x_2 = 0$. En appliquant f à (2), on obtient

$$x_1 z_1 e_1 + x_2 z_2 e_2 = 0. \tag{3}$$

$z_1 \times$(2)-(3) donne $x_2(z_1 - z_2)e_2 = 0$, et donc $x_2 = 0$ puisque $z_1 \neq z_2$. En revenant à (2), on obtient aussi $x_1 = 0$.

Le cas général se démontre de manière similaire: en appliquant f^j à (3) pour $j = 0, \dots, n-1$ on obtient les n équations:

$$x_1 z_1^j e_1 + x_2 z_2^j e_2 + \dots + x_n z_n^j e_n = 0, \quad j = 0, \dots, n-1. \tag{4}$$

On démontre que la seule solution de ce système d'équations est $x_1 = x_2 = \dots = x_n = 0$ en utilisant l'inversibilité de la matrice de Vandermonde $[z_i^j]_{1 \leq \substack{i \\ j} \leq n}$. □

Exemple 2.10. Soit T une matrice triangulaire supérieure dont les éléments diagonaux sont distincts deux à deux. Alors T est diagonalisable.

2.c. Diagonalisation et sous-espaces propres. On rappelle que les sous-espaces $E_1, \dots, E_j$ de E sont *supplémentaires* dans E quand tout élément x de E s'écrit de manière unique $x_1 + x_2 + \dots + x_j$, avec $x_k \in E_k$ pour tout j entre 1 et k. On note alors $E = E_1 \oplus E_2 \oplus \dots \oplus E_j = \bigoplus_{k=1}^{j} E_k$.

Exercice 2.11. Montrer que E_1, E_2 et E_3 sont supplémentaires dans $\mathbb{R}^4$

$$E_1 = \operatorname{vect}\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right),\ E_2 = \operatorname{vect}\left(\begin{pmatrix}0\\0\\1\\-1\end{pmatrix}\right),\ E_3 = \operatorname{vect}\left(\begin{pmatrix}0\\0\\1\\1\end{pmatrix}\right).$$

Fixons pour tout $k \in \{1, \dots, j\}$, une base $\mathcal{B}_k$ de E_k. On peut montrer que les espaces $E_1, \dots, E_j$ sont supplémentaires si et seulement si $\bigcup_{k=1}^{j} \mathcal{B}_k$ est une base de E.

On en déduit que $E_1, \dots, E_j$ sont supplémentaires dans E si et seulement si 2 des 3 propriétés suivantes sont vérifiées:

(i) $E = E_1 + \dots + E_j$ (i.e. tout vecteur de E s'écrit $x_1 + x_2 + \dots + x_j$ avec $k \in E_k$, $k = 1, \dots, j$.

(ii)

$$\forall (x_1, \dots, x_j) \in E_1 \times \dots \times E_j, \quad x_1 + x_2 + \dots + x_j = 0_E \Longrightarrow x_1 = x_2 = \dots = x_j = 0_E.$$

(iii) $\dim E_1 + \dots + \dim E_j = \dim E$.

En particulier, deux sous-espaces E_1 et E_2 de E sont supplémentaires si et seulement si $\dim E_1 + \dim E_2 = \dim E$ et $E_1 \cap E_2 = \{0_E\}$.

Proposition 2.12. *L'endomorphisme f de E est diagonalisable si et seulement si ses sous-espaces propres sont supplémentaires dans E.*

Démonstration. En effet, si les sous-espaces propres de f sont supplémentaires dans E, on peut construire une base de vecteurs propres de E en choisissant une base de chaque sous-espace propre de E et en faisant la réunion de toutes ces bases.

Réciproquement, si f est diagonalisable, il existe une base $\mathcal{B}$ de vecteurs propres de E. Le sous espace propre associé à la valeur propre λ est l'espace vectoriel engendré par les $v \in \mathcal{B}$ tel que v est un vecteur propre associé à λ. On déduit facilement du fait que $\mathcal{B}$ est une base que ces sous-espaces propres sont supplémentaires. $\square$

On en déduit:

COROLLAIRE 2.13. *L'endomorphisme f est diagonalisable si et seulement si la somme des dimensions de ses sous-espaces propres est exactement la dimension de E.*

EXERCICE 2.14. Calculer les sous-espaces propres des matrices suivantes. Sont-elles diagonalisables?

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

2.d. Application de la diagonalisation au calcul des puissances successives d'une matrice. Si $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, ses puissances successives sont données par $D^j = \operatorname{diag}(\lambda_1^j, \dots, \lambda_n^j)$. Soit maintenant $A \in \mathcal{M}_n(\mathbb{K})$ une matrice diagonalisable. Il existe donc une matrice P inversible telle que $P^{-1}AP = D$ est une matrice diagonale. On en déduit

$$(P^{-1}AP)^j = D^j.$$

Or

$$(P^{-1}AP)^j = P^{-1}A^jP,$$

ce qui donne

$$A^j = PD^jP^{-1}.$$

On peut donc très facilement calculer les puissances d'une matrice qu'on a diagonalisée.

3. Quelques propriétés suppémentaires des valeurs propres

3.a. Ordre de multiplicité des valeurs propres. Soit $f \in \mathcal{L}(E)$ et $\lambda \in \operatorname{Sp}(E)$. λ est donc une racine du polynôme caractéristique χ_f. A λ, on associe deux notions de multiplicité:

DÉFINITION 3.1.
- La *multiplicité géométrique* $m_g(\lambda)$ est l'ordre de multiplicité de λ en tant que racine de χ_f.
- La *multiplicité algébrique* $m_a(\lambda)$ est la dimension du sous-espace propre $E_\lambda(f)$.

PROPOSITION 3.2. *Si $\lambda_0 \in \operatorname{Sp}(f)$,*

$$m_g(\lambda) \geq m_a(\lambda).$$

DÉMONSTRATION. En effet, soit $(e_1, \dots, e_m)$ une base de $E_{\lambda_0}(f)$ (avec $m = m_a(f)$). On complète cette base en une base $\mathcal{B}$ de E. ALors

$$\operatorname{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_0 I_m & B \\ 0 & C \end{pmatrix},$$

où B est une matrice $m \times (n-m)$, 0 est la matrice nulle $(n-m) \times m$ et C est une matrice $(n-m) \times (n-m)$. On voit alors que $(\lambda_0 - \lambda)^m$ divise $\chi_f(\lambda)$ ce qui montre le résultat. $\square$

Soit $f \in \mathcal{L}(E)$ et $\{\lambda_1, \dots, \lambda_k\}$ ses valeurs propres (comptées avec multiplicité). Supposons χ_f scindé, ce qui signifie exactement

$$m_g(\lambda_1) + \ldots + m_g(\lambda_p) = \dim E.$$

On rappelle que f est diagonalisable si et seulement si

$$\dim E_{\lambda_1} + \ldots \dim E_{\lambda_p} = \dim E,$$

i.e.

$$m_a(\lambda_1) + \ldots m_a(\lambda_p) = \dim E.$$

Par la proposition 3.2, ceci ne peut avoir lieu que si $m_g(\lambda_j) = m_a(\lambda_j)$ pour tout j. On a donc

THÉORÈME 3.3. *Un endormorphisme f est diagonalisable si et seulement si son polynôme caractéristique est scindé et la mutliplicité algébrique de chaque valeur propre de f est égale à sa multiplicité géométrique.*

3.b. Trace. Soit $A = [a_{ij}]_{1 \leq \substack{i \\ j} \leq n} \in M_n(\mathbb{K})$. On appelle *trace* de A, et on note $\operatorname{Tr} A$ la somme des coefficients diagonaux de A:

$$\operatorname{Tr} A = \sum_{i=1}^{n} a_{ii}.$$

Remarquons que $A \mapsto \operatorname{Tr} A$ est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$ (c'est à dire une application linéaire de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$).

THÉORÈME 3.4. *La trace de A ne dépend que du polynôme caractéristique de A. En conséquence, si P est une matrice inversible,* $\operatorname{Tr}(P^{-1}AP) = \operatorname{Tr} A$.

DÉMONSTRATION. Par définition,

$$\chi_A(\lambda) = \det(A - \lambda I_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} (a_{j\sigma(j)}) - \delta_{j,\sigma(j)}\lambda).$$

Supposons que $\sigma \in \mathcal{S}_n$ n'est pas l'identité de $\{1, \dots, n\}$. Il existe donc $k \in \{1, \dots, n\}$ tel que $\ell = \sigma(k) \neq k$. Mais alors $\sigma(\ell) \neq \ell$ aussi (sinon on aurait $\sigma(\ell) = \sigma(k)$, contredisant l'injectivité de σ). Il y a donc au moins deux indices k tels que $\sigma(k) \neq k$. Ceci montre que $\prod_{j=1}^{n}(a_{j\sigma(j)}) - \delta_{j,\sigma(j)}\lambda)$ est un polynôme de degré au plus $n-2$ en λ. Donc

$$\chi_A(\lambda) = \prod_{j=1}^{n} (a_{jj} - \lambda) + P_{n-2}(\lambda),$$

où P_{n-2} est un polynôme de degré au plus $n-2$. En développant, on voit que le coefficient de λ^{n-1} dans χ_A est exactement $(-1)^{n-1} \operatorname{Tr} A$, ce qui montre le théorème. $\square$

EXERCICE 3.5. Montrer que si $A, B \in \mathcal{M}_n(\mathbb{K})$, $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. On pourra commencer par le cas où $B = E_{ij}$ (la matrice dont tous les coefficients sont égaux à 0 sauf le coefficient b_{ij} qui vaut 1).

COROLLAIRE 3.6 (Trace d'un endomorphisme). *Si f est un endormophisme d'un espace vectoriel E de dimension finie,* $\operatorname{Tr} \operatorname{Mat}_{\mathcal{B}}(f)$ *ne dépend pas du choix de la base $\mathcal{B}$ de E et est appelée* trace *de f.*

Démonstration. Par le théorème, $\operatorname{Tr}\operatorname{Mat}_{\mathcal{B}}(f)$ ne dépend que du polynôme caractéristique de f (qui est égal au polynôme caractéristique de $\operatorname{Mat}_{\mathcal{B}}(f)$ pour n'importe quelle base $\mathcal{B}$, ce qui montre le résultat. □

Proposition 3.7. *Si le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ (respectivement $f \in \mathcal{L}(E)$) est scindé, alors $\operatorname{Tr} A$ (respectivement $\operatorname{Tr} f$) est la somme de ses valeurs propres, comptées avec leurs multiplicités géométriques.*

Démonstration. Soit $f \in \mathcal{L}(E)$. Si χ_f est scindé, alors

$$\chi_f(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda),$$

où les λ_i sont les valeurs propres de f (comptées avec multiplicité géométrique). On voit alors immédiatement que le coefficient de λ^{n-1} est exactement $(-1)^{n-1} \sum_{i=1}^{n} \lambda_i$, ce qui montre le résultat, puisque par la démonstration du théorème 3.4, cette quantité est aussi étale à $(-1)^{n-1} \operatorname{Tr} f$. (La démonstration est la même lorsque l'on remplace l'endomorphisme f par une matrice). □

Nous verrons plus loin qu'un endomorphisme dont le polynôme caractéristique est scindé est trigonalisable. En écrivant la matrice de f dans une base de trigonalisation, on retrouve immédiatement le résulat précédent.

Exercice 3.8. Soit p un projecteur (cf travaux dirigés). Quelle lien y a-t-il la trace de p et l'image de p?

4. Trigonalisation

4.a. Définition et caractérisation.

Définition 4.1. On dit que l'endomorphisme f de E est trigonalisable quand il existe une base $\mathcal{B}$ de E telle que $\operatorname{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure. $\mathcal{B}$ est alors appelée *base de trigonalisation.*

Si f est trigonalisable, en écrivant la matrice de représentation de f dans une base de trigonalisation, on voit que son polynôme caractéristique vérifie

$$\chi_f(\lambda) = \prod_{j=1}^{n} (\lambda_j - \lambda),$$

où les λ_j sont les valeurs propres de f (comptées avec multiplicité géométrique). Le polynôme χ_f est donc scindé. On admet la réciproque de cette propriété:

Théorème 4.2. *L'endomorphisme f est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

En particulier, tout endomorphisme d'un $\mathbb{C}$-espace vectoriel est trigonalisable.

Définition 4.3. La matrice $A \in \mathcal{M}_n(\mathbb{K})$ est *trigonalisable* si et seulement si l'endomorphisme $X \mapsto AX$ représenté par A dans la base canonique de $\mathbb{K}^n$ est trigonalisable.

Comme pour la diagonalisation, on démontre facilement que est trigonalisable si et seulement si un endomorphisme (respectivement tous les endomorphismes) représentant A dans une base $\mathcal{B}$ d'un espace vectoriel est trigonalisable. De manière équivalent, A est trigonalisable si et seulement si il existe une matrice $P \in \mathcal{M}_n(\mathbb{K})$, inversible, telle que $P^{-1}AP$ est triangulaire supérieure.

Il existe des méthodes systématiques de trigonalisation des matrices et des endomorphismes (les réductions de Dunford et de Jordan par exemple). Nous ne verrons pas ces méthodes dans le cas général. Nous allons en revanche, à titre d'exemple, montrer comment trigonaliser des endomorphismes d'espaces vectoriels en dimensions 2 et 3.

Dans toute la suite, f est un endomorphisme sur un espace vectoriel E de dimension finie n, que l'on suppose trigonalisable (i.e. χ_f est scindé).

4.b. Trigonalisation en dimension 2. Supposons $n = 2$.

4.b.1. Si les racines de χ_f sont simples, f est diagonalisable.

4.b.2. Si χ_f a une seule racine λ, qui est donc double, ou bien f est diagonalisable, et donc $\operatorname{Ker}(f - \lambda \operatorname{Id}_E) = E$, ce qui signifie $f = \lambda \operatorname{Id}_E$, ou bien f n'est pas diagonalisable. Dans ce dernier cas, une base de triagonalisation est donnée par $\mathcal{B} = (v_1, v_2)$ où v_1 est un vecteur propre associé à la valeur propre λ, et v_2 n'importe quel vecteur associé à λ. La matrice $\operatorname{Mat}_{\mathcal{B}}(f)$ est de la forme $\left(\begin{smallmatrix} \lambda & a \\ 0 & \lambda \end{smallmatrix}\right)$. On détermine $a \in \mathbb{K}$ en calculant $f(v_2) = av_1 + \lambda v_2$. On remarque que l'on peut toujours choisir $a = 1$ en remplaçant la base $\mathcal{B}$ par la base $\mathcal{B}' = (av_1, v_2)$.

4.c. Trigonalisation en dimension 3. On suppose $n = 3$.

4.c.1. Si les racines de χ_f sont simples, f est diagonalisable.

4.c.2. Si χ_f a une racine simple λ_1 et une racine double λ_2, f est diagonalisable si et seulement si $\operatorname{Ker}(f - \lambda \operatorname{Id}_E)$ est de dimension 2. Si cette dernière condition n'est pas vérifié, on peut facilement trigonaliser f. Une base de trigonalisation est donnée par (v_1, v_2, v_3), où v_1 est un vecteur propre associé à λ_1, v_2 un vecteur propre associé à λ_2 et v_3 n'importe quel vecteur qui n'est pas dans l'espace vectoriel engendré par (v_1, v_2). La matrice de f dans la base $\mathcal{B}$ est de la forme $\operatorname{Mat}_{\mathcal{B}}(f) = \left(\begin{smallmatrix} \lambda_1 & 0 & a \\ 0 & \lambda_2 & b \\ 0 & 0 & \lambda_2 \end{smallmatrix}\right)$, où a et b sont tels que $f(v_3) = av_1 + bv_2 + \lambda_2 v_3$. *Exercice:* trouver une base $\mathcal{B}'$ telle que $\operatorname{Mat}_{\mathcal{B}'}(f)$ est de la forme précédente avec $a = 0$, $b = 1$.

4.c.3. Si χ_f a une racine triple λ, la situation dépend de la dimension de $\operatorname{Ker}(f - \lambda \operatorname{Id}_E)$.

• Si $\dim \operatorname{Ker}(f - \lambda \operatorname{Id}_E = 3$, $f = \lambda \operatorname{Id}_E$, et donc f est trivialement diagonalisable.

• Si $\dim \operatorname{Ker}(f - \lambda \operatorname{Id}_E) = 2$, f n'est pas diagonalisable. Une base de trigonalisation est donnée par (v_1, v_2, v_3), où (v_1, v_2) est une base de $\operatorname{Ker}(f - \lambda \operatorname{Id}_E)$ et v_3 est n'importe quel vecteur de E qui n'est pas dans $\operatorname{Ker}(f - \lambda \operatorname{Id}_E)$. La matrice de f dans la base $\mathcal{B}$ est de la forme $\operatorname{Mat}_{\mathcal{B}}(f) = \left(\begin{smallmatrix} \lambda & 0 & a \\ 0 & \lambda & b \\ 0 & 0 & \lambda \end{smallmatrix}\right)$, où a et b sont tels que $f(v_3) = av_1 + bv_2 + \lambda v_3$. *Exercice:* trouver une base $\mathcal{B}'$ telle que $\operatorname{Mat}_{\mathcal{B}'}(f)$ est de la forme précédente avec $a = 0$, $b = 1$.

• Si $\dim \operatorname{Ker}(f - \lambda \operatorname{Id}_E = 1$, la trigonalisation est un tout petit peu plus compliquée. On va montrer que l'on peut trouver une base $\mathcal{B} = (v_1, v_2, v_3)$ telle que

$$\operatorname{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

ou de manière équivalente:

$$(f - \lambda \operatorname{Id}_E)v_1 = 0, \quad (f - \lambda \operatorname{Id}_E)v_2 = v_1, \quad (f - \lambda \operatorname{Id}_E)v_3 = v_2. \tag{5}$$

Pour cela, on commence par remarquer que la dimension de l'image de $f - \lambda \operatorname{Id}_E$ est 2 par le théorème du rang. En particulier, cette image ne peut pas être incluse dans le noyau de $f - \lambda \operatorname{Id}_E$, qui par hypothèse est de dimension 1. En d'autre termes, $(f - \lambda \operatorname{Id}_E)^2 \neq 0$. On se donne alors v_3 tel que $(f - \lambda \operatorname{Id}_E)^2(v_3) \neq 0$, et on pose, en suivant (5) $v_2 = (f - \lambda \operatorname{Id}_E)(v_3)$, $v_1 = (f - \lambda \operatorname{Id}_E)(v_2) = (f - \lambda \operatorname{Id}_E)^2(v_3)$. En utilisant ces définitions, et le fait que $v_1 \neq 0$ par le choix de v_3, on montre facilement que (v_1, v_2, v_3) est libre, et donc que c'est une base de E. Par construction, $\operatorname{Mat}_{\mathcal{B}}(f)$ est de la forme recherchée.

CHAPITRE 3

Espaces euclidiens

Ce chapitre est une version étendue du chapitre correspondant du polycopié écrit par Benoît Rittaud pour le cours 2019-2020.

À la fin de ce chapitre, vous devrez notamment savoir :

- Utiliser les propriétés des produits scalaires et des normes euclidiennes.
- Représenter matriciellement un produit scalaire.
- Manier les bases orthogonales et orthonormales.
- Appliquer le procédé d'orthonormalisation de Gram-Schmidt.
- Déterminer une distance à un sous-espace à partir d'une projection orthogonale.

1. Produits scalaires

1.a. Forme bilinéaire. On fixe un $\mathbb{K}$-espace vectoriel E, où $\mathbb{K}$ est pour l'instant n'importe quel corps (on peut prendre $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$ pour fixer les idées).

DÉFINITION 1.1. Une application $\varphi : E\times E\to\mathbb{K}$ est une *forme bilinéaire* sur E lorsque

- pour tout y de E, $x\mapsto\varphi(x,y)$ est linéaire et
- pour tout x de E, $y\mapsto\varphi(x,y)$ est linéaire.

De manière équivalente,

$$\forall(\alpha_1,\alpha_2,\beta_1,\beta_2)\in\mathbb{K}^4,\ \forall(u_1,u_2,v_1,v_2)\in E^4,\quad \varphi(\alpha_1u_1+\alpha_2u_2,\beta_1v_1+\beta_2v_2)$$
$$=\alpha_1\beta_1\varphi(u_1,v_1)+\alpha_1\beta_2\varphi(u_1,v_2)+\alpha_2\beta_1\varphi(u_2,v_1)+\alpha_2\beta_2\varphi(u_2,v_2).$$

EXEMPLES 1.2. La formule

$$\varphi_1(x,y)=\sum_{j=1}^n x_jy_j,\quad x=(x_1,\dots,x_n),\ y=(y_1,\dots y_n)$$

définit une forme bilinéaire sur $\mathbb{K}^n$.

L'application φ_2 de $\mathbb{R}^3\times\mathbb{R}^3$ dans $\mathbb{R}$ définie par

$$\varphi_2(x,y)=2x_1y_1+x_2y_3-x_3y_1+4x_3y_2$$

est une forme bilinéaire sur $\mathbb{R}^3$.

L'application définie sur $\mathbb{R}_n[X]\times\mathbb{R}_n[X]$ par

$$\varphi_3(P,Q)=\int_0^1 P(t)Q(t)dt$$

est une forme bilinéaire sur $\mathbb{R}_n[X]$.

1.b. Matrice associée à une forme bilinéaire. On suppose E de dimension finie, et on fixe une base $\mathcal{B}=(e_1,\dots,e_n)$ de E. Soient x, y deux vecteurs de E, et

$$X=(x_i)_{1\leq i\leq n},\quad Y=(y_i)_{1\leq i\leq n}$$

les vecteurs colonnes de leurs coordonnées dans $\mathcal{B}$. En utilisant la bilinéarité de φ, on obtient

$$\varphi(x,y)=\varphi\left(\sum_{i=1}^n x_ie_i,\sum_{j=1}^n y_je_j\right)=\sum_{i=1}^n x_i\varphi\left(e_i,\sum_{j=1}^n y_je_j\right)$$
$$=\sum_{1\leq {i \atop j}\leq n} x_iy_j\varphi(e_i,e_j)={}^tXAY,$$

où A est la matrice $[\varphi(e_i,e_j)]_{1\leq {i \atop j}\leq n}$.

Plus précisément, ${}^tXAY$ est une matrice à une ligne et une colonne, identifiée à un élément de $\mathbb{K}$.

Définition 1.3. La *matrice de la forme bilinéaire* φ dans la base $\mathcal{B}=(e_1,\dots,e_n)$ est la matrice carrée $[\varphi(e_i,e_j)]_{1\leq {i \atop j}\leq n}$.

Exemples 1.4. Soit $A\in\mathcal{M}_n(\mathbb{K})$. La formule

$$\varphi_A(X,Y)={}^tXAY$$

définit une forme bilinéaire sur $\mathcal{M}_{n,1}(\mathbb{K})$. Sa matrice dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$ est exactement la matrice A.

La matrice de la forme bilinéaire φ_1 de l'exemple 1.2 dans la base canonique de $\mathbb{R}^n$ est I_n.

La matrice de la forme bilinéaire φ_2 de l'exemple 1.2 dans la base canonique de $\mathbb{R}^3$ est $\begin{pmatrix} 2 & 0 & 0\\ 0 & 0 & 1\\ -1 & 4 & 0\end{pmatrix}$.

Exercice 1.5. Déterminer la matrice de la forme bilinéaire φ_3 de l'exemple 1.2 dans la base $(1,X,\dots,X^n)$.

Exercice 1.6. Déterminer la matrice de la forme bilinéaire φ_2 de l'exemple 1.2 dans la base $\left(\begin{pmatrix}-1\\2\\0\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ de $\mathbb{R}^3$.

Exercice 1.7. Soit φ la forme bilinéaire sur $\mathbb{R}^3$ de matrice $\begin{pmatrix}0 & 0 & 1\\ 0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix}$ dans la base canonique de $\mathbb{R}^3$. Déterminer $\varphi(x,y)$ en fonction des coordonnées (x_1,x_2,x_3) et (y_1,y_2,y_3) de x et y.

Comme pour les applications linéaires, on a une *formule de changement de bases* des matrices de formes bilinéaires. On se donne une forme bilinéaire φ sur le $\mathbb{K}$-espace vectoriel E de dimension finie, et deux bases $\mathcal{B}$ et $\widetilde{\mathcal{B}}$ de cet espace vectoriel. On note A (respectivement $\widetilde{A}$) la matrice de φ dans $\mathcal{B}$ (respectivement $\widetilde{B}$).

On note $P=P_{\mathcal{B},\widetilde{\mathcal{B}}}$ la matrice de passage de $\mathcal{B}$ à $\widetilde{\mathcal{B}}$. On rappelle que les vecteurs colonnes de cette matrice sont les coordonnées, dans la base $\mathcal{B}$, des vecteurs de $\widetilde{\mathcal{B}}$. Par ailleurs, la multiplication matricielle par P transforme les coordonnées d'un vecteur dans $\widetilde{\mathcal{B}}$ en ses coordonnées dans $\mathcal{B}$ (attention à l'inversion des deux bases par rapport au nom "matrice de passage de $\mathcal{B}$ à $\widetilde{\mathcal{B}}$).

Soit x et y deux vecteurs de E, X et Y (respectivement $\widetilde{X}$ et $\widetilde{Y}$) leurs coordonnées dans la base $\mathcal{B}$ (respectivement $\widetilde{\mathcal{B}}$). Alors

$$\varphi(x,y) = {}^tXAY = {}^t\widetilde{X}\widetilde{A}\widetilde{Y}.$$

D'autre part, puisque $X = P\widetilde{X}$ et $Y = P\widetilde{Y}$,

$${}^tXAY = {}^t(P\widetilde{X})AP\widetilde{Y} = {}^t\widetilde{X}\,{}^tPAP\widetilde{Y}.$$

On a donc

$$\forall \widetilde{X}, \widetilde{Y} \in \mathcal{M}_{n,1}(\mathbb{K}), \quad {}^t\widetilde{X}({}^tPAP)\widetilde{Y} = {}^t\widetilde{X}\widetilde{A}\widetilde{Y},$$

ce dont on déduit facilement (par exemple en applicant la formule précédentes aux vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$):

$$\widetilde{A} = {}^tPAP, \quad P = P_{\mathcal{B},\widetilde{\mathcal{B}}}.$$

C'est la *formule de changement de bases pour les formes bilinéaires*, à comparer à la formule analogue pour les applications linéaires.

Exercice 1.8. Vérifier le résultat de l'exerice 1.6 en utilisant la formule de changement de bases.

1.c. Forme bilinéaire symétrique.

Définition 1.9. La forme bilinéaire $\varphi : E \times E \to \mathbb{K}$ est *symétrique* quand

$$\forall (x,y) \in E^2, \quad \varphi(x,y) = \varphi(y,x).$$

Exemples 1.10. Reprenons les formes bilinéaires de l'exemple 1.2. Les formes bilinéaires φ_1 (sur $\mathbb{R}^n$) et φ_3 sur $\mathbb{R}_n[X]$, définies par

$$\varphi_1(x,y) = \sum_{j=1}^{n} x_j y_j, \quad \varphi_3(P,Q) = \int_0^1 P(t)Q(t)dt$$

sont symétriques. La forme bilinéaire φ_2 sur $\mathbb{R}^3$, définie par $\varphi_2(x,y) = 2x_1y_1 + x_2y_3 - x_3y_1 + 4x_3y_2$ n'est pas symétrique.

Proposition 1.11. *On suppose E de dimension finie. Soit $\mathcal{B}$ une base de E. La forme bilinéaire φ sur E est symétrique si et seulement si sa matrice dans la base $\mathcal{B}$ est symétrique.*

Démonstration. Soit A la matrice de φ dans $\mathcal{B}$. La forme bilinéaire φ est symétrique si et seulement si

$$\forall (x,y) \in E^2, \quad \varphi(x,y) = \varphi(y,x).$$

Soient X et Y deux éléments de $\mathcal{M}_{n,1}(\mathbb{K})$. Si x et y sont les éléments de E de coordonnées respectives X et Y dans $\mathcal{B}$, on a $\varphi(x,y) = {}^tXAY$ et $\varphi(y,x) = {}^tYAX$. On obtient donc que φ est symétrique si et seulement si

$$\forall (X,Y) \in \mathcal{M}_{n,1}(\mathbb{K})^2, \quad {}^tXAY = {}^tYAX.$$

Puisque ${}^tXAY$ est une matrice 1×1, elle est égale à sa transposée:

$${}^tXAY = {}^t({}^tXAY) = {}^tY{}^tA\,{}^t({}^tX) = {}^tY{}^tAX.$$

La matrice φ est donc symétrique si et seulement si

$$\forall (X,Y) \in \mathcal{M}_{n,1}(\mathbb{K})^2, \quad {}^tYAX = {}^tY{}^tAX,$$

ce qui signifie exactement $A = {}^tA$, i.e. que A est symétrique. □

1.d. Forme quadratique.

Définition 1.12. Soit φ une forme bilinéaire sur E. L'application $q : E \mapsto \mathbb{K}$ définie par $q(x) = \varphi(x, x)$ est appelée *forme quadratique associée à* φ. Une application $q : E \to \mathbb{K}$ est appelée une *forme quadratique* quand il existe une forme bilinéaire φ sur E telle que q soit la forme quadratique associée à φ.

Exemples 1.13. La forme quadratique associée à la forme bilinéaire définie par

$$\varphi_1(x, y) = \sum_{j=1}^{n} x_j y_j$$

est q_1 telle que

$$q_1(x) = \sum_{j=1}^{n} x_j^2, \quad x \in \mathbb{K}^n$$

La forme quadratique associée à la forme bilinéaire sur $\mathbb{K}^n$, $n \geq 2$, définie par

$$\varphi_4(x, y) = x_1 y_2 - x_2 y_1$$

est l'application constante égale à 0 sur $\mathbb{K}^n$.

Remarque 1.14. Il n'y a pas unicité de la forme bilinéaire correspondant à une forme quadratique q. De fait, on peut ajouter à une forme bilinéaire φ toute forme bilinéaire ψ telle que

$$\forall x \in E, \quad \psi(x, x) = 0$$

sans changer la forme quadratique correspondante. Une telle forme bilinéaire est dite alternée, et il en existe sur tous les espaces vectoriels de dimension finie supérieure à 2, comme le montre l'exemple précédent de φ_4.

Soit q une forme quadratique et φ sa forme bilinéaire associée. Soit $\lambda \in \mathbb{K}$ et $x \in E$, alors

$$q(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda^2 q(x).$$

On dit qu'une forme quadratique est *homogène de degré* 2. On a de plus l'identité remarquable suivante:

$$q(x + y) = \varphi(x + y, x + y) = q(x) + q(y) + \varphi(x, y) + \varphi(y, x). \tag{6}$$

En utilisant la même identité, où y est remplacé par $-y$, on en déduit *l'identité du parallèlogramme*:

$$q(x + y) + q(x - y) = 2(q(x) + q(y)).$$

De plus, si on suppose φ symmétrique, et que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ (ou n'importe quel corps $\mathbb{K}$ de caractéristique différente de 2, ce qui permet de diviser par 2 dans $\mathbb{K}$) (6) s'écrit:

$$\varphi(x, y) = \frac{1}{2} \left(q(x + y) - q(x) - q(y) \right). \tag{7}$$

Théorème 1.15. *On suppose $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit q une forme quadratique sur E. Alors il existe une* unique *forme bilinéaire symétrique φ telle que*

$$\forall x \in E, \quad \varphi(x, x) = q(x).$$

Cette forme bilinéaire φ est donnée par la formule (7).

DÉFINITION 1.16. Si q est une forme quadratique, la forme bilinéaire symétrique définie par (7) est appelée *forme polaire* de q.

DÉMONSTRATION. *Unicité.* Soit φ une forme bilinéaire symétrique associée à q. Alors φ doit vérifier (7), ce qui implique l'unicité.

Existence. Puisque q est une forme quadratique, il existe une forme bilinéaire ψ sur E telle que $q(x) = \psi(x,x)$ pour tout x de E. On définit φ par la formule (7). Alors d'une part

$$\varphi(x,x) = \frac{1}{2}\left(q(2x) - q(x) - q(x)\right) = \frac{1}{2}(4q(x) - 2q(x)) = q(x)$$

(où dans la deuxième inégalité, on a utilisé que q était homogène de degré 2). Et d'autre part, par la formule (7) définissant φ, et la formule (6) appliquée à ψ,

$$\varphi(x,y) = \frac{1}{2}\left(\psi(x,y) + \psi(y,x)\right), \tag{8}$$

ce qui montre que φ est bien bilinéaire (puisque ψ l'est). □

EXEMPLE 1.17. On considère sur $\mathbb{R}^3$ la forme quadratique définie par

$$q(x) = x_1x_2 + x_3^2.$$

Alors l'unique forme bilinéaire symétrique φ associée à q est donnée par

$$\varphi(x,y) = \frac{1}{2}(x_1y_2 - x_2y_1) + x_3^2.$$

REMARQUE 1.18. Soit ψ une forme bilinéaire sur E, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Alors la formule (8) définie une forme bilinéaire symétrique, appelée *symétrisée* de ψ, et qui est l'unique forme bilinéaire symétrique de même forme quadratique que ψ.

1.e. Forme bilinéaire positive. On suppose maintenant $\mathbb{K} = \mathbb{R}$.

DÉFINITION 1.19. La forme bilinéaire φ sur E (ou sa forme quadratique associée) est dite *positive* quand

$$\forall x \in E, \quad \varphi(x,x) \geq 0.$$

Elle est dite *définie positive* quand

$$\forall x \in E \setminus \{0_E\}, \quad \varphi(x,x) > 0.$$

Donc, si φ est définie positive, le seul vecteur de E tel que $\varphi(x,x) = 0$ est le vecteur nul.

EXEMPLE 1.20. La forme bilinéaire φ_1 définie sur $\mathbb{R}^n$ par $\varphi_1(x,y) = \sum_{j=1}^n x_jy_j$ est définie positive. La forme bilinéaire φ_3 définie sur $\mathbb{R}_n[X]$ par $\varphi_3(P,Q) = \int_0^1 P(t)Q(t)dt$ est définie positive. Les formes bilinéaires définies sur $\mathbb{R}^2$ par

$$\varphi_5(x,y) = x_1y_1 - x_2y_2, \quad \varphi_6(x,y) = x_1y_2$$

ne sont pas positives. En effet, $\varphi_5((0,1),(0,1)) = -1$ et $\varphi_6((-1,1),(-1,1)) = -1$. La forme bilinéaire définie sur $\mathbb{R}^2$ par $\psi(x,y) = x_1y_1$ est positive, mais pas définie positive (car $\psi((0,1),(0,1)) = 0$).

DÉFINITION 1.21. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *positive* quand

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \quad {}^tXAX \geq 0.$$

Elle est *définie positive* quand

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0\}, \quad {}^tXAX > 0.$$

En d'autres termes, les matrices positives (respectivement définies positives) sont les matrices des formes bilinéaires positives (respectivement définies positives).

1.f. Produit scalaire. On suppose encore $\mathbb{K} = \mathbb{R}$.

DÉFINITION 1.22. Un produit scalaire sur E est une forme bilinéaire symétrique définie positive.

EXEMPLES 1.23. Les exemples φ_1 (produit scalaire euclidien) et φ_3 (cf Exemples 1.2) sont des produits scalaires, respectivement sur $\mathbb{R}^n$ et $\mathbb{R}_n[X]$.

La forme bilinéaire ψ sur $\mathbb{R}^2$ de l'exemple précédent n'est pas un produit scalaire: elle est symétrique et positive, mais pas définie positive.

Un produit scalaire sur un $\mathbb{R}$-espace vectoriel E est souvent noté $(x, y) \mapsto x \cdot y$ ou $(x, y) \mapsto \langle x, y\rangle$. Le nombre $x \cdot y$ est appelé *produit scalaire des vecteurs x et y.*

2. Norme euclidienne

2.a. Norme.

DÉFINITION 2.1. Une *norme* sur un $\mathbb{R}$-espace vectoriel E est une application $N : E \to [0, +\infty[$ vérifiant les trois propriétés suivantes

(i) $\forall u \in E,\ N(u) = 0 \Longrightarrow u = 0_E$.
(ii) $\forall \lambda \in \mathbb{R},\ N(\lambda u) = |\lambda| N(u)$.
(iii) $\forall (u, v) \in E^2,\ N(u + v) \leq N(u) + N(v)$ (inégalité triangulaire).

Le couple (E, N), où N est une norme sur E est appelé *espace vectoriel normé.* Si $u \in E$, $N(u)$ est appelé norme du vecteur u.

On utilise souvent une notation de type 'valeur absolue", mais avec deux barres, pour dénoter une norme $\|\cdot\|$ (parfois avec un indice). La norme d'un vecteur u est donc notée $\|u\|$.

DÉFINITION 2.2. La *distance* de deux vecteurs u et v dans un espace vectoriel normé $(E, \|\cdot\|)$ est la quantité $\|u - v\|$.

EXEMPLE 2.3. Les formules suivantes définissent deux normes sur $\mathbb{R}^n$:

$$\|x\|_1 = \sum_{k=1}^{n} |x_k|, \quad \|x\|_\infty = \max_{1 \leq k \leq n} |x_k|.$$

EXERCICE 2.4. Démontrer que ce sont bien des normes.

2.b. Inégalité de Cauchy-Schwarz. Nous allons montrer que la racine carrée de la forme quadratique associée au produit scalaire définit une norme sur un espace euclidien. Pour cela, nous aurons besoin de la propriété fondamentale suivante:

THÉORÈME 2.5. *Soit E un $\mathbb{R}$-espace vectoriel et $\langle \cdot, \cdot \rangle$ un produit scalaire sur E. Alors*

$$\forall (x, y) \in E^2, \quad |\langle x, y\rangle|^2 \leq |\langle x, x\rangle| \, |\langle y, y\rangle| \,,$$

avec égalité si et seulement si x et y sont colinéaires.

DÉMONSTRATION. Notons pour simplifier $\|x\|^2 = \langle x, x\rangle$.

Si $y = 0_E$, la propriété énoncé est triviale. De fait, il y a égalité et x et y sont colinéaires.

Supposons $y \neq 0_E$. Alors $\|y\| > 0$ car le produit scalaire est défini positif. Pour tout $\lambda \in \mathbb{R}$, on a

$$0 \leq \|x + \lambda y\|^2 = \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2.$$

C'est un trinôme du second degré (puisque $\|y\| \neq 0$) en λ qui est toujours positif. Sont discriminant est donc négatif. Ceci signifie exactement:

$$\langle x, y\rangle^2 - \|y\|^2\|x\|^2 \leq 0,$$

qui est la propriété recherchée. Le discriminant est nul si et seulement si le trinôme a une racine, donc si et seulement si

$$\exists \lambda \in \mathbb{R}, \quad \|x + \lambda y\| = 0.$$

Ceci est bien équivalent au fait que x et y sont colinéaires. $\square$

Exercice 2.6. Vérifier que l'inégalité de Cauchy-Schwarz (mais pas le critère d'égalité) reste vrai si le produit scalaire est remplacée par une forme bilinéaire symétrique positive (non nécessairement définie positive).

Exercice 2.7. En examinant la preuve précédente, montrer que

$$\|x\|^2\|y\|^2 - \langle x, y\rangle^2 = \|z\|^2,$$

pour un certain $z \in E$ que l'on déterminera.

Exemple 2.8. Pour des réels $x_1, \ldots, x_n$, $y_1, \ldots, y_n$,

$$\Big|\sum_{j=1}^n x_j y_j\Big|^2 \leq \sum_{j=1}^n x_j^2 \sum_{j=1}^n y_j^2.$$

Si f et g sont des fonctions continues sur $[a, b]$, $a < b$,

$$\left|\int_a^b f(t)g(t)dt\right| \leq \sqrt{\int_a^b f^2(t)dt}\sqrt{\int_a^b g^2(t)dt}.$$

2.c. Norme euclidienne.

Théorème 2.9. *Soit $\big(E, \langle\cdot,\cdot,\rangle\big)$ un espace euclidien ou pré-hilbertien réel. On note $\|x\| = \sqrt{\langle x, x\rangle}$. Alors $\|\cdot\|$ est une norme sur E, appelée* norme euclidienne.

Démonstration. Remarquons que par positivité du produit scalaire, la quantité $\|x\|$ est bien définie comme racine carrée d'un nombre positif. Elle est positive par construction. De plus, le produit scalaire étant défini positif, on a

$$\|x\| = 0 \Longrightarrow \langle x, x\rangle = 0 \Longrightarrow x = 0_E.$$

Par ailleurs, si $\lambda \in \mathbb{R}$, par bilinéarité du produit scalaire,

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

Pour montrer l'inégalité triangulaire, on utilise l'inégalité de Cauchy-Schwarz:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

$\square$

EXEMPLE 2.10. La norme euclidienne usuelle sur $\mathbb{R}^n$ est la norme définie par:

$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

La distance définie par cette norme est la distante usuelle.

Il existe une infinité d'autres normes euclidiennes sur $\mathbb{R}^n$. Par exemple, si a_1, $\ldots$, a_n sont des nombres strictement positifs,

$$\|x\| = \sqrt{\sum_{j=1}^{n} a_j x_j^2}$$

définit une norme euclidenne sur $\mathbb{R}^n$.

EXERCICE 2.11. Dessiner la sphère unité $\{x \in \mathbb{R}^2,\ \|x\| = 1\}$ lorsque $\|\cdot\|$ est la norme

$$\|x\| = \sqrt{2x_1^2 + 3x_2^2}.$$

EXEMPLE 2.12. La formule suivant définit une norme euclidienne sur $\mathbb{R}[X]$ (ou sur tout sous-espace de $\mathbb{R}[X]$, par exemple $\mathbb{R}_n[X]$):

$$\|P\| = \sqrt{\int_0^1 P^2(x)dx}.$$

La même formule définit une norme sur l'espace vectoriel des fonctions continues sur $[0,1]$ à valeurs dans $\mathbb{R}$.

2.d. Quelques égalités et inégalités. Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien ou pré-hilbertien réel. On rappelle les formules suivantes, que l'on peut démontrer par calcul direct:

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

(Les formules de la dernière ligne, déjà vues, sont des égalités de polarisation, permettant de retrouver le produit scalaire à partir de la norme euclidienne). On a aussi les deux inégalités triangulaires:

$$\|x+y\| \leq \|x\| + \|y\|, \quad |\|x\| - \|y\|| \leq \|x-y\|.$$

La première ce ces inégalités à été démontrée à l'aide de l'inégalité de Cauchy-Schwarz. La deuxième découle de la première, appliquée à $X = x - y$, $Y = y$ puis à $X = y - x$, $Y = x$.

3. Orthogonalité

3.a. Définition. Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien ou pré-hilbertien réel.

DÉFINITION 3.1. Les vecteurs u et v de E sont dits *orthogonaux* lorsque $\langle u, v\rangle = 0$. On note alors $u \perp v$. Deux sous-ensembles U et V de E sont dits orthogonaux lorsque

$$\forall u \in U,\ \forall v \in V, \quad u \perp v.$$

On note de nouveau $U \perp V$. Enfin, une famille $\mathcal{F}$ de vecteurs de E est dite orthogonale lorsque les éléments de $\mathcal{F}$ sont orthogonaux deux à deux.

Exemple 3.2. Les vecteurs $u = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ sont orthogonaux dans $\mathbb{R}^2$, muni du produit scalaire usuel. Les vecteurs u et $w = \left(\begin{smallmatrix}-1\\0\end{smallmatrix}\right)$ ne le sont pas.

On munit maintenant $\mathbb{R}^2$ du produit scalaire

$$\langle x, y\rangle_s = (x_1 + x_2)(y_1 + y_2) + x_2 y_2.$$

Pour ce produit scalaire, les vecteurs u et v ne sont pas orthogonaux, mais les vecteurs u et w le sont.

On munit $\mathbb{R}[X]$ du produit scalaire défini par $\langle P, Q\rangle = \int_0^1 P(x)Q(x)dx$. Alors X et $\frac{2}{3} - X$ sont orthogonaux.

Exercice 3.3. On munit $\mathbb{R}[X]$ du même produit scalaire que dans l'exemple précédent. A quelle condition sur a les polynômes X^2 et $X-a$ sont-ils orthogonaux?

3.b. Théorème de Pythagore.

Théorème 3.4. *Soit $(u_1, \dots, u_p)$ une famille de vecteurs orthogonaux. Alors*

$$\|u_1 + \dots + u_p\|^2 = \|u_1\|^2 + \dots + \|u_p\|^2.$$

Démonstration. Ceci découle immédiatement de l'orthogonalité des vecteurs et de la bilinéarité du produit scalaire:

$$\|u_1 + \dots + u_p\|^2 = \left\langle \sum_{j=1}^{p} u_j, \sum_{k=1}^{p} u_k \right\rangle = \sum_{1 \leq \substack{j \\ k} \leq p} \underbrace{\langle u_j, u_k\rangle}_{0 \text{ si } j\neq k} = \sum_{j=1}^{p} \langle u_j, u_j\rangle = \sum_{j=1}^{p} \|u_j\|^2.$$

□

Exemple 3.5. Soit ABC un triangle rectangle en A. Alors

$$BC^2 = AB^2 + AC^2.$$

C'est le théorème de Pythagore usuel, que l'on obtient en appliquant le théorème 3.4 aux vecteurs $\overrightarrow{BA}$ et $\overrightarrow{AC}$.

Exemple 3.6. Appliquons le théorème de Pythagore aux familles de 2 vecteurs de l'exemple 3.2:

$$\left\|\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)\right\|^2 + \left\|\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\right\|^2 = \left\|\left(\begin{smallmatrix}2\\0\end{smallmatrix}\right)\right\|^2 = 4$$

(où $\|\cdot\|$ désigne la norme euclidienne usuelle de $\mathbb{R}^2$).

$$\left\|\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)\right\|_s^2 + \left\|\left(\begin{smallmatrix}-1\\0\end{smallmatrix}\right)\right\|_s^2 = \left\|\left(\begin{smallmatrix}0\\-1\end{smallmatrix}\right)\right\|_s^2 = 2.$$

$$\int_0^1 t^2 dt + \int_0^1 \left(\frac{2}{3} - t\right)^2 dt = \int_0^1 \frac{4}{9} dt = \frac{4}{9}.$$

Exemple 3.7. La grande diagonale d'un cube de côté c (reliant deux sommets opposés) a pour longueur:

$$\sqrt{c^2 + c^2 + c^2} = c\sqrt{3}.$$

Une conséquence importante du théorème de Pythagore est la suivante:

Corollaire 3.8. *Soit $(u_1, \dots, u_p)$ une famille orthogonale* sans vecteur nul. *Alors $(u_1, \dots, u_p)$ est libre.*

Démonstration. Soit $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$ tel que

$$\sum_{j=1}^{p} \lambda_j u_j = 0.$$

On remarque que la famille $(\lambda_1 u_1, \dots, \lambda_p u_p)$ est orthogonale. Par le théorème de Pythagore,

$$0 = \Big\| \sum_{j=1}^{p} \lambda_j u_j \Big\|^2 = \sum_{j=1}^{p} \lambda_j^2 \|u_j\|^2.$$

On en déduit (puisque par hypothèse $\|u_j\| \neq 0$ pour tout j) $\lambda_1 = \ldots = \lambda_p = 0$. On a bien montré que la famille était libre. □

3.c. Orthogonal d'une partie de E.

Définition 3.9. Soit A un sous-ensemble ou une famille de vecteurs de E. L'orthogonal de A, noté $A^\perp$, est l'ensemble

$$A^\perp = \{x \in E,\ \forall y \in A,\ \langle x, y\rangle = 0\}.$$

Proposition 3.10. *L'orthogonal de A est un sous-espace vectoriel de E.*

Démonstration. Soit $(x, y) \in A^\perp$ et $(\lambda, \mu) \in \mathbb{R}^2$. Soit $z \in A$. Alors

$$\langle \lambda x + \mu y, z\rangle = \lambda \langle x, z\rangle + \mu \langle y, z\rangle = 0.$$

On en déduit $\lambda x + \mu y \in A^\perp$. □

Exemple 3.11. Dans $\mathbb{R}^3$ muni du produit scalaire usuel,

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right\}^\perp = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ x - y + 2z = 0 \right\}$$

qui est bien un sous-espace vectoriel de $\mathbb{R}^3$.

Dans les deux exercices suivants, on se place de nouveau dans $\mathbb{R}^3$ muni du produit scalaire usuel.

Exercice 3.12. Donner une base de

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right\}^\perp.$$

Exercice 3.13. Quel est l'orthogonal de $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ x > 0,\ y < 0 \right\}$?

Proposition 3.14. *Soit A et B deux parties de E. Alors*

$$B \subset A \Longrightarrow A^\perp \subset B^\perp.$$

Démonstration. Soit $x \in A^\perp$ et $y \in B$. Puisque $B \subset A$, on a $y \in A$. Donc

$$\langle x, y\rangle = 0.$$

On a montré $\forall y \in B$, $\langle x, y\rangle = 0$, ce qui montre $x \in B^\perp$. □

Exercice 3.15. Soit $A \subset E$. Montrer que $A \subset (A^\perp)^\perp$. Calculer $E^\perp$, $\{0_E\}^\perp$. Montrer que si $u \in E$, $u^\perp = (\operatorname{vect} u)^\perp$.

La dernière question de l'exercice précédent se généralise de la manière suivante:

Proposition 3.16. *Soit F un sous-espace vectoriel de E de dimension finie p, et $\mathcal{B} = (f_1, \dots, f_p)$ une base de F. Alors*

$$F^\perp = \{f_1, \dots, f_p\}^\perp .$$

Démonstration. Puisque $\{f_1, \dots, f_p\} \subset F$, on a

$$F^\perp \subset \{f_1, \dots, f_p\}^\perp$$

par la proposition précédente. Montrons l'autre inclusion. Soit $x \in \{f_1, \dots, f_p\}^\perp$. Soit $y \in F$, et $(y_1, \dots, y_p)$ ses coordonnées dans la base $\mathcal{B}$. Alors

$$\langle x, y\rangle = \left\langle x, \sum_{j=1}^{p} y_j f_j \right\rangle = \sum_{j=1}^{p} y_j \langle x, f_j\rangle.$$

Puisque $x \in \{f_1, \dots, f_p\}^\perp$, les produits scalaires $\langle x, f_j\rangle$ sont tous nuls. On a donc bien $\langle x, y\rangle = 0$. Le vecteur y étant un élément quelconque de F, on a montré $x \in F^\perp$. $\square$

Remarque 3.17. Par la même démonstration, on obtient que pour toute partie A de E,

$$(\operatorname{vect} A)^\perp = A^\perp.$$

4. Base orthogonale

On se place dans un espace *euclidien* $(E, \langle\cdot,\cdot\rangle)$.

4.a. Définition.

Définition 4.1. Une *base orthogonale* d'un espace euclidien E est une base $\mathcal{B}$ de E dont les éléments sont orthogonaux 2 à 2. Si de plus tous ces éléments sont de norme 1, on dit que $\mathcal{B}$ est une *base orthonormale*.

Remarque 4.2. Soit E un espace euclidien de dimension n. Alors toute famille orthogonale de cardinal n sans vecteur nul est une base orthogonale de E. En effet, par le Corollaire 3.8, une telle famille est libre. On en déduit que c'est une base parce que son cardinal est égal à la dimension de E.

Exemple 4.3. La base canonique de $\mathbb{R}^n$ est une base orthonormale pour le produit scalaire euclidien usuel.

La base $(1, X, \dots, X^n)$ de $\mathbb{R}_n[X]$ n'est pas orthogonale pour le produit scalaire défini par $\langle P, Q\rangle = \int_0^1 P(t)Q(t)dt$.

La famille $(X, \frac{2}{3} - X)$ est une base orthogonale de $\mathbb{R}_1[X]$ pour le produit scalaire précédent. Ce n'est pas une base orthonormale.

4.b. Coordonnées dans une base orthonormale.

Proposition 4.4. *Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de l'espace euclidien E et $x \in E$. Les coordonnées de x dans $(e_1, \dots, e_n)$ sont données par $\big(\langle x, e_j\rangle\big)_{1\leq j\leq n}$.*

Démonstration. Soit $(x_j)_{1\leq j\leq n}$ les coordonnées de x dans E. Alors

$$x = \sum_{k=1}^{n} x_k e_k.$$

On fixe $j \in \{1, \dots, n\}$ et on fait le produit scalaire de cette égalité avec e_j. On obtient
$$\langle x, e_j \rangle = \sum_{k=1}^{n} x_k \langle e_k, e_j \rangle = x_j,$$
puisque $\langle e_k, e_j \rangle = \delta_{j,k}$. □

EXEMPLE 4.5. Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, et $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Alors on a bien
$$x_j = \langle x, e_j \rangle,$$
(où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien usuel).

COROLLAIRE 4.6. *Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale d'un espace euclidien E, et $x \in E$. Alors*
$$\|x\|^2 = \sum_{j=1}^{n} \langle x, e_j \rangle^2.$$

DÉMONSTRATION. On a
$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j,$$
et la famille de vecteurs $\Big(\langle x, e_j \rangle e_j\Big)_{1 \leq j \leq n}$ est orthogonale. Le résultat découle alors du théorème de Pythagore. □

EXERCICE 4.7. Comme dans l'exemple 3.2, on munit maintenant $\mathbb{R}^2$ du produit scalaire
$$\langle x, y \rangle_s = (x_1 + x_2)(y_1 + y_2) + x_2 y_2.$$
Soit $u = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ et $w = \left(\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right)$.

Montrer que (u, w) est une base orthonormale de E. Donner les coordonnées d'un vecteur $x = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ de $\mathbb{R}^2$ dans cette base. Calculer $\|x\|_s^2$.

REMARQUE 4.8. Soit $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ une base orthogonale de E. Alors $\mathcal{B}' = \left(\frac{e_j}{\|e_j\|}\right)_{1 \leq j \leq n}$ est une base orthonormale de E. Si $x \in E$, les coordonnées de x dans $\mathcal{B}'$ sont $\left(\frac{\langle e_j, x \rangle}{\|e_j\|}\right)_{1 \leq j \leq n}$. En d'autres termes:
$$x = \sum_{j=1}^{n} \frac{\langle e_j, x \rangle}{\|e_j\|^2} e_j.$$
Les coordonnées de x dans $\mathcal{B}$ sont donc $\left(\frac{\langle e_j, x \rangle}{\|e_j\|^2}\right)_{1 \leq j \leq n}$.

EXERCICE 4.9. Dans le contexte de la remarque précédente, Calculer $\|x\|^2$.

EXERCICE 4.10. On se place dans l'espace E des fonctions 2π-périodiques f sur $\mathbb{R}$ qui sont de la forme
$$f(x) = a + b \cos x + c \sin x.$$
Pour $(f, g) \in E$, on pose
$$\langle f, g \rangle = \int_0^{2\pi} f(x) g(x) dx.$$

a. Montrer que c'est un produit scalaire sur E, et que $(1, \cos, \sin)$ est une base orthogonale pour ce produit scalaire.

b. En déduire une base orthonormale de E. Si $f = a + b\cos + c\sin$, exprimer a, b, c et $\|f\|$ en utilisant des produits scalaires avec les fonctions 1, cos et sin.

4.c. Procédé d'orthogonalisation de Gram-Schmidt. Un espace euclidien admet-il toujours une base orthonormale? La réponse _positive_ est donnée par le procédé d'orthogonalisation de Gram-Schmidt:

THÉORÈME 4.11. *Soit* $(e_1, \ldots, e_p)$ *une famille libre d'un espace euclidien (ou pré-hilbertien réel)* E. *Alors il existe une famille libre* orthogonale $(f_1, \ldots, f_p)$ *telle que*

$$\forall j \in \{1, \ldots, p\}, \quad f_j \in \operatorname{vect}(e_1, \ldots, e_j).$$

Si $(e_1, \ldots, e_p)$ est une base de E, $(f_1, \ldots, f_p)$ est une famille libre orthogonale de E ayant $p = \dim E$ éléments, donc une base orthogonale de E. Quitte à normaliser les vecteurs f_j, on obtient:

COROLLAIRE 4.12. *Tout espace euclidien admet une base orthonormale.*

DÉMONSTRATION. La démonstration est importante, puisqu'elle donne une méthode de construction de la famille orthogonale $(f_j)_{1\leq j\leq p}$.

On construit la famille libre orthogonale $(f_j)_{1\leq j\leq p}$ par récurrence, de la manière suivante.

On pose $f_1 = e_1$.

Soit $j \geq 2$. Supposons construits $(f_1, \ldots, f_{j-1})$ avec les propriétés demandées. On cherche f_j de la forme

$$f_j = e_j + \sum_{k=1}^{j-1} \alpha_k f_k.$$

On veut $\langle f_j, f_\ell\rangle = 0$ pour tout entier ℓ entre 1 et $j-1$. En d'autres termes, en utilisant l'orthogonalité des vecteurs $(f_k)_{1\leq k\leq j-1}$,

$$\forall \ell \in \{1, \ldots, j-1\}, \quad 0 = \langle e_j, f_\ell\rangle + \alpha_\ell \|f_\ell\|^2,$$

c'est à dire $\alpha_\ell = -\frac{\langle e_j, f_\ell\rangle}{\|f_\ell\|^2}$. On obtient donc le vecteur

$$f_j = e_j - \sum_{k=1}^{j-1} \frac{\langle e_j, f_k\rangle}{\|f_k\|^2} f_k,$$

qui vérifie les propriétés demandées.

Pour une interprétation de f_j en terme de projection, voir la remarque 4.18 plus bas. □

Un exemple important d'espace euclidien est donné par un sous-espace vectoriel F d'un espace vectoriel euclidien E, muni de la restriction du produit scalaire de E à $F \times F$. Le procédé d'orthogonalisation de Gram-Schmidt peut être utilisé dans ce contexte, par exemple pour construire des bases orthogonales de sous-espaces vectoriels de $\mathbb{R}^n$ (muni du produit scalaire usuel), comme dans l'exercice suivant.

EXERCICE 4.13. Construire une base orthonormale (pour le produit scalaire usuelle de $\mathbb{R}^4$) de

$$F = \operatorname{vect}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right).$$

Donnons un deuxième exemple d'utilisation:

EXERCICE 4.14. On munit $\mathbb{R}_2[X]$ du produit scalaire

$$\langle P, Q\rangle = \int_0^1 P(t)Q(t)dt.$$

Construire une base orthogonale de $\mathbb{R}_2[X]$ à l'aide du procédé d'orthogonalisation de Gram-Schmidt à partir de la base $(1, X, X^2)$

4.d. Projection orthogonale.

THÉORÈME 4.15. *Soit F un sous espace vectoriel de dimension fini d'un espace euclidien ou préhilbertien réel E. Alors $F^\perp \oplus F = E$. La projection π_F sur F parallèlement à $F^\perp$ est appelée projection orthogonale sur F. Si $(e_1, \dots, e_p)$ est une base* orthonormale *de F, on a*

$$\forall x \in E, \quad \pi_F(x) = \sum_{j=1}^{p} \langle e_j, x\rangle e_j.$$

DÉMONSTRATION. Il est facile de voir que $F \cap F^\perp = \{0_E\}$. En effet, si $x \in F \cap F^\perp$, on a $\langle x, x\rangle = 0$, ce qui implique $x = 0_E$.

Il reste à montrer $F + F^\perp = E$. On fixe une base orthonormale $(e_1, \dots, e_p)$ de F (une telle base existe par le procédé d'orthogonalisation de Gram-Schmidt et puisque F est de dimension finie). Pour $x \in E$, on pose

$$\varphi(x) = \sum_{k=1}^{p} \langle e_k, x\rangle e_k \in F.$$

Soit $y = x - \varphi(x)$. Montrons $y \in F^\perp$. Pour cela il suffit de vérifier $\forall j \in \{1, \dots, p\}$, $\langle y, e_j\rangle = 0$ (cf Proposition 3.16). Or

$$\langle y, e_j\rangle = \langle x, e_j\rangle - \sum_{k=1}^{p} \langle e_k, x\rangle\langle e_k, e_j\rangle = \langle x, e_j\rangle - \sum_{k=1}^{p} \langle e_k, x\rangle \delta_{j,k} = 0.$$

Finalement, on a

$$x = \underbrace{y}_{\in F^\perp} + \underbrace{\varphi(x)}_{\in F},$$

ce qui montre comme annoncé que $E = F \oplus F^\perp$, et que φ est la projection sur F parallèlement à $F^\perp$. □

REMARQUE 4.16. On a bien évidemment $\pi_{F^\perp} = \mathrm{Id}_E - \pi_F$.

REMARQUE 4.17. Lorsque E est un espace euclidien, le théorème implique $\dim E = \dim F + \dim F^\perp$. Par ailleurs, en l'appliquant le théorème à $F^\perp$, on obtient

$$E = F^\perp \oplus (F^\perp)^\perp.$$

Donc

$$\dim E = \dim F^\perp + \dim(F^\perp)^\perp,$$

ce qui montre que $\dim F = \dim(F^\perp)^\perp$. Puisque $F \subset (F^\perp)^\perp$, on en déduit

$$F = (F^\perp)^\perp.$$

Attention ce raisonnement n'est valable que lorque la dimension de E est finie!

REMARQUE 4.18. Revenons à la preuve du Théorème 4.11. Fixons $j \in \{1, \dots, p\}$ et posons

$$F_j = \mathrm{vect}(f_j, \dots, f_p).$$

La base $(f_k)_{1 \le k \le p}$ étant orthogonale, on a

$$F_j^{\perp} = \mathrm{vect}(f_1, \dots, f_{j-1}) = \mathrm{vect}(e_1, \dots, e_{j-1}),$$

et donc, par la remarque précédente,

$$F_j = (F_j^{\perp})^{\perp} = \mathrm{vect}(e_1, \dots, e_{j-1})^{\perp}.$$

De plus,

$$e_j = \underbrace{f_j}_{\in F_j} + \underbrace{\sum_{k=1}^{j-1} \frac{\langle e_j, f_k \rangle}{\|f_k\|^2} f_k}_{\in F_j^{\perp}}$$

En d'autres termes, f_j est exactement la projection orthogonale de e_j sur $\mathrm{vect}(e_1, \dots, e_{j-1})^{\perp}$.

EXERCICE 4.19. Soit $E = \mathbb{R}^4$ (muni du produit scalaire usuel). On pose

$$F = \mathrm{vect}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

Donner $F^{\perp}$ et π_F (la projection orthogonale sur F).

EXERCICE 4.20. Calculer la projection orthogonale sur le sous-espace F de l'exemple 4.13.

Terminons ce chapitre par une caractérisation de la projection orthogonale sur un espace vectoriel en terme de distance:

THÉORÈME 4.21. *Soit F un sous-espace vectoriel de l'espace euclidien E et $u \in E$. Alors la fonction*

$$x \mapsto \|x - u\|$$

sur F admet un unique minimum global, atteint en $x = \pi_F(u)$.

DÉMONSTRATION. On a

$$u = \pi_F(u) + \pi_{F^{\perp}}(u).$$

Par le théorème de Pythagore, si $x \in F$,

$$\|u - x\|^2 = \|\pi_F(u) - x\|^2 + \|\pi_{F^{\perp}}(u)\|^2.$$

On voit que cette quantité est minimale quand $\|\pi_F(u) - x\|$ est minimal, c'est à dire quand $x = \pi_F(u)$. $\square$