Notes de cours - Algèbre Linéaire

Table des matières

Chapitre 1

Introduction

1.1

Chapitre 2

Les Bases de l'algèbre linéaire

2.1 Espaces vectoriels

C'est Giuseppe Peano, vers la fin du 19ème siècle, qui dégage le premier les notions d'espaces vectoriels et d'applications linéaires abstraites que nous étudions dans ce cours.

Les éléments d'un espace vectoriels sont appelés vecteurs. Comme les vecteurs de $\mathbb{R}^2$ ou $\mathbb{R}^3$, on peut additionner les vecteurs et les multiplier par un nombre. Ces opérations vérifient quelques propriétés que nous allons isoler dans la définition suivante.

Definition 1. *Un $\mathbb{R}$-espace vectoriel est un triplet $(E, +, .)$ formé d'un ensemble E dont les éléments sont appelés vecteurs, d'une loi d'addition, notée $+$, qui est une application $E \times E \to E$ qui à deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ de E associe un vecteur $\boldsymbol{u} + \boldsymbol{v} \in E$, qu'on appellera somme des deux vecteurs, et d'une loi de multiplication par un scalaire notée $\cdot$ qui est une application $\mathbb{R} \times E \to E$ qui associe à un nombre réel λ et un vecteur $\boldsymbol{u}$ le vecteur $\lambda\boldsymbol{u}$ qu'on appellera produit du vecteur $\boldsymbol{u}$ par le réel λ.*

Axiomes de la somme :

1. **Associativité** *pour tous $\boldsymbol{u}$, $\boldsymbol{v}$ et $\boldsymbol{w}$ de E, $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$.*
2. **Commutativité** *pour tous $\boldsymbol{u}$ et $\boldsymbol{v}$ de E, $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$.*
3. **Neutre** *il existe un élément de E noté 0 tel que, pour tout $\boldsymbol{u}$ de E, $\boldsymbol{u} + 0 = \boldsymbol{u}$.*
4. **Opposé** *Pour tout $\boldsymbol{u}$ de E, il existe $\boldsymbol{v} \in E$ tel que $\boldsymbol{u} + \boldsymbol{v} = 0$.*

Axiomes de compatibilité pour la multiplication :

1. **produit de réels** *pour tout $\boldsymbol{u}$ de E, λ et μ de $\mathbb{R}$, $(\lambda\mu)\boldsymbol{u} = \lambda(\mu\boldsymbol{u})$.*
2. **Somme de réels** *pour tout $\boldsymbol{u}$ de E, λ et μ de $\mathbb{R}$, $(\lambda + \mu)\boldsymbol{u} = \lambda\boldsymbol{u} + \mu\boldsymbol{u}$.*
3. **Somme de vecteurs** *pour tous $\boldsymbol{u}$ et $\boldsymbol{v}$ de E, λ de $\mathbb{R}$, $\lambda(\boldsymbol{u}+\boldsymbol{v}) = \lambda\boldsymbol{u}+\lambda\boldsymbol{v}$.*
4. **Unité** *Pour tout $\boldsymbol{u}$ de E, $1\boldsymbol{u} = \boldsymbol{u}$.*

Conséquences immédiates

— L'associativité permet d'éviter de mettre des parenthèses dans les sommes de vecteurs.

— La commutativité permet d'échanger les termes d'une somme.
— On a $0 + \mathbf{u} = \mathbf{u}$ pour tout $\mathbf{u} \in E$.
— L'opposé de $\mathbf{u}$ est noté $-\mathbf{u}$ et $\mathbf{u} + (-\mathbf{v})$ est noté $\mathbf{u} - \mathbf{v}$ de sorte qu'on a, pour tout $\mathbf{u}$, $\mathbf{u} - \mathbf{u} = 0$.
— Si $\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{w}$, alors $\mathbf{v} = \mathbf{w}$.
— $0\mathbf{u} = 0$.
— $\lambda 0 = 0$.
— $(-1)\mathbf{u} = -\mathbf{u}$.
— $2\mathbf{u} = \mathbf{u} + \mathbf{u}$.

2.1.1 Exemples d'espaces vectoriels

$\mathbb{R}^n$ muni de la somme de vecteurs et du produit par un réel.
L'ensemble des fonctions d'un ensemble I dans $\mathbb{R}$, muni de la somme de fonctions et de la multiplication par un réel.
L'ensemble des suites à valeurs réelles.
L'espace $\mathbb{R}[X]$ des polynômes à coefficients réels, qu'on identifiera aux fonctions de la forme $P(x) = \sum_{k=0}^{n} a_k x^k$.
L'ensemble des fonctions d'un ensemble X à valeurs dans un espace vectoriel E.
Exercice : vérifier que ces ensembles vérifient bien les axiomes définissant les espaces vectoriels.

2.1.2 Espaces vectoriels sur des corps plus généraux

Nous nous limiterons dans ce cours, à de rares exceptions près, aux espaces vectoriels sur $\mathbb{R}$. Les nombres réels s'additionnent et se multiplient, les deux lois sont associatives et commutatives, il existe des éléments 0 et 1 qui sont des éléments neutres pour l'addition et la multiplication, tout réel a un opposé pour l'addition et tout réel non nul a un inverse pour la multiplication.

Il existe d'autres ensembles vérifiant ces propriétés, comme l'ensemble des nombres complexes ou l'ensemble des nombres rationnels. Dedekind donne à de tels ensembles le nom de corps. Les résultats que nous allons énoncer pour les espaces vectoriels sur $\mathbb{R}$ restent valables si $\mathbb{R}$ est remplacé par un autre corps K.

2.1.3 Sous-espaces vectoriels

Definition 2. *Soit E un espace vectoriel et F un sous-ensemble de E. Si F est non vide et vérifie*

1. *Pour tous $\boldsymbol{u}$ et $\boldsymbol{v}$ de F et tout $\lambda \in \mathbb{R}$, $\boldsymbol{u} + \lambda \boldsymbol{v} \in F$.*

Alors F est un espace vectoriel appelé sous-espace vectoriel de E.

Exercice : Vérifier que F est un espace vectoriel.

Proposition 3. *L'intersection d'une famille de sous-espace vectoriels est un sous-espace vectoriel.*

Preuve : Exercice. □

2.1.4 Exemples de sous-espaces vectoriels

Sous-espaces de l'ensemble des fonctions de I dans $\mathbb{R}$.
— L'ensemble des fonctions continues de I dans $\mathbb{R}$.
— L'ensemble des fonctions de classe C^k de I dans $\mathbb{R}$.
— L'ensemble des fonctions f de classe C^k de I dans $\mathbb{R}$ telles que

$$a_k f^{(k)} + \ldots + a_1 f' + a_0 f = 0 \ .$$

— L'ensemble des fonctions bornées de I dans $\mathbb{R}$.
Sous-espaces de l'ensemble des suites.
— L'espace des suites convergentes.
— L'espace des suites bornées.
— L'espace des suites telles que $\sum_{n\geqslant 0} |u_n| < \infty$.
Sous-espaces de l'ensemble des polynômes.
— L'espace $\mathbb{R}_n[X]$ des polynômes de degré inférieur à n.
Sous-espaces de l'ensemble de $\mathbb{R}^n$.
— L'ensemble des vecteurs $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ vérifiant un système d'équations linéaires :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n = 0 \\ \vdots \\ a_{p,1}x_1 + a_{p,2}x_2 + \ldots + a_{p,n}x_n = 0 \end{cases}$$

L'ensemble $L(E, F)$ des applications linéaires de E dans F.

2.1.5 Espace engendré

Definition 4. *Soit E un espace vectoriel et A un sous-ensemble de E. On appelle espace engendré par A l'ensemble des combinaisons linéaires d'éléments de A, c'est à dire l'ensemble des vecteurs $\boldsymbol{v}$ de la forme*

$$\boldsymbol{v} = \sum_{i=0}^{n} a_i \boldsymbol{u}_i \ ,$$

où les a_i sont réels et les $\boldsymbol{u}_i$ des éléments de A. On note cet ensemble Vect(A).

Par exemple, une droite est un espace engendré par un vecteur non nul.

Proposition 5. *L'espace engendré par A est l'intersection des sous-espaces vectoriels de E contenant A.*

$$\mathrm{Vect}(A) = \cap_{F\supset A} F \ .$$

En particulier donc, $\mathrm{Vect}(A)$ *est un espace vectoriel.*

Preuve : **Exercice.** □

2.1.6 Sommes de sous-espaces

Definition 6. *Soit F, G des sous-espaces vectoriels de E. On appelle $F + G$ l'ensemble des vecteurs $\boldsymbol{v} \in E$ de la forme $\boldsymbol{v} = \boldsymbol{u}_F + \boldsymbol{u}_G$, où $\boldsymbol{u}_F \in F$ et $\boldsymbol{u}_G \in G$.*

Proposition 7. *$F + G$ est un sous-espace vectoriel de E.*

Preuve : **Exercice.** □

2.2 Bases et dimension

L'algèbre linéaire s'est développé au début du 20ème siècle pour étudier des problèmes d'analyse fonctionnelle. Ces problèmes font intervenir des espaces de dimension infinie. Plus récemment, des problèmes de statistiques et d'informatiques ont motivé le développement de nouveaux résultats d'algèbre linéaire en dimension finie. Une des clés de cet essor est le concept de base. Muni d'une base, les éléments d'un espace vectoriel de dimension finie sont "encodables" dans des vecteurs qu'on peut manipuler algorithmiquement et sur lesquels on peut faire des calculs. Dans cette section, on rappelle les éléments permettant de définir ces notions ainsi que les premières propriétés fondamentales de ces objets.

2.2.1 Famille génératrice

Soit E un espace vectoriel.

Definition 8. *Une collection G de vecteurs de E tel que* $\mathrm{Vect}(G) = E$ *est appelée famille génératrice de E.*

Definition 9. *On dit que l'espace E est de dimension finie s'il existe une famille génératrice de E de cardinal fini.*

— Toute famille contenant une famille génératrice est génératrice.
— L'espace vectoriel $\mathbb{R}_n[X]$ des polynômes de degré inférieur à n est un espace de dimension finie. (donner une famille génératrice !)

2.2.2 Famille libre

Definition 10. *Une collection G de vecteurs de E est appelée famille libre si tout vecteur de* $\mathrm{Vect}(G)$ *s'écrit de manière unique comme combinaison linéaire d'éléments de G. Si G n'est pas libre on dit qu'elle est liée. Si $G = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$ est une famille libre, on dit aussi que les vecteurs $\boldsymbol{u}_i$ sont linéairement indépendants.*

— Une famille contenant le vecteur nul n'est jamais libre.
— La famille vide est libre.
— Une famille extraite d'une famille libre est libre.

Proposition 11. *Soit $G = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$ une famille finie de vecteurs de E.*

1. *G est libre si et seulement si $a_1\boldsymbol{u}_1 + \ldots + a_p\boldsymbol{u}_p = 0$ implique $a_1 = \ldots = a_n = 0$.*

2. *G est liée si et seulement si il existe $a_1, \ldots, a_p$ non tous nuls, tels que $a_1 \boldsymbol{u}_1 + \ldots + a_p \boldsymbol{u}_p = 0$.*
3. *Si G est liée, un vecteur de G est combinaison linéaire des autres.*

Preuve : Exercice. □

Exemple : Dans $\mathbb{R}^n$, une famille triangulaire sans vecteur nul est libre.

2.2.3 Base

Definition 12. *Une famille B libre et génératrice de E est appelée base de E.*

Théorème 13. *Soit E un espace de dimension finie. On peut extraire de toute famille génératrice G de E une base de E.*

Preuve : Comme E est de dimension finie, il existe une famille finie F engendrant E. Puisque G engendre E, il existe une famille finie H de vecteurs de G engendrant F, donc E. Soit $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ une sous-famille de H engendrant F, de cardinal minimal. Si elle était liée, un des vecteurs, disons $\mathbf{u}_p$ serait combinaison linéaire des autres, donc $(\mathbf{u}_1, \ldots, \mathbf{u}_{p-1})$ serait une famille engendrant F de cardinal inférieur. C'est absurde, donc $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ est libre. Comme elle est génératrice, c'est une base. □

Théorème 14. *Soit E un espace vectoriel de dimension finie. Toute famille libre de E, $L = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$, peut être complétée en une base $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ de E.*

Preuve : Si L engendre E, L est une base. Supposons donc que L n'engendre pas E. Comme E est de dimension finie, il existe une famille finie $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ engendrant E. On pose $L^{(0)} = L$ et récursivement, si pour tout $i \in \{0, \ldots p-1\}$, si $L^{(i)} \cup \mathbf{v}_{i+1}$ est libre $L^{(i+1)} = L^{(i)} \cup \mathbf{v}_{i+1}$, sinon $L^{(i+1)} = L^{(i)}$. Par construction, $L^{(p)}$ est une famille libre engendrant $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ donc E, donc c'est une base de E. □

Definition 15. *Soit E un espace vectoriel de dimension finie et $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E. Tout vecteur $\mathbf{u}$ de E se décompose de manière unique sur la base B : il existe un unique vecteur $\boldsymbol{u}_c = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$ tel que*

$$\boldsymbol{u} = u_1 \boldsymbol{e}_1 + \ldots + u_n \boldsymbol{e}_n \quad .$$

Le vecteur $\mathbf{u}_c$ est appelé vecteur des coordonnées de E dans la base B.

2.2.4 Dimension

Proposition 16. *Soit E un espace vectoriel et $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E. Toute famille de m vecteurs, avec $m > n$ est liée.*

Preuve : Soit $\mathbf{u}_1, \ldots, \mathbf{u}_m$ une famille de m vecteurs. On décompose chacun de ces vecteurs sur la base B :

$$\mathbf{u}_i = \sum_{j=1}^{n} u_{i,j} \mathbf{e}_j \quad .$$

On peut appliquer la méthode du pivot au tableau $u_{i,j}$. Comme $m > n$, il existe au moins une ligne nulle à la fin de l'algorithme, ce qui assure que la famille est liée. □

Théorème 17. *Dans un espace de dimension finie, toutes les bases ont même cardinal.*

Preuve : Exercice. □

Definition 18. *Le cardinal commun n des cardinaux des bases d'un espace de dimension finie E est appelé la dimension de E et est noté* $\dim(E)$*.*

Proposition 19. *Dans un espace de dimension n.*
- *Toute famille libre de cardinal n est une base.*
- *Toute famille génératrice de cardinal n est une base.*

Preuve : Exercice. □

Proposition 20. *Soit E un espace de dimension finie et F un sous-espace de E.*
- *F est de dimension finie.*
- $\dim(F) \leqslant \dim(E)$.
- *Si* $\dim(F) = \dim(E)$, $F = E$.

Preuve : Soit $\mathbf{u}_1, \dots, \mathbf{u}_p$ une famille libre de F. Elle est libre dans E, donc elle peut être complétée en une base $\mathbf{u}_1, \dots, \mathbf{u}_n$ de E. Ainsi, $p \leqslant n$. F ne contient aucune famille libre de plus de n éléments, c'est donc un espace de dimension finie, de dimension inférieure à n.

Si $\dim(F) = n$, F contient une famille libre de n éléments. Ces éléments forment une base de E, donc $F \supset E$. □

Proposition 21. *Soit F et G deux sous-espaces d'un espace E de dimension finie, alors*

$$\dim(F) + \dim(G) = \dim(F + G) + \dim(F \cap G) \ .$$

Preuve : Exercice. *Indication :* Prendre une base de $F \cap G$, la compléter en une base de F et de G et construire avec ces vecteurs une base de $F + G$. □

Definition 22. *Un hyperplan de E est un espace de dimension* $\dim(E) - 1$*.*

Un sous-espace F contenant un hyperplan H vérifie $F = H$ ou $F = E$.
Exemples : Une famille triangulaire sans vecteurs nul de $\mathbb{R}^n$ est une base.
La famille $1, X, \dots, X^n$ de $\mathbb{R}_n[X]$ est une base.

2.3 Applications linéaires

Definition 23. *Soient E et F deux espace vectoriels. Une application $f : E \to F$ est dite linéaire si elle vérifie*

$$\forall \boldsymbol{u}, \boldsymbol{v} \in E, \forall \lambda \in K, \qquad f(\boldsymbol{u} + \lambda \boldsymbol{v}) = f(\boldsymbol{u}) + \lambda f(\boldsymbol{v}) \ .$$

Les applications linéaires sont celles compatibles avec la structure d'espace vectoriel. On déduit de la définition que $f(0) = 0$ et si $\mathbf{u}_1, \ldots, \mathbf{u}_k$ sont des vecteurs de E et $\lambda_1, \ldots, \lambda_k$ sont des scalaires,

$$f(\sum_{i=1}^{k} \lambda_i \mathbf{u}_i) = \sum_{i=1}^{k} \lambda_i f(\mathbf{u}_i) \ .$$

La composée d'applications linéaires est linéaire.

Definition 24. *Soient E et F deux espace vectoriels, soient f et g deux applications linéaires $f, g : E \to F$ et soit $\lambda \in K$ un scalaire.*

La somme des applications f et g est l'application définie par $f+g : E \to F$, $\boldsymbol{u} \mapsto (f+g)(\boldsymbol{u}) = f(\boldsymbol{u}) + g(\boldsymbol{u})$.

La multiplication de f par le scalaire λ est l'application $\lambda f : E \to F$, $\boldsymbol{u} \mapsto (\lambda f)(\boldsymbol{u}) = \lambda f(\boldsymbol{u})$.

Proposition 25. *Muni de ces opérations, l'ensemble $L(E, F)$ de toutes les applications linéaires de E vers F est un espace vectoriel.*

Les éléments de $L(E, E)$ sont appelés les endomorphismes de E. On notera $L(E, E)$ plus simplement $L(E)$ dans la suite.

2.3.1 Exemples

Pour tout $a \in K$, l'application $h_a : E \to E$, définie par $h_a(\mathbf{u}) = a\mathbf{u}$ est une application linéaire. On l'appelle homothétie de rapport a.

Soit E l'espace vectoriel des fonctions de classe C^∞ de $\mathbb{R}$ dans lui-même.

- Pour tout $a \in \mathbb{R}$, l'application $\mathrm{ev}_a : E \to E$, $\mathrm{ev}_a(f) = f(a)$ est linéaire.
- L'application $D : E \to E$ définie par $D(f) = f'$ est linéaire.
- Pour tous réels a et b, l'application $I_{a,b} : E \to \mathbb{R}$ définie par $I_{a,b}(f) = \int_a^b f(t)\mathrm{d}t$ est linéaire.
- Pour tout $a \in \mathbb{R}$, l'application $\tau_a : E \to E$ définie par $\tau_a(f) = f(\cdot - a)$ est linéaire, de même que l'application $\Delta_a(f) = \tau_a(f) - f$.

2.3.2 Propriété universelle

Proposition 26. *Soit E un espace vectoriel de dimension finie n et $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E. Soit F un espace vectoriel et $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ des vecteurs de F. Il existe une unique application linéaire $f : E \to F$ telle que $f(\boldsymbol{e}_i) = \boldsymbol{u}_i$ pour tout $i \in \{1, \ldots, n\}$.*

Preuve : Exercice ! □

2.3.3 Noyau

Definition 27. *Soient E, F deux espaces vectoriels et $f : E \to F$ une application linéaire. Le noyau de f, noté $\ker(f)$ est la préimage de 0 par f, c'est à dire*

$$\ker(f) = \{\boldsymbol{x} \in E : f(\boldsymbol{x}) = 0\} \ .$$

Proposition 28. *Le noyau d'une application linéaire $f : E \to F$ est un sous-espace vectoriel de E.*

Preuve : Exercice ! □

Soit E l'espace vectoriel des fonctions de classe C^∞ de $\mathbb{R}$ dans lui-même. Soit $\theta : E \to E$, $f \mapsto \theta(f) = f'' - 3f' + 2f$. θ est une application linéaire, son noyau est l'ensemble des fonctions solutions de l'équation différentielle

$$f'' - 3f' + 2f = 0 \ .$$

Definition 29. *Une application f entre deux ensembles E et F est injective si, pour tout x et y de E, $x \neq y$, on a $f(x) \neq f(y)$.*

Proposition 30. *Soient E, F deux espaces vectoriels et $f : E \to F$ une application linéaire. f est injective si et seulement* $\ker(f) = \{0\}$.

Preuve : Exercice ! □

Proposition 31. *Soient E, F deux espaces vectoriels et $f : E \to F$ une application linéaire injective. Si $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ est une famille libre de E, alors $f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_n)$ est une famille libre de F.*

Preuve : Exercice ! □

2.3.4 Image d'une application linéaire

Soient E, F deux espaces vectoriels et $f : E \to F$ une application linéaire.

Definition 32. *L'image d'un sous-ensemble G de E par f est l'ensemble défini par*

$$f(G) = \{\boldsymbol{y} \in F : \exists \boldsymbol{x} \in G,\ f(\boldsymbol{x}) = \boldsymbol{y}\} \ .$$

On appelle image de f l'image de l'espace E

$$\mathrm{Im}(f) = f(E) \ .$$

Proposition 33. *Si G est un sous-espace vectoriel de E, $f(G)$ est un sous-espace vectoriel de F.*

Preuve : Exercice ! □

Definition 34. *Si f est un endomorphisme de E et G est un sous-espace de E, alors G est dit stable par f si $f(G) \subset G$.*

Proposition 35. *L'image de l'espace engendré par une famille G de vecteurs de E est l'espace engendré par la famille $f(G)$ dans F, i.e. $f(\mathrm{Vect}(G)) = \mathrm{Vect}(f(G))$.*

En particulier, si G est une famille génératrice de E, $\mathrm{Im}(f)$ est l'espace engendré par $f(G)$.

Preuve : Exercice ! □

2.3.5 Théorème du rang

Théorème 36. *Soit E un espace de dimension finie, F un espace vectoriel quelconque et $f \in L(E,F)$. Les espaces* $\ker(f)$ *et* $\mathrm{Im}(f)$ *sont de dimensions finies et ces dimensions vérifient*

$$\dim(E) = \dim(\ker(f)) + \dim(\mathrm{Im}(f)) \ .$$

Preuve : Comme $\ker(f)$ est un sous-espace vectoriel de E, il est de dimension finie. Soit $\mathbf{e}_1, \ldots, \mathbf{e}_r$ une base de $\ker(f)$ qu'on complète en une base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E. On sait que $f(B)$ est une famille génératrice de $\mathrm{Im}(f)$, donc $(f(\mathbf{e}_{r+1}), \ldots, f(\mathbf{e}_n))$ est également une famille génératrice de $\mathrm{Im}(f)$. Supposons que $a_{r+1}, \ldots, a_n$ soient des scalaires tels que $\sum_{i=r+1}^n a_i f(\mathbf{e}_i) = 0$. On aurait alors

$$\sum_{i=r+1}^{n} a_i \mathbf{e}_i \in \ker(f) \ ,$$

donc par unicité de l'écriture des vecteurs de E sur la base B, $a_{r+1} = \ldots = a_n = 0$. Ainsi, la famille $(f(\mathbf{e}_{r+1}), \ldots, f(\mathbf{e}_n))$ est une famille libre, c'est donc une base de $\mathrm{Im}(f)$, ce qui montre le théorème. □

Proposition 37. *Soit $\boldsymbol{y} \in F$, on s'intéresse à l'équation $f(\boldsymbol{x}) = \boldsymbol{y}$. Supposons $\boldsymbol{y} \in \mathrm{Im}(f)$ et $\boldsymbol{x}_0$ vérifie $f(\boldsymbol{x}_0) = \boldsymbol{y}$. Alors l'ensemble des solutions de l'équation $f(\boldsymbol{x}) = \boldsymbol{y}$ est*

$$\boldsymbol{x}_0 + \ker(f) = \{\boldsymbol{x} \in E : \boldsymbol{x} - \boldsymbol{x}_0 \in \ker(f)\} \ .$$

Preuve : Exercice ! □

2.3.6 Résolution d'un système linéaire

Soit $\mathbf{e}_1, \ldots, \mathbf{e}_p$ la base canonique de $\mathbb{R}^p$ et $\mathbf{e}'_1, \ldots \mathbf{e}'_n$ la base canonique de $\mathbb{R}^n$. Soient $\mathbf{v}_1, \ldots, \mathbf{v}_p$ des vecteurs de $\mathbb{R}^n$. On peut écrire, pour tout $i \in \{1, \ldots p\}$,

$$\mathbf{v}_i = \sum_{j=1}^{n} v_{j,i} \mathbf{e}'_j \ .$$

La propriété universelle assure l'existence et l'unicité de l'application linéaire f telle que $f(\mathbf{e}_i) = \mathbf{v}_i$.

Soit $\mathbf{x} = \sum_{i=1}^p x_i \mathbf{e}_i$ un vecteur de $\mathbb{R}^p$, son image par f est donc

$$f(\mathbf{x}) = \sum_{i=1}^{p} x_i \mathbf{v}_i = \sum_{i=1}^{p} x_i \sum_{j=1}^{n} v_{j,i} \mathbf{e}'_j = \sum_{j=1}^{n} \big(\sum_{i=1}^{p} v_{j,i} x_i \big) \mathbf{e}'_j \ .$$

Le vecteur $\mathbf{x}$ appartient au noyau de f si et seulement si ses coordonnées sont solutions du système linéaire

$$\begin{cases} \sum_{i=1}^p v_{1,i} x_i = 0 \\ \vdots \\ \sum_{i=1}^p v_{n,i} x_i = 0 \end{cases}$$

Soit r le rang de ce système linéaire. On a d'après le théorème du rang $\dim(\ker(f)) = n - r$.

2.3.7 Isomorphismes

Definition 38. *Soient E et F deux espaces vectoriels et soit $f \in L(E,F)$. f est appelée isomorphisme d'espaces vectoriels s'il existe $g \in L(F,E)$ telle que $f \circ g = id_F$ et $g \circ f = id_E$. Si g existe, on l'appelle inverse de f et on la note f^{-1}. Les espaces E et F sont dits isomorphes s'il existe un isomorphisme $f : E \to F$.*

Un isomorphisme entre E et E est appelé automorphisme.

Proposition 39. *La composée de deux isomorphismes $f : E \to F$ et $g : F \to G$ est un isomorphisme et on a*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} \quad .$$

Preuve : Exercice! □

Proposition 40. *Si f est une application linéaire bijective, c'est un isomorphisme.*

Preuve : Exercice! □

Proposition 41. *Soient E et F deux espaces de même dimension n et $f \in L(E,F)$.*
- *Si f est injective, c'est un isomorphisme.*
- *Si f est surjective, c'est un isomorphisme.*

Preuve : Exercice! □

Proposition 42. *Deux espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils sont de même dimension.*

Tout K-espace vectoriel de dimension n est isomorphe à K^n.

Preuve : Exercice! □

Proposition 43. *Soit G un sous-espace de E et $f : E \to F$ un isomorphisme, alors la restriction de f à G est un isomorphisme de G vers $f(G)$.*

Preuve : Exercice! □

2.4 Matrice

Soit E un espace vectoriel de dimension p muni d'une base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_p)$, F un espace vectoriel de dimension n muni d'une base $B' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$. La donnée d'une application $f \in L(E,F)$ est équivalente à la donnée du tableau suivant

$$\mathbf{M} = \begin{bmatrix} a_{1,1} & \ldots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,p} \end{bmatrix} \quad ,$$

où, pour tout $i \in \{1, \ldots, p\}$,

$$f(\mathbf{e}_i) = \sum_{j=1}^{n} a_{j,i} \mathbf{e}'_j \quad .$$

En effet, la donnée de f implique bien sûr la donnée de $\mathbf{M}$ et réciproquement, si on connait les coefficients de $\mathbf{M}$, alors, pour tout $\mathbf{u} = \sum_{i=1}^{p} u_i \mathbf{e}_i \in E$, on a

$$f(\mathbf{u}) = \sum_{i=1}^{p} u_i f(\mathbf{e}_i) = \sum_{j=1}^{n} \left(\sum_{i=1}^{p} a_{j,i} u_i \right) \mathbf{e}'_j \ .$$

Pour bâtir la matrice $\mathbf{M}$, on met donc dans les colonnes les coefficients des vecteurs $f(\mathbf{e}_i)$ décomposés dans la base B'.

2.4.1 Matrices

Definition 44. *Une matrice à coefficients dans K à n lignes et p colonnes est un tableau de np scalaires $(a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$*

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{bmatrix} \ .$$

Le terme général $a_{i,j}$ de cette matrice est celui situé à la i-ème ligne et à la j-ème colonne. Une matrice à n lignes et p colonnes est dite de type (n,p). On note $M_{n,p}$ ou $M_{n,p}(K)$ l'ensemble des matrices de type (n,p) à coefficients dans K. Si $n = p$, on dit que la matrice est carrée de taille n. On note M_n ou $M_n(K)$ l'ensemble des matrices carrées de taille n à coefficients dans K.

2.4.2 Matrice d'une application linéaire

Soit E un espace vectoriel de dimension p muni d'une base $B = (\mathbf{e}_1, \dots, \mathbf{e}_p)$, F un espace vectoriel de dimension n muni d'une base $B' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$. On peut définir l'application Φ qui à tout $f \in L(E,F)$ associe la matrice

$$\mathbf{M} = \begin{bmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{bmatrix} \ ,$$

telle que, pour tout $i \in \{1, \dots, p\}$,

$$f(\mathbf{e}_i) = \sum_{j=1}^{n} a_{j,i} \mathbf{e}'_j \ .$$

Proposition 45. *L'application Φ est une bijection.*

Preuve : Exercice ! □

Definition 46. *Soit $\boldsymbol{M}, \boldsymbol{M'}$ deux matrices de $M_{n,p}$ de coefficients génériques respectifs $a_{i,j}$ et $a'_{i,j}$ et soit $\lambda \in K$*

La somme des matrices $\boldsymbol{M}$ et $\boldsymbol{M'}$ est la matrice $\boldsymbol{M} + \boldsymbol{M'}$ de coefficient générique $a_{i,j} + a'_{i,j}$.

La multiplication de la matrice $\boldsymbol{M}$ par le scalaire λ est la matrice $\lambda \boldsymbol{M}$ de coefficient générique $\lambda a_{i,j}$.

Proposition 47. *Muni de ces deux opérations, $M_{n,p}$ est un espace vectoriel et l'application Φ est un isomorphisme d'espaces vectoriels.*

Preuve : Exercice ! □

2.4.3 Matrices particulières

Matrices diagonales. Une matrice carrée telle que $a_{i,j} = 0$ lorsque $i \neq j$ est appelée matrice diagonale. On note parfois $\mathrm{diag}(a_1, \ldots, a_n)$ la matrice carrée de taille n diagonale dont les coefficients de la diagonale sont $a_1, \ldots, a_n$.

Matrice unité. La matrice unité est la matrice diagonale dont les coefficients diagonaux valent tous 1. On note celle de taille n $\mathbf{I}_n$. C'est la matrice de l'application identité de E dans n'importe quelle base.

Matrice triangulaire. Une matrice dont tous les coefficients $a_{i,j}$ avec $i > j$ sont nuls est appelée matrice triangulaire supérieure.

Si tous les coefficients $a_{i,j}$ avec $i < j$ sont nuls, la matrice est dite triangulaire inférieure.

Le matrices diagonales sont les matrices triangulaires supérieure et inférieure.

Matrices lignes, colonnes. Si $n = 1$, la matrice n'a qu'une ligne, on l'appelle matrice ligne.

Si $p = 1$, la matrice n'a qu'une colonne, on dit que c'est une matrice colonne.

On identifiera toujours un vecteur avec la matrice colonne de ses coefficients.

Base canonique. Pour tout $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, p\}$, on note $E_{i,j}$ la matrice dont tous les coefficients sont nuls, sauf celui à la i-ème ligne et j-ème colonne qui vaut 1. Les matrices $E_{i,j}$ forment une base de $M_{n,p}$.

Matrice de transvection. Une matrice carrée de taille n de la forme $\mathbf{T}_{i,j}(\alpha) = \mathbf{I}_n + \alpha E_{i,j}$ pour un couple (i, j) tel que $i \neq j$ est appelée matrice de transvection.

Matrice de transposition. Une matrice carrée de taille n de la forme $P_{i,j} = \mathbf{I}_n + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$ pour un couple (i, j) tel que $i \neq j$ est appelée matrice de transposition.

2.4.4 Matrice de la composée

On considère trois espaces vectoriels E, F et G de dimension finie :
- E est muni d'une base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$,
- F est muni d'une base $B' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$,
- G est muni d'une base $B'' = (\mathbf{e}''_1, \ldots, \mathbf{e}''_p)$.

On se donne aussi des applications linéaires $f : E \to F$ et $g : F \to G$.

On note $R \in M_{n,m}$ la matrice de f entre les bases B et B', $S \in M_{p,n}$ la matrice de g entre les bases B' et B''.

On sait que $g \circ f$ est aussi une application linéaire de E dans G. Notons $T \in M_{p,m}$ sa matrice entre les bases B et B''.

Soit $r_{i,j}$ le coefficient générique de R, $s_{i,j}$ celui de S et $t_{i,j}$ celui de T.

On a, pour tout $i \in \{1, \ldots, m\}$,

$$\sum_{i=1}^{p} t_{i,k}\mathbf{e}''_i = g\circ f(\mathbf{e}_k) = g\left(\sum_{j=1}^{n} r_{j,k}\mathbf{e}'_j\right) = \sum_{j=1}^{n} r_{j,k} \sum_{i=1}^{p} s_{i,j}\mathbf{e}''_i = \sum_{i=1}^{p} \left(\sum_{j=1}^{n} s_{i,j} r_{j,k}\right)\mathbf{e}''_i \ .$$

Par unicité de la décomposition d'un vecteur dans une base, on a donc, pour tout $i \in \{1, \ldots, p\}$ et tout $k \in \{1, \ldots, m\}$,

$$t_{i,k} = \sum_{j=1}^{n} s_{i,j} r_{j,k} \quad .$$

Isolons la i-ème ligne de S et la k-ème colonne de R,

$$\begin{bmatrix} s_{i,1} & \ldots & s_{i,n} \end{bmatrix} \begin{bmatrix} r_{1,k} \\ \vdots \\ r_{n,k} \end{bmatrix} \quad ,$$

$t_{i,k}$ s'obtient en multipliant les termes de même rang des ces vecteurs et en prenant la somme de ces produits.

Pour calculer la matrice T, il faut faire mp fois cette opération.

Produit de matrices

Definition 48. *Soit $S \in M_{p,n}$ et $R \in M_{n,m}$ deux matrices telles que le nombre de colonnes de S soit égal au nombre de ligne de R. Soient $s_{i,j}$ le coefficient générique de S et $r_{i,j}$ celui de R. Le produit des matrices S et R, noté SR est la matrice $T \in M_{p,m}$ de coefficient générique*

$$t_{i,k} = \sum_{j=1}^{n} s_{i,j} r_{j,k} \quad .$$

Par construction donc SR est la matrice de la composée $g \circ f$ des applications $g : F \to G$ et $f : E \to F$ de matrices respectives S et R.

2.4.5 Propriétés du produit

Proposition 49. *Pour toute matrice $\boldsymbol{M} \in M_{n,p}$, $\boldsymbol{M}\boldsymbol{I}_p = \boldsymbol{M}$ et $\boldsymbol{I}_n\boldsymbol{M} = \boldsymbol{M}$.*

Preuve : Exercice ! □

Proposition 50. *Soient $\boldsymbol{A} \in M_{n,p}$, $\boldsymbol{B} \in M_{p,q}$, $\boldsymbol{C} \in M_{q,r}$. On a*

$$\boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}) = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} \quad .$$

Preuve : Exercice ! □

Proposition 51. *Pour toutes matrices pour lesquelles ces opérations sont licites, on a*

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) &= \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C} \quad , \\ (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} &= \boldsymbol{A}\boldsymbol{C} + \boldsymbol{B}\boldsymbol{C} \quad , \\ \boldsymbol{A}(\lambda\boldsymbol{B}) = (\lambda\boldsymbol{A})\boldsymbol{B} &= \lambda(\boldsymbol{A}\boldsymbol{B}) \quad . \end{aligned}$$

Preuve : Exercice ! □

Definition 52. *Une matrice carrée $\boldsymbol{A}$ de taille n est inversible s'il existe une matrice $\boldsymbol{B}$ telle que $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}_n$. Lorsqu'elle existe, une telle matrice $\boldsymbol{B}$ est unique, on l'appelle l'inverse de $\boldsymbol{A}$ et on la note $\boldsymbol{A}^{-1}$.*

Vérifier que $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. Si $\mathbf{A}$ n'est pas carrée elle ne peut pas avoir d'inverse. Si $\mathbf{A}$ et $\mathbf{B}$ sont des matrices carrées inversibles, vérifier que $\mathbf{AB}$ est inversible et que

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad .$$

Si f est un automorphisme de E, sa matrice $\mathbf{A}$ dans une base B est inversible et l'inverse de f, f^{-1} a pour matrice $\mathbf{A}^{-1}$ dans la base B.

2.4.6 Changement de base

Soit E, F deux espaces vectoriels de dimension finie, soient B_E et B'_E deux bases de E, B_F et B'_F deux bases de F. Soit $f \in L(E,F)$. Soit $\mathbf{A}$ la matrice de f de E muni de B_E dans F muni de B_F, soit $\mathbf{A}'$ la matrice de f de E muni de B'_E dans F muni de B'_F. Soit $\mathbf{P}_E$ la matrice de l'identité de (E, B'_E) dans (E, B_E) (cette matrice est appelée matrice de passage de B_E à B'_E) et $\mathbf{P}_F$ la matrice de l'identité de (F, B'_F) dans (F, B_F).

Le tableau suivant résume la situation

$$\begin{array}{ccc} & \mathbf{A} & \\ (E, B_E) & \rightarrow & (F, B_F) \\ & f & \\ \mathbf{P}_E \uparrow \mathrm{id}_E & & \mathrm{id}_F \uparrow \mathbf{P}_F \\ & f & \\ (E, B'_E) & \rightarrow & (F, B'_F) \\ & \mathbf{A}' & \end{array} \quad .$$

Clairement, on a $f = \mathrm{id}_F \circ f \circ \mathrm{id}_E$, ce qui se réécrit matriciellement.

$$\mathbf{A} = \mathbf{P}_F\mathbf{A}'\mathbf{P}_E^{-1} \quad .$$

Cette relation est connue sous le nom de *formule de changement de base.*

Important : Pour ne pas se tromper dans cette relation, mieux vaut éviter de l'apprendre et chercher plutôt à la retrouver avec le petit schéma précédent !

Definition 53. *Deux matrices carrées $\boldsymbol{A}$ et $\boldsymbol{A}'$ sont dites semblables s'il existe une matrice inversible $\boldsymbol{P}$ telle que $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{A}'\boldsymbol{P}^{-1}$.*

Definition 54. *Le rang d'une application $f \in L(E,F)$ est la dimension de $Im(f)$. Le rang d'une matrice $\boldsymbol{A}$ est le rang de l'application linéaire f qu'elle représente. C'est la dimension de l'espace engendré par les colonnes de $\boldsymbol{A}$.*

Definition 55. *La trace de la matrice $\boldsymbol{A}$ carrée de taille n est la somme de ces coefficients diagonaux : si $a_{i,j}$ est le coefficient générique de $\boldsymbol{A}$:*

$$Tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{i,i} \quad .$$

Proposition 56. *Pour toutes matrices carrées $\boldsymbol{A}$ et $\boldsymbol{B}$ de taille n, on a*

$$Tr(\boldsymbol{AB}) = Tr(\boldsymbol{BA}) \quad .$$

En particulier, la trace d'une matrice est égale à celle de toute matrice qui lui est semblable.

Preuve : Exercice ! □

Definition 57. *Soit $\boldsymbol{A} \in M_{n,p}$ une matrice de coefficient générique $a_{i,j}$. La transposée de $\boldsymbol{A}$ est la matrice notée $\boldsymbol{A}^T$ de $M_{p,n}$ de coefficient générique $b_{i,j}$ défini par*

$$b_{i,j} = a_{j,i} \ .$$

2.5 Sommes directes

2.5.1 Décomposition en somme directe

Definition 58. *Soit E un espace vectoriel et F et G deux sous-espaces de E. On dit que E est somme directe de F et G et on note $E = F \oplus G$ si tout vecteur $\boldsymbol{u}$ de E se décompose de manière unique sous la forme $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$, avec $\boldsymbol{v} \in V$ et $\boldsymbol{w} \in W$. On dit alors que F et G sont supplémentaires dans E ou que F est un supplémentaire de G dans E.*

Proposition 59. *On a $E = F \oplus G$ si et seulement si $E = F + G$ et $F \cap G = \{0\}$.*

Preuve : Supposons que $E = F \oplus G$. Alors clairement $E = F + G$ et, si $\mathbf{u} \in F \cap G$, on a $\mathbf{u} = \mathbf{u} + 0$ avec $\mathbf{u} \in F$ et $0 \in G$ et $\mathbf{u} = 0 + \mathbf{u}$ avec $0 \in F$ et $\mathbf{u} \in G$, donc, par unicité, $\mathbf{u} = 0$.

Réciproquement, supposons que $E = F + G$ et $F \cap G = \{0\}$. Soit $\mathbf{u} \in E$, comme $E = F + G$, il existe $\mathbf{v} \in F$ et $\mathbf{w} \in G$ tels que $\mathbf{u} = \mathbf{v} + \mathbf{w}$. Supposons maintenant qu'il existe également $\mathbf{v}' \in F$ et $\mathbf{w}' \in G$ tels que $\mathbf{u} = \mathbf{v}' + \mathbf{w}'$. Alors $\mathbf{v} - \mathbf{v}' = \mathbf{w}' - \mathbf{w}$. Comme $\mathbf{v} - \mathbf{v}' \in F$ et $\mathbf{w}' - \mathbf{w} \in G$, ces vecteurs sont dans $F \cap G$, ils sont donc nuls. □

Proposition 60. *Soit E un espace vectoriel de dimension finie et soient F et G deux sous-espaces vectoriels de E tels que $E = F \oplus G$.*

1. *Si B est une base de F et B' une base de G, alors $B \cup B'$ est une base de E.*
2. $\dim(E) = \dim(F) + \dim(G)$.

Preuve : Exercice ! □

Proposition 61. *Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F admet un supplémentaire dans E.*

Preuve : Exercice ! □

Si $\dim(E) = n$ et $\dim(F) = k$, tout supplémentaire de F est de dimension $n - k$. Les espaces E et $\{0\}$ sont supplémentaires dans E. En général, un supplémentaire de sous-espace n'est pas unique : dans $\mathbb{R}^3$, si $F = \mathrm{Vect}(\mathbf{e}_1, \mathbf{e}_2)$, tout vecteur de la forme $\mathbf{u} = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3$ tel que $c \neq 0$ vérifie $G = \mathrm{Vect}(\mathbf{u})$ est supplémentaire de F dans E.

2.5.2 Sommes directes finies

On peut généraliser les résultats de la section précédente au cas de n sous-espaces $F_i, 1 \leqslant i \leqslant n$.

Definition 62. *On dit que E est somme directe des sous-espaces $F_i, 1 \leqslant i \leqslant n$ si tout vecteur $\boldsymbol{u}$ de E s'écrit de manière unique $\boldsymbol{u} = \boldsymbol{v}_1 + \ldots + \boldsymbol{v}_n$, avec $\boldsymbol{v}_i \in F_i$ pour tout $1 \leqslant i \leqslant n$. On écrit alors $E = \oplus_{i=1}^n F_i$ ou $E = F_1 \oplus \ldots \oplus F_n$. On note aussi $E_{(-i)} = \oplus_{j \in \{1,\ldots,n\} \setminus \{i\}} F_j$.*

Proposition 63. *On a $E = \oplus_{i=1}^n F_i$ si et seulement si $E = F_1 + \ldots + F_n$ et, pour tout $i \in \{1, \ldots, n\}$, $F_i \cap E_{(-i)} = \{0\}$.*

Preuve : Supposons $E = \oplus_{i=1}^n F_i$, alors, $E = F_1 + \ldots + F_n$. Soit maintenant $i \in \{1, \ldots, n\}$ et $\mathbf{u} \in F_i \cap E_{(-i)}$. Alors s'écrit de manière unique $\mathbf{u} = \sum_{j=1}^n \mathbf{v}_j$, avec $\mathbf{v}_j \in F_j$, donc, comme $\mathbf{u} \in F_i$, tous les $\mathbf{v}_j$, avec $j \neq i$, vérifient $\mathbf{v}_j = 0$, et comme $\mathbf{u} \in E_{(-i)}$, on a $\mathbf{v}_i = 0$, donc $\mathbf{u} = 0$.

Supposons $E = F_1 + \ldots + F_n$ et, pour tout $i \in \{1, \ldots, n\}$, $F_i \cap E_{(-i)} = \{0\}$. Tout vecteur $\mathbf{u}$ de E s'écrit $\mathbf{u} = \sum_{i=1}^n \mathbf{v}_i$, avec $\mathbf{v}_i \in F_i$ pour tout $i \in \{1, \ldots, n\}$. Supposons que $\mathbf{u} = \sum_{i=1}^n \mathbf{v}'_i$ avec $\mathbf{v}'_i \in F_i$, pour tout $i \in \{1, \ldots, n\}$. Alors, pour tout $i \in \{1, \ldots, n\}$, on a

$$\mathbf{v}_i - \mathbf{v}'_i = \sum_{j \in \{1,\ldots,n\}} \mathbf{v}'_j - \mathbf{v}_j \ .$$

Ainsi $\mathbf{v}_i - \mathbf{v}'_i \in F_i \cap E_{(-i)}$, donc $\mathbf{v}_i = \mathbf{v}'_i$. Ceci étant vrai pour tout i, on a $\mathbf{u} = 0$. □

Proposition 64. *Soit E un espace vectoriel et F_i, $1 \leqslant i \leqslant n$ des sous-espaces vectoriels tels que $E = \oplus_{i=1}^n F_i$.*

1. *Si pour tout $i \in \{1, \ldots, n\}$, B_i est une base de F_i, alors $B_1 \cup \ldots \cup B_n$ est une base de E.*
2. $\dim(E) = \sum_{i=1}^n \dim(F_i)$.

Preuve : Exercice ! □

2.5.3 Produit de deux espaces vectoriels

Definition 65. *Soient E_1 et E_2 deux espaces vectoriels sur le même corps K. L'espace produit $E_1 \times E_2$ des couples $(\boldsymbol{u}_1, \boldsymbol{u}_2)$, où $\boldsymbol{u}_i \in E_i$ pour $i \in \{1, 2\}$ est un espace vectoriel sur K si on le munit des opérations + et . suivantes :*

$$(\boldsymbol{u}_1, \boldsymbol{u}_2) + (\boldsymbol{v}_1, \boldsymbol{v}_2) = (\boldsymbol{u}_1 + \boldsymbol{v}_1, \boldsymbol{u}_2 + \boldsymbol{v}_2) \ ,$$
$$\lambda(\boldsymbol{u}_1, \boldsymbol{u}_2) = (\lambda \boldsymbol{u}_1, \lambda \boldsymbol{u}_2) \ .$$

Les applications $p_1 : E_1 \times E_2 \to E_1$ et $p_2 : E_1 \times E_2 \to E_2$, définies par $p_1(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1$, $p_2(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_2$ sont clairement linéaires et surjectives, leurs noyaux sont $\ker(p_1) = \{(0, \mathbf{u}_2), \mathbf{u}_2 \in E_2\}$ qui est isomorphe à E_2 et $\ker(p_2) = \{(\mathbf{u}_1, 0), \mathbf{u}_1 \in E_1\}$ qui est isomorphe à E_1.

Les applications $j_1 : E_1 \to E_1 \times E_2$ et $j_2 : E_2 \to E_1 \times E_2$, définies par $j_1(\mathbf{u}_1) = (\mathbf{u}_1, 0)$, $j_2(\mathbf{u}_2) = (0, \mathbf{u}_2)$ sont clairement linéaires et injectives, leurs images sont $\mathrm{Im}(p_1) = \ker(p_2)$et $\mathrm{Im}(p_2) = \ker(p_1)$. Enfin, on a $p_1 \circ j_1 = \mathrm{id}_{E_1}$, $p_2 \circ j_2 = \mathrm{id}_{E_2}$.

Proposition 66. *On a* $E_1 \times E_2 = j_1(E_1) \oplus j_2(E_2)$ *et* $\dim(E_1 \times E_2) = \dim(E_1) + \dim(E_2)$.

Preuve : Exercice ! □

2.5.4 Projecteurs

Definition 67. *On appelle projecteur de* E *toute application linéaire* p *telle que* $p \circ p = p$.

Proposition 68. *Soit* E *un espace vectoriel et* p *un projecteur de* E*, on a* $E = \ker(p) \oplus \mathrm{Im}(p)$.

Preuve : Exercice ! □

Si p est un projecteur $q = \mathrm{id} - p$ aussi, et on a $\ker(p) = \mathrm{Im}(q)$, $\mathrm{Im}(p) = \ker(q)$.

2.6 Permutations

2.6.1 Groupes

Definition 69. *Un groupe* $(G, .)$ *où le point désigne une application* $G \times G \to G$ *est un groupe si la loi . vérifie les axiomes suivants :*
- *associativité :* $(xy)z = x(yz)$*, pout tout* x, y, z *de* G.
- *il existe un élément neutre, noté* 1 *vérifiant* $1x = x1 = x$*, pour tout* $x \in G$.
- *pour tout* $x \in G$*, il existe un élément noté* x^{-1} *et appelé* inverse *de* x *tel que* $xx^{-1} = x^{-1}x = 1$.

Exemples : L'ensemble $\{-1, 1\}$ muni du produit est un groupe, dont l'élément neutre est 1. L'ensemble $\mathbb{Z}$ muni de la loi $+$ est un groupe d'élément neutre 0.

Definition 70. *Soient* $(G, .)$ *et* $(G', .)$ *deux groupes. Une application* $f : G \to G'$ *est un morphisme de groupe si pour tout* x, y *de* G*,*

$$f(xy) = f(x)f(y) \ .$$

Exemples : Les fonctions $z \mapsto az$ sont des morphismes de $(\mathbb{Z}, +)$ dans lui-même, si $a \in \mathbb{Z}$. La fonction log est un morphisme de $(\mathbb{R}_+^*, .)$ dans $(\mathbb{R}, +)$.

2.6.2 Groupe des permutations

Definition 71. *Soit* E *un ensemble et* S_E *l'ensemble des bijections de* E *dans lui même.* S_E *muni de la loi de composition est un groupe (vérifiez le !).*

Si $E = \{1, \ldots, n\}$*,* S_E *est appelé groupe symétrique et les éléments de* S_E *sont appelés permutations de* $\{1, \ldots, n\}$*. On note* S_E *par* S_n *dans ce cas.*

Notation : Soit $\sigma \in S_n$, on note souvent

$$\sigma = \begin{pmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{pmatrix} \ .$$

Definition 72. *On appelle transposition tout élément de S_n qui échange deux éléments de $\{1, \ldots, n\}$ et laisse les autres fixes. On note $\tau = (i, j)$ la transposition qui échange i et j. On a $\tau\tau = id$.*

Definition 73. *On appelle cycle de longueur $r > 1$ toute permutation $\sigma \in S_n$ pour laquelle il existe des éléments $x_1, \ldots, x_r$ de E tels que $\sigma(x) = x$ si $x \notin \{x_1, \ldots, x_r\}$ et $\sigma(x_1) = x_2$, …, $\sigma(x_{r-1}) = x_r$, $\sigma(x_r) = x_1$.*

Deux cycles $(x_1, \ldots, x_r)$ et $(y_1, \ldots, y_s)$ sont disjoints si $\{x_1, \ldots, x_r\} \cap \{y_1, \ldots, y_s\} = \emptyset$. Deux cycles disjoints commutent.

Proposition 74. *Toute permutation se décompose en produit de cycles disjoints (de manière unique, à l'ordre des cycles près).*

Preuve : Prendre un exemple. □

Théorème 75. *Le groupe S_n est engendré par l'ensemble des transpositions.*

Preuve : Il suffit de le vérifier pour les cycles, et la formule

$$(x_1, \ldots, x_r)(x_1, x_r) = (x_2, \ldots, x_r) \ ,$$

permet de montrer le résultat par récurrence. □

2.6.3 Signature d'une permutation

Théorème 76. *Soit $\sigma \in S_n$. Si σ s'écrit de deux manières comme produit de transpositions : $\sigma = t_1 \circ \ldots \circ t_r = t'_1 \circ \ldots \circ t'_{r'}$, alors r et r' ont même parité.*

Pour montrer ce théorème, on s'appuie sur la notion suivante.

Definition 77. *Soit $O = \{(i, j) \in \{1, \ldots, n\}^2 : 1 \leqslant i < j \leqslant n\}$. La signature de $\sigma \in S_n$ est définie par*

$$\varepsilon(\sigma) = \prod_{(i,j) \in O} \frac{\sigma(j) - \sigma(i)}{j - i} \ .$$

On montre alors le théorème suivant.

Théorème 78. *L'application $\varepsilon : S_n \to \{-1, 1\}$, $\sigma \mapsto \varepsilon(\sigma)$ est un isomorphisme de groupe (i.e. un morphisme bijectif dont l'inverse est un morphisme).*

On a alors, sous les hypothèses du théorème 76, $\varepsilon(\sigma) = (-1)^r = (-1)^{r'}$, donc r et r' ont même parité comme annoncé.

2.7 Exercices

Exercice 1

1. Soient a_1, a_2 deux réels distincts, montrer que la famille $(1, a_1)$, $(1, a_2)$ est une famille libre de $\mathbb{R}^2$.
2. Soient a_1, a_2, a_3 trois réels distincts, montrer que la famille $(1, a_1, a_1^2)$, $(1, a_2, a_2^2)$, $(1, a_3, a_3^2)$ est une famille libre de $\mathbb{R}^3$.
3. Soient a_1, a_2, a_3, a_4 4 réels distincts, montrer que la famille $(1, a_1, a_1^2, a_1^3)$, $(1, a_2, a_2^2, a_2^3)$, $(1, a_3, a_3^2, a_3^3)$, $(1, a_4, a_4^2, a_4^3)$ est une famille libre de $\mathbb{R}^4$.

Exercice 2

Les familles suivantes de fonctions sont elles libres ou liées ?

1. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille f_1, f_2, f_3 avec $f_1 : x \mapsto 1$, $f_2 : x \mapsto \cos^2(x)$, $f_3 : x \mapsto sin^2(x)$.
2. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille f_1, f_2, f_3 avec $f_1 : x \mapsto e^x$, $f_2 : x \mapsto e^{2x}$, $f_3 : x \mapsto e^{3x}$.
3. Dans $\mathrm{F}((0,+\infty),\mathbb{R})$, la famille f_1, f_2, f_3 avec $f_1 : x \mapsto \ln(x)$, $f_2 : x \mapsto \ln^2(x)$, $f_3 : x \mapsto \ln^3(x)$.
4. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille f_1, f_2, f_3 avec $f_1 : x \mapsto e^x$, $f_2 : x \mapsto e^{-x}$, $f_3 : x \mapsto \cosh(x)$.
5. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille $(f_n, n \in \mathbb{N})$ avec $f_n : x \mapsto e^{nx}$.
6. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille $(f_n, n \in \mathbb{N})$ avec $f_n : x \mapsto |x-n|$.
7. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille $(f_k, k \in \{1,\ldots,n\})$ avec $f_k : x \mapsto \sin(kx)$.
8. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille $(f_k, k \in \{1,\ldots,n\})$ avec $f_k : x \mapsto \cos(kx)$.
9. Dans $\mathrm{F}(\mathbb{R},\mathbb{R})$, la famille $(f_k, g_k, k \in \{1,\ldots,n\})$ avec $f_k : x \mapsto \sin(kx)$, $g_k : x \mapsto \cos(kx)$.
10. Dans $\mathrm{F}((-2,2),\mathbb{R})$, la famille f_1, f_2, f_3 avec $f_1 : x \mapsto 1/(x-2)$, $f_2 : x \mapsto 1/(x+2)$, $f_3 : x \mapsto (4x+7)/(x^2-4)$.

Exercice 3

Montrer que les familles suivantes sont des bases des espaces indiqués.

1. Dans $\mathbb{R}_4[X]$, la famille $(1, X-1, (X-1)^2, (X-1)^3, (X-1)^4)$. Donner la décomposition d'un polynôme P de $\mathbb{R}_4[X]$ dans cette base en fonction de la valeur de P et de ses dérivées en 1.
2. Dans $\mathbb{R}_2[X]$, la famille $(X, X(X+1), (X+1)^2)$. Donner la décomposition de 1, X, X^2 dans cette base, puis la décomposition d'un polynôme $P(X) = a + bX + cX^2$ de $\mathbb{R}_2[X]$ dans cette base.
3. Dans $\mathbb{R}_4[X]$, la famille
$$\left(1, X, \frac{X(X-1)}{2}, \frac{X(X-1)(X-2)}{6}, \frac{X(X-1)(X-2)(X-3)}{24}\right) .$$
Donner l'écriture des polynômes $1+X+X^2+X^3$ et $1+X+X^2+X^4$ dans cette base.
4. Généraliser la base de la question précédente pour obtenir une base de $\mathbb{R}_n[X]$. Montrer que les polynômes de cette base vérifient $P(z) \in \mathbb{Z}$, pour tout $z \in \mathbb{Z}$. Montrer que tout polynôme vérifiant cette propriété est combinaison linéaire à coefficients entiers des éléments de la base.
5. Soit $n \geqslant 0$ un entier et pour tout $k \in \{0,\ldots,n\}$, soit
$$B_k(X) = \binom{n}{k} X^k (1-X)^{n-k} .$$
Pour toute fonction $f : [0,1] \to \mathbb{R}$, on définit
$$P_f = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_k .$$

P_f est un polynôme de $\mathbb{R}_n[X]$ appelé polynôme de Bernstein associé à f.

(a) Montrer que $(B_k, k \in \{0, \dots, n\})$ est une base de $\mathbb{R}_n[X]$.

(b) Exprimer les polynômes $1, X, X^2$ dans cette base. (On pourra calculer P_f pour chacune de ces fonctions).

Exercice 4

Soient $a_1, \dots, a_{n+1}$ des entiers tous distincts, et, pour tout $i \in \{1, \dots, n+1\}$, soient

$$p_i(X) = \prod_{j \in \{1,\dots,n+1\} \setminus \{i\}} (X - a_j), \quad A = p_i(a_i), \quad P_i(X) = \frac{1}{A_i} p_i(X) \ .$$

1. Montrer que $P_1, \dots, P_{n+1}$ est une base de $\mathbb{R}_n[X]$.
2. Soient $b_1, \dots, b_{n+1}$ des réels quelconques. Déterminer un polynôme de $\mathbb{R}_n[X]$ tel que $P(a_i) = b_i$ pour tout $i \in \{1, \dots, n+1\}$.
3. Déterminer un polynôme de $\mathbb{R}_2[X]$ tel que $P(0) = 3$, $P(1) = -2$, $P(2) = 5$. En existe-t-il d'autres ?

Exercice 5

Soit $\Delta : \mathbb{R}[X] \to \mathbb{R}[X]$, $P \mapsto \Delta(P)$ avec $\Delta(P)(X) = P(X+1) - P(X)$.

1. Décrire le noyau et l'image de la restriction de Δ à $\mathbb{R}_4[X]$. Donner la dimension de ces espaces.
2. Décrire le noyau et l'image de la restriction de Δ à $\mathbb{R}_n[X]$. Donner la dimension de ces espaces. Montrer que Δ est surjective.
3. Décrire l'ensemble des solutions de l'équation $\Delta(P) = Q$.

Exercice 6

Soit $f : \mathbb{R}_4[X] \to \mathbb{R}_4[X]$ défini par $f(P) = (X-1)P' - P$.

1. Déterminer $\ker(f)$.
2. Déterminer l'ensemble des vecteurs Q pour lesquels l'équation $f(P) = Q$ a une solution. (on pourra calculer $f(X-1)^k)$, pour tout $k \in \{1, \dots, 4\}$).
3. Résoudre $(X-1)P' - P = X^2 - 2X + 2$.

Exercice 7

Soit $f \in L(E, F)$ une application injective. Montrer qu'il existe une application linéaire $g : F \to E$ telle que $g \circ f = \mathrm{id}_E$.

Exercice 8

Soit $f \in L(E, F)$ une application surjective. Montrer qu'il existe une application linéaire $g : F \to E$ telle que $f \circ g = \mathrm{id}_F$.

Exercice 9

Soient f et g deux endomorphismes de E tels que $f \circ g = g \circ f$. Montrer que $\ker f$ et $\mathrm{Im}(f)$ sont stables par g.

Exercice 10

Soit $E = \mathrm{Fonc}(\mathbb{R}, \mathbb{R})$ et $\Delta : E \to E$, $f \mapsto \Delta(f) = f(\cdot + 1) - f$.

1. Déterminer $\ker(\Delta)$.
2. Déterminer $\ker(\Delta^2)$.
3. Déterminer $\ker(\Delta^n)$, pour tout $n \geqslant 1$.

Exercice 11

Soit f un endomorphisme de E.

1. Montrer que $\ker(f) \subset \ker(f^2)$ et, plus généralement, que $\ker(f^n) \subset \ker(f^{n+1})$.
2. Montrer que, si $\ker(f^s) = \ker(f^{s+1})$, alors $\ker(f^s) \subset \ker(f^{s+n})$ pour tout $n \geqslant 1$.
3. Montrer que, si E est de dimension finie, il existe s tel que $\ker(f^s) = \ker(f^{s+1})$.

Exercice 12

Soit E l'ensemble des fonctions de classe C^∞ sur $\mathbb{R}$ et soit $D : E \to E$ tel que $D(f) = f'$. Supposons qu'il existe un endomorphisme L de E et un entier $n \geqslant 2$ tels que $D = L^n$.

1. Déterminer $\ker(D)$.
2. Montrer que $\ker L \neq \{0\}$.
3. Montrer que $\ker L^k = \ker D$, pour tout $k \geqslant 1$.
4. Obtenir une contradiction en considérant $\ker(D^2)$.

Exercice 13

Calculer, pour tout $n \geqslant 1$, $\mathbf{A}^n$, où

$$\mathbf{A} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} .$$

Exercice 14

Une matrice $\mathbf{A}$ est dite nilpotente s'il existe r tel que $\mathbf{A}^r = 0$.

1. Si $\mathbf{A}$ est nilpotente, est-elle inversible ?
2. Montrer qu'un matrice triangulaire supérieure est nilpotente si et seulement si ses coefficients diagonaux sont nuls.
3. Montrer que, si $\mathbf{A}$ est nilpotente $\mathbf{I} + \mathbf{A}$ est inversible et exprimer son inverse en fonction des puissances de $\mathbf{A}$.

Exercice 15

1. Déterminer le nombre d'opérations de base (additions et multiplications) nécessaire pour effectuer le calcul du produit de deux matrices carrées de taille 2.
2. On pose $\mathbf{A}$, $\mathbf{B}$ carrée de taille 2 de coefficient générique respectifs $a_{i,j}$, $b_{i,j}$ et on note $c_{i,j}$ le coefficient générique de $\mathbf{AB}$.

$$\begin{array}{lll} E = (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2}), & F = (a_{2,1} + a_{2,2})b_{1,1} & \\ G = a_{1,1}(b_{1,2} - b_{2,2}), & H = a_{2,2}(b_{2,1} - b_{1,1}), & J = (a_{1,1} + a_{1,2})b_{2,2} \ , \\ K = (a_{2,1} - a_{1,1})(b_{1,1} + b_{1,2}), & L = (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2}) \ . & \end{array}$$

   Vérifier que

$$c_{1,1} = E+H-J+L, \quad c_{1,2} = G+J, \quad c_{2,1} = F+H, \quad c_{2,2} = E+G-F+K \ .$$

3. Supposons maintenant que $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices de tailles 2^k. On peut décomposer ces matrices par blocs en 4 matrices de tailles 2^{k-1}. Vérifier qu'on peut appliquer les formules précédentes pour déterminer le produit de $\mathbf{A}$ par $\mathbf{B}$ à partir des produits $\mathbf{A}_{i,j}\mathbf{B}_{k,l}$.
4. En appliquant récursivement la méthode de la question précédente, déterminer le nombre d'opérations d'opérations de base nécessaires pour calculer le produit $\mathbf{AB}$ avec cette méthode. Comparer au nombre d'opérations de base nécessaires par la méthode directe.
5. Comment généraliser cette méthode à des matrices de taille quelconque ?

Exercice 16

Soient m et n deux entiers. Pour quelle valeur de p $\mathbb{R}^p$ est-il isomorphe à

1. $\mathbb{R}^n \times \mathbb{R}^m$?
2. $\mathbb{R}^n \oplus \mathbb{R}^m$?

Exercice 17

Soit E un espace vectoriel de dimension n et $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ une base de E. Montrer que $E = \oplus_{i=1}^n \operatorname{Vect}(\mathbf{e}_i)$.

Exercice 18

Soit E un espace vectoriel B une base de E et B_1, B_2 une partition de B. Montrer que $E = \operatorname{Vect}(B_1) \oplus \operatorname{Vect}(B_2)$.

Exercice 19

Soit E un espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E tels que $F \cap G = \{0\}$, $\dim(F) + \dim(G) = \dim(E)$. Montrer que $E = F \oplus G$.

Exercice 20

Si $E = F_1 \oplus F_2$ et G est un sous-espace vectoriel de E, a-t-on $G = G \cap F_1 \oplus G \cap F_2$?

Exercice 21

Montrer que $E = F_1 \oplus F_2 \oplus F_3$ si et seulement si $E = F_1 + F_2 + F_3$ et $F_1 \cap F_2 = \{0\}$, $F_3 \cap (F_1 + F_2) = \{0\}$.

Exercice 22

Soit E un espace de dimension finie n, F et G deux sous-espaces de E de même dimension k. Montrer qu'il existe un sous-espace de E qui est supplémentaire de F et G.

Exercice 23

Soit S l'ensemble des matrices $\mathbf{M}$ de $M_n(\mathbb{R})$ telle que $\mathbf{M}^T = \mathbf{M}$ et A l'ensemble des matrices $\mathbf{M}$ de $M_n(\mathbb{R})$ telle que $\mathbf{M}^T = -\mathbf{M}$. Montrer que S et A sont des sous-espaces supplémentaires de $M_n(\mathbb{R})$.

Exercice 24

Soient p et q deux projecteurs de E. A quelles conditions sur les noyaux et images de p et q équivalent les relations $p \circ q = p$ et $p \circ q = q$?

Exercice 25

1. Donner un exemple de projecteurs p et q tels que $p \circ q = q \circ p$.
2. Donner un exemple de projecteurs p et q tels que $p \circ q \neq q \circ p$.
3. Soient $E_1 = \ker(p) \cap \ker(q)$, $E_2 = \mathrm{Im}(p) \cap \ker(q)$, $E_3 = \ker(p) \cap \mathrm{Im}(q)$, $E_4 = \mathrm{Im}(p) \cap \mathrm{Im}(q)$. Montrer que, $E = \oplus_{i=1}^{4} E_i$ si et seulement si $p \circ q = q \circ p$.
4. Montrer qu'alors $p \circ q$ et $p + q - p \circ q$ sont des projecteurs. Déterminer leurs images et leurs noyaux.

Exercice 26

Le jeu de taquin se joue sur un carré de 16 cases dont une est vide. La seule façon de changer la disposition des cases est de faire glisser une case dans la case vide. Si la position de départ est la suivante

13	14	15	vide
9	10	11	12
5	6	7	8
1	2	3	4

,

peut-on trouver une suite de mouvement amenant à la position suivante ?

13	15	14	vide
9	10	11	12
5	6	7	8
1	2	3	4

Chapitre 3

Déterminants

3.1 Un peu d'histoire

En 1750, Cramer donne le premier une règle permettant de déterminer les solutions d'un système de n équations à n inconnues. Il raisonne sur les cas où $n = 2$, $n = 3$, puis infère à partir de là la formule générale. Pour $n = 2$, le problème s'écrit : on recherche x_1, x_2 tels que

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 = b_2 \end{cases} .$$

Lorsque $a_{1,1}a_{2,2} - a_{2,1}a_{1,2} \neq 0$, il trouve

$$x_1 = \frac{b_1a_{2,2} - b_2a_{1,2}}{a_{1,1}a_{2,2} - a_{2,1}a_{1,2}} ,$$
$$x_2 = \frac{b_2a_{1,1} - b_1a_{2,1}}{a_{1,1}a_{2,2} - a_{2,1}a_{1,2}} .$$

Pour $n = 3$, on recherche x_1, x_2, x_3 tels que

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3} = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3} = b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3} = b_3 \end{cases} .$$

Lorsque

$$a_{1,1}a_{2,2}a_{3,3}+a_{2,1}a_{3,2}a_{1,3}+a_{3,1}a_{1,2}a_{2,3}-a_{3,1}a_{2,2}a_{1,3}-a_{1,1}a_{3,2}a_{2,3}-a_{2,1}a_{1,2}a_{3,3} \neq 0 ,$$

il trouve

$$x_1 = \frac{b_1(a_{2,2}a_{3,3} - a_{3,2}a_{2,3}) + b_2(a_{3,2}a_{1,3} - a_{1,2}a_{3,3}) + b_3(a_{1,2}a_{2,3} - a_{2,2}a_{1,3})}{a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}} ,$$
$$x_2 = \frac{b_2(a_{1,1}a_{3,3} - a_{1,3}a_{3,1}) + b_3(a_{2,1}a_{1,3} - a_{1,1}a_{2,3}) + b_1(a_{3,1}a_{2,3} - a_{2,1}a_{3,3})}{a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}} ,$$
$$x_3 = \frac{b_3(a_{1,1}a_{2,2} - a_{2,1}a_{1,2}) + b_1(a_{2,1}a_{3,2} - a_{3,1}a_{2,2}) + b_2(a_{3,1}a_{1,2} - a_{1,1}a_{3,2})}{a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}} .$$

Cramer cherche alors la forme du dénominateur pour le cas ou n est un entier quelconque. Partant de sa formule pour $n = 3$, il remarque que tous les termes de la forme $a_{i,1}a_{j,2}a_{k,3}$ ou i, j et k sont tous distincts apparaissent une fois et une seule. En introduisant l'ensemble des permutations σ de $\{1, 2, 3\}$, chaque terme est donc de la forme $a_{\sigma(1),1}a_{\sigma(2),2}a_{\sigma(3),3}$ et est précédé d'un signe $+$ ou $-$, suivant ce que Cramer appelle le nombre de dérangements. Un dérangement est un couple (i, j) tel que $i < j$ et $\sigma(i) > \sigma(j)$. Le nombre de dérangements est donc la signature de la permutation σ.

Dans le cas général, Cramer conjecture que le déterminant comprendra un nombre de termes égal au nombre de permutations de $\{1, \dots, n\}$, chaque terme sera de la forme $a_{\sigma(1),1}a_{\sigma(2),2} \dots a_{\sigma(n),n}$ et son signe sera la signature $\varepsilon(\sigma)$ de la permutation σ. Notons S_n l'ensemble des permutations de $\{1, \dots, n\}$. Cramer conjecture la forme générale suivante

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1}a_{\sigma(2),2} \dots a_{\sigma(n),n} \quad . \tag{3.1}$$

Cramer considérait des équations avec autant d'équations que d'inconnues. Ces systèmes peuvent être écrits sous la forme matricielle suivante

$$\mathbf{Ax} = \mathbf{b} \quad ,$$

où $\mathbf{A}$ est la matrice carrée $\mathbf{A} = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ et où $\mathbf{x}, \mathbf{b}$ sont les vecteurs (colonnes) de $\mathbb{R}^n$ définis respectivement par

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad .$$

Le *déterminant* de $\mathbf{A}$, défini par la formule (3.1), sera noté dans la suite $\mathrm{Det}(\mathbf{A})$. Lorsque $\mathbf{A}$ est définie par ses coefficients numériques, par exemple si

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & -3 \\ 1 & 2 & 0 & 0 \\ -1 & -1 & 1 & -1 \\ 0 & 0 & 6 & 2 \end{bmatrix} \quad , \tag{3.2}$$

alors, on notera aussi le déterminant de A en utilisant la notation de Cayley suivante

$$\mathrm{Det}(\mathbf{A}) = \begin{vmatrix} 1 & 0 & 2 & -3 \\ 1 & 2 & 0 & 0 \\ -1 & -1 & 1 & -1 \\ 0 & 0 & 6 & 2 \end{vmatrix} \quad .$$

Un des objectifs de ce chapitre est de donner des méthodes permettant de calculer le déterminant d'une matrice plus rapidement qu'avec les formules de Cramer.

3.2 Calcul des déterminants

Bézout reprend le travail de Cramer. Il remarque qu'on peut écrire le dénominateur pour $n = 3$ en faisant apparaître trois déterminants du cas $n = 2$.

En effet, on a par exemple

$$\begin{aligned}
&a_{1,1}a_{2,2}a_{3,3}+a_{2,1}a_{3,2}a_{1,3}+a_{3,1}a_{1,2}a_{2,3}-a_{3,1}a_{2,2}a_{1,3}-a_{1,1}a_{3,2}a_{2,3}-a_{2,1}a_{1,2}a_{3,3} \\
&= a_{1,1}(a_{2,2}a_{3,3}-a_{3,2}a_{2,3})-a_{1,2}(a_{2,1}a_{3,3}-a_{3,1}a_{2,3})+a_{1,3}(a_{2,1}a_{3,2}-a_{3,1}a_{2,2}) \ .
\end{aligned} \tag{3.3}$$

Pour comprendre cette formule, il est utile d'introduire la notion de *matrice extraite.*

Definition 79. *Etant donnée une matrice carrée* $\boldsymbol{A}$ *d'ordre* n *et de coefficient générique* $(a_{i,j})$ *et un couple* (k,l) *d'entiers tels que* $1 \leqslant k,l \leqslant n$*, on note* $\boldsymbol{A}_{-(k,l)}$ *la matrice carrée d'ordre* $n-1$ *obtenue en supprimant la* k*-ième ligne et la* l*-ième colonne de* $\boldsymbol{A}$*.*

Exemple. Pour la matrice A définie à l'équation (3.2), on a

$$\mathbf{A}_{-(2,3)} = \begin{bmatrix} 1 & 0 & -3 \\ -1 & -1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \ .$$

La remarque de Bézout suggère une nouvelle définition pour le déterminant. Dans cette approche, le déterminant est défini par récurrence sur le rang n de la matrice A de la façon suivante. Nous notons $\det(\mathbf{A})$ (avec un d minuscule) le déterminant défini ainsi. Nous verrons en Section 3.4 que cette nouvelle définition coïncide avec la précédente.

Definition 80. *Soit* $\boldsymbol{A}$ *une matrice carrée d'ordre* n*. Le déterminant de* $\boldsymbol{A}$*,* $\det(\boldsymbol{A})$*, est défini récursivement de la façon suivante :*

— *si* $n=0$, $\det(\boldsymbol{A}) = 1$ *par convention,*
— *si* $n=1$, $\det(\boldsymbol{A}) = a_{1,1}$*,*
— *si* $n>1$*,*

$$\begin{aligned}
\det(\boldsymbol{A}) &= a_{1,1}\det(\boldsymbol{A}_{-(1,1)}) - a_{1,2}\det(\boldsymbol{A}_{-(1,2)}) + \ldots + (-1)^{n+1}a_{1,n}\det(\boldsymbol{A}_{-(1,n)}) \\
&= \sum_{i=1}^{n}(-1)^{i+1}a_{1,i}\det(\boldsymbol{A}_{-(1,i)}) \ .
\end{aligned}$$

Lorsqu'on applique la formule de récurrence de Bézout, on dit qu'on développe le déterminant par rapport à la première ligne.

Exemple. Si $n=2$, on a

$$\det(\mathbf{A}) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} \ .$$

On retrouve la formule de Cramer pour le cas $n=2$. Si maintenant $n=3$,

$$\begin{aligned}
\det(\mathbf{A}) &= a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\
&= a_{1,1}(a_{2,2}a_{3,3} - a_{2,3}a_{3,2}) - a_{1,2}(a_{2,1}a_{3,3} - a_{2,3}a_{3,1}) + a_{1,3}(a_{2,1}a_{3,2} - a_{2,2}a_{3,1}) \ .
\end{aligned}$$

D'après la remarque de Bézout (3.3), on a donc bien

$$\det(\mathbf{A}) = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} \ .$$

Proposition 81. *Pour tout $n \geqslant 1$,*

$$\det(\boldsymbol{I}_n) = 1 \ .$$

Preuve : Vérifiez le ! $\square$

Calculons le déterminant de la matrice $\mathbf{A}$ définie à l'équation (3.2).

$$\begin{aligned}
\det(\mathbf{A}) &= \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 6 & 2 \end{bmatrix} + 2\begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 2 \end{bmatrix} + 3\begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 0 & 6 \end{bmatrix} \\
&= 2\begin{bmatrix} 1 & -1 \\ 6 & 2 \end{bmatrix} + 2\left(\begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix} - 2\begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix}\right) + 3\left(\begin{bmatrix} -1 & 1 \\ 0 & 6 \end{bmatrix} - 2\begin{bmatrix} -1 & 1 \\ 0 & 6 \end{bmatrix}\right) \\
&= 2(2+6) + 2(-2-2(-2)) + 3(-6-2(-6)) \\
&= 16 + 4 + 18 = 38 \ .
\end{aligned}$$

3.3 Le caractère alterné

Definition 82. *Soit E un espace vectoriel sur un corps K (généralement $K = \mathbb{R}$, mais on peut vérifier sans peine que les résultats restent valables sur le corps $\mathbb{C}$ des nombres complexes). Soit $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E et soient $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ n vecteurs de E. Soit $\boldsymbol{A}$ la matrice dont la colonne i contient les coefficients de u_i dans la base B. On définit*

$$\det{}_B(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = \det(\boldsymbol{A}) \ .$$

Vandermonde et Laplace étudient les déterminants en les voyant ainsi comme des fonctions de n vecteurs. Vandermonde montre que ces fonctions
- s'annulent lorsque deux de ces vecteurs sont identiques,
- prennent des valeurs opposées lorsqu'on permute deux vecteurs.

Laplace redémontre (de façon propre) les résultats de Vandermonde. Résumons les en une proposition.

Proposition 83. *Soit σ une permutation de $\{1, \ldots, n\}$. Alors, pour tout n-uplets de vecteurs $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ de $\mathbb{R}^n$ et toute base B de $\mathbb{R}^n$,*

$$\det{}_B(\boldsymbol{u}_{\sigma(1)}, \ldots, \boldsymbol{u}_{\sigma(n)}) = \varepsilon(\sigma)\det{}_B(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \ .$$

En particulier, le déterminant change de signe quand on échange deux colonnes et il est nul lorsque deux vecteurs $\boldsymbol{u}_i$ et $\boldsymbol{u}_j$ sont égaux.

Preuve : On raisonne par récurrence sur n. Comme S_n est engendré par les transpositions, il suffit de montrer le cas particulier ou σ est une transposition $t_{k,l}$. Notons A' la matrice obtenue en échangeant les colonnes k et l de A.

$$\begin{aligned}
\det(A') = a_{1,1}\det(A'_{-(1,1)}) - a_{1,2}\det(A'_{-(1,2)}) + \ldots + (-1)^{1+k}\det(A'_{-(1,k)}) + \\
\ldots + (-1)^{1+l}\det(A'_{-(1,l)}) + \ldots + (-1)^{1+n}\det(A'_{-(1,n)}) \ .
\end{aligned}$$

Si $j \notin \{k, l\}$, on a $\det(A'_{-(1,j)}) = -\det(A_{-(1,j)})$ par l'hypothèse de récurrence. D'autre part,

$$A'_{-(1,k)} = \begin{bmatrix} a_{1,1} & \ldots & a_{1,k-1} & a_{1,k+1} & \ldots & a_{1,l-1} & a_{1,k} & a_{1,l+1} & \ldots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,k-1} & a_{n,k+1} & \ldots & a_{n,l-1} & a_{n,k} & a_{n,l+1} & \ldots & a_{n,n} \end{bmatrix} \ .$$

$$A_{-(1,l)} = \begin{bmatrix} a_{1,1} & \dots & a_{1,k-1} & a_{1,k} & \dots & a_{1,l-1} & a_{1,l+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,k-1} & a_{n,k} & \dots & a_{n,l-1} & a_{n,l+1} & \dots & a_{n,n} \end{bmatrix} .$$

Par l'hypothèse de récurrence, on a donc

$$\begin{aligned} \det(A'_{-(1,k)}) &= \varepsilon((k, k+1, \dots, l-1)) \det(A_{-(1,l)}) \\ &= (-1)^{l-k+1} \det(A_{-(1,l)}) \ . \end{aligned}$$

On montre de la même façon que $\det(A'_{-(1,l)}) = (-1)^{k-l+1} \det(A_{-(1,k)})$. On a donc établi que

$$\begin{aligned} \det(A') = -a_{1,1} \det(A_{-(1,1)}) + a_{1,2} \det(A'_{-(1,2)}) + \dots + (-1)^{1+k}(-1)^{1+l-k} \det(A_{-(1,l)}) + \\ \dots + (-1)^{1+l}(-1)^{1+k-l} \det(A_{-(1,k)}) + \dots - (-1)^{1+n} \det(A_{-(1,n)}) \\ = -\det(A) = \varepsilon(t_{k,l}) \det(A) \ . \end{aligned}$$

Si les vecteurs u_k et u_l sont égaux, alors, on a d'une part

$$\det_B(u_1, \dots, u_k, \dots, u_l, \dots, u_n) = \det_B(u_1, \dots, u_l, \dots, u_k, \dots, u_n) \ ,$$

puisque ces deux matrices sont égales. Mais d'autre part, d'après la première partie du théorème,

$$\det_B(u_1, \dots, u_k, \dots, u_l, \dots, u_n) = -\det_B(u_1, \dots, u_l, \dots, u_k, \dots, u_n) \ .$$

On en déduit que, nécessairement,

$$\det_B(u_1, \dots, u_k, \dots, u_l, \dots, u_n) = 0 \ .$$

□

3.4 Multilinéarité

Definition 84. *Soient E un K-espace vectoriel et p un entier. On appelle forme p-linéaire sur E une application $\varphi : E^p \to K$ telle que, pour tout $i \in \{1, \dots, p\}$ et tout famille $(u_i)_{i \in \{1,\dots,p\}}$, l'application*

$$u \mapsto \varphi(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_p)$$

est linéaire.

Par exemple, si $p = 3$, $\varphi : E^3 \to K$ est une forme 3-linéaire si toutes les applications $u \mapsto \varphi(u, u_2, u_3)$, $u \mapsto \varphi(u_1, u, u_3)$ et $u \mapsto \varphi(u_1, u_2, u)$ sont linéaires. On parle également de formes multilinéaires sans préciser le p.

Proposition 85. *Soit B une base de E. L'application $\det_B : E^n \to K$ est n-linéaire.*

Preuve : Raisonnons par récurrence sur n. Soit $\mathbf{A}$ une matrice carrée dont les colonnes définissent des vecteurs $\mathbf{u}_1, \dots, \mathbf{u}_n$. On suppose qu'il existe k, $\mathbf{u}'_k$, $\mathbf{u}''_k$ et $\lambda \in K$ tels que $\mathbf{u}_k = \mathbf{u}'_k + \lambda \mathbf{u}''_k$. On écrit

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_k & \dots & \mathbf{u}_n \end{bmatrix} \ , \\ \mathbf{A}' &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}'_k & \dots & \mathbf{u}_n \end{bmatrix} \ , \\ \mathbf{A}'' &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}''_k & \dots & \mathbf{u}_n \end{bmatrix} \ . \end{aligned}$$

On a

$$\det(A) = a_{1,1}\det(A_{-(1,1)}) + \ldots + (-1)^{1+k}(a'_{1,k} + \lambda a''_{1,k})\det(A_{-(1,k)}) \\ + \ldots + (-1)^{1+n}\det(A_{-(1,n)})$$

Par hypothèse de récurrence, pour tout $j \neq k$,

$$\det(A_{-(1,j)}) = \det(A'_{-(1,j)}) + \lambda\det(A''_{-(1,j)}) \ .$$

De plus, $\det(A_{-(1,k)}) = \det(A'_{-(1,k)}) = \det(A''_{-(1,k)})$, donc

$$\det(A) = \det(A') + \lambda\det(A) \ .$$

□

Attention On a, si $\mathbf{A}$ est une matrice carrée dont les colonnes sont les coordonnées des vecteurs $\mathbf{u}_1, \ldots, \mathbf{u}_n$ dans une base B et si $\lambda \in K$,

$$\det(\lambda\mathbf{A}) = \det{}_B(\lambda\mathbf{u}_1, \ldots, \lambda\mathbf{u}_n) = \lambda^n \det{}_B(\mathbf{u}_1, \ldots, \mathbf{u}_n) = \lambda^n\det(\mathbf{A}) \ .$$

En particulier donc $\det(\lambda\mathbf{I}_n) = \lambda^n$.

Remarque : La propriété de multilinéarité du déterminant a été découverte par Catalan, 96 ans après les travaux de Cramer !!

Definition 86. *Une forme p-linéaire $\varphi : E^p \to K$ est dite alternée si, pour tout $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ de E^p, $\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$ s'annule dès qu'il existe $i \neq j$ tels que $\boldsymbol{u}_i = \boldsymbol{u}_j$.*

On a vu que le déterminant était une forme n-linéaire alternée. On peut généraliser les autres conclusions de la Proposition 83 aux formes p-linéaires alternées.

Proposition 87. *Soit $\varphi : E^p \to K$ une forme p-linéaire alternée et $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ dans E. Alors, on a, pour tous $i < j$ de $\{1, \ldots, n\}$,*

$$\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i, \ldots, \boldsymbol{u}_j, \ldots, \boldsymbol{u}_p) = -\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j, \ldots, \boldsymbol{u}_i, \ldots, \boldsymbol{u}_p) \ .$$

De plus, pour tout $\sigma \in S_n$,

$$\varphi(\boldsymbol{u}_{\sigma(1)}, \ldots, \boldsymbol{u}_{\sigma(p)}) = \varepsilon(\sigma)\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p) \ .$$

Démonstration. Pour la première partie, on a, comme φ est alternée,

$$\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i + \mathbf{u}_j, \ldots, \mathbf{u}_i + \mathbf{u}_j, \ldots, \mathbf{u}_p) = 0 \ .$$

Comme φ est p-linéaire, on a de plus

$$\begin{aligned}\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i + \mathbf{u}_j, \ldots, \mathbf{u}_i + \mathbf{u}_j, \ldots, \mathbf{u}_p) &= \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_p) \\ &+ \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_p) \\ &+ \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_p) \\ &+ \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_p) \ .\end{aligned}$$

Le premier et le dernier terme du membre de droite sont nuls car φ est alternée, donc

$$\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_p) + \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_p) = 0 \ .$$

Pour la seconde partie, on décompose $\sigma = t_1 \circ \ldots \circ t_k$ en transpositions et on applique récursivement le résultat de la première partie de la proposition. □

Proposition 88. *Soit $\varphi : E^p \to K$ une forme p-linéaire alternée. Alors, $\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p) = 0$ si la famille $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ est liée. De plus, on a, pour tout i et tout $\lambda_1, \ldots, \lambda_p$ de K,*

$$\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i, \ldots, \boldsymbol{u}_p) = \varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i + \sum_{j \neq i} \lambda_j \boldsymbol{u}_j, \ldots, \boldsymbol{u}_p) \ .$$

Preuve : On montre d'abord la seconde partie de la proposition. On a par multilinéarité

$$\begin{aligned}\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i + \sum_{j \neq i} \lambda_j \mathbf{u}_j, \ldots, \mathbf{u}_p) &= \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_p) \\ &\quad + \sum_{j \neq i} \lambda_j \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_p) \ .\end{aligned}$$

Tous les termes de la somme du second membre de droite sont nuls car φ est alternée.

Pour la première partie, si $\mathbf{u}_1, \ldots, \mathbf{u}_n$ sont liés, il existe $\lambda_1, \ldots, \lambda_p$ non tous nuls tels que

$$\sum_{i=1}^{n} \lambda_i \mathbf{u}_i = 0 \ .$$

Supposons que $\lambda_i \neq 0$, on a alors

$$\mathbf{u}_i + \sum_{j \neq i} \lambda'_j \mathbf{u}_j = 0, \qquad \text{où} \qquad \lambda'_j = \frac{\lambda_j}{\lambda_i} \ .$$

Comme φ est p-linéaire,

$$0 = \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i + \sum_{j \neq i} \lambda'_j \mathbf{u}_j, \ldots, \mathbf{u}_p) \ .$$

D'après la seconde partie de la proposition, on en déduit donc que

$$\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_i, \ldots, \mathbf{u}_p) = 0 \ .$$

□

Supposons jusqu'à la fin de cette section que E est un espace vectoriel de dimension n. On étudie les formes n-linéaires alternées sur E.

Proposition 89. *Soit $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E. Soit $\varphi : E^n \to K$ une forme n-linéaire alternée. Alors, pour toute famille $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ de vecteurs de E, on a*

$$\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = \mathrm{Det}_B(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \varphi(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \ .$$

Preuve : Soit $\mathbf{u}_1, \ldots, \mathbf{u}_n$ une famille de vecteurs de E et soit $\mathbf{A}$ la matrice telle que, pour tout i, la i-ème colonne est formée des coefficients du vecteur $\mathbf{u}_i$ dans la base B. Ainsi,

$$\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_n) = \varphi\left(\sum_{i=1}^{n} a_{i,1} \mathbf{e}_i, \ldots, \sum_{i=1}^{n} a_{i,n} \mathbf{e}_i \right) \ .$$

En développant cette expression par multilinéarité et en utilisant le caractère alterné de φ, on voit que tous les termes de la somme sont nuls, sauf ceux de la forme $a_{\sigma(1),1} \dots a_{\sigma(n),n}\varphi(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)})$, où σ est une permutation de $\{1, \dots, n\}$. D'après la Proposition 87, ces termes sont de la forme

$$a_{\sigma(1),1} \dots a_{\sigma(n),n}\varepsilon(\sigma)\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) \ .$$

On a ainsi montré que

$$\begin{aligned}\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) \\ &= \mathrm{Det}(\mathbf{A})\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = \mathrm{Det}_B(\mathbf{A})\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) \ .\end{aligned}$$

□

La Proposition 89 permet de montrer l'égalité des deux définitions de déterminant.

Corollaire 90. *Les applications* det *et* Det *coïncident.*

Preuve : Soit B une base de E. D'après les Propositions 83 et 85, $\det_B$ est une forme n-linéaire alternée. D'après la Proposition 89, on a donc

$$\det_B(\mathbf{u}_1, \dots, \mathbf{u}_n) = \mathrm{Det}_B(\mathbf{u}_1, \dots, \mathbf{u}_n)\det_B(\mathbf{e}_1, \dots, \mathbf{e}_n) \ .$$

Enfin, d'après la Proposition 81, $\det_B(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, ce qui conclut la preuve du corollaire. □

Le Corollaire 90 permet de reformuler la Proposition 89.

Proposition 91. *Soit $B = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E. Soit $\varphi : E^n \to K$ une forme n-linéaire alternée. Alors, pour toute famille $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ de vecteurs de E, on a*

$$\varphi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \det_B(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)\varphi(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) \ .$$

3.5 Règles de calcul de déterminants

La règle de Sarrus pour les matrices de taille 3 procède de la façon suivante. On écrit la matrice $\mathbf{A}_c$ obtenue en recopiant les deux premières colonnes de A de la façon suivante.

$$\mathbf{A}_c = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \end{bmatrix} \ .$$

Alors, le déterminant de A est égal à la somme des produits des coefficients de chaque diagonale de $\mathbf{A}_c$ moins la somme des produits des coefficients de chaque antidiagonale de $\mathbf{A}_c$.

$$\det(\mathbf{A}) = (a_{1,1}a_{2,2}a_{3,3}+a_{1,2}a_{2,3}a_{3,1}+a_{1,3}a_{2,1}a_{3,2})-(a_{1,2}a_{2,1}a_{3,3}+a_{1,1}a_{2,3}a_{3,2}+a_{1,3}a_{2,2}a_{3,1}) \ .$$

Proposition 92 (Déterminant de la transposée)**.** *Soit $\boldsymbol{A}$ une matrice carrée d'ordre n. Alors,* $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$.

Preuve : Soit $a_{i,j}$ le coefficient générique de $\mathbf{A}$ et $b_{i,j}$ celui de $\mathbf{A}^T$. Rappelons que $b_{i,j} = a_{j,i}$ pour tout couple d'entiers i, j. Par la formule de Cramer, on a

$$\det(\mathbf{A}^T) = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1),1} \ldots b_{\sigma(n),n} = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1,\sigma(1)} \ldots a_{n,\sigma(n)} \ .$$

En réarrangeant les termes du produit, on a donc

$$\det(\mathbf{A}^T) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma^{-1}(1),1} \ldots a_{\sigma^{-1}(n),n} \ .$$

Comme $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$, on a donc

$$\det(\mathbf{A}^T) = \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) a_{\sigma^{-1}(1),1} \ldots a_{\sigma^{-1}(n),n} \ .$$

Dans cette dernière somme, tous les termes de la forme $\varepsilon(\sigma) a_{\sigma(1),1} \ldots a_{\sigma(n),n}$ apparaissent une fois et une seule, donc, finalement

$$\det(\mathbf{A}^T) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \ldots a_{\sigma(n),n} = \det(\mathbf{A}) \ .$$

□

La Proposition 92 permet d'obtenir immédiatement les propriétés suivantes.
— Le déterminant change de signe si on échange deux lignes.
— Le déterminant est linéaire par rapport à chaque ligne.
— Le déterminant ne change pas si on ajoute à une ligne une combinaison linéaire des autres.

Proposition 93 (Déterminant des matrices triangulaires)**.** *Soit $\boldsymbol{A}$ une matrice triangulaire de coefficient générique $a_{i,j}$. Alors* $\det(\boldsymbol{A}) = a_{1,1} \ldots a_{n,n}$.

Preuve : On procède par récurrence sur n et on développe par rapport à la première colonne si $\mathbf{A}$ est une matrice triangulaire supérieure. Le résultat pour les matrices triangulaires inférieures se déduit alors de la Proposition 92. □

La méthode du pivot Soit $\mathbf{A}$ une matrice d'ordre n dont les colonnes définissent des vecteurs $\mathbf{u}_1, \ldots, \mathbf{u}_n$ par rapport à la base canonique. Pour calculer le déterminant de $\mathbf{A}$, on peut utiliser la méthode du pivot sur les vecteurs $\mathbf{u}_1, \ldots, \mathbf{u}_n$ puisque le déterminant change de signe si on permute deux colonnes et que sa valeur ne change pas si on ajoute à une colonne une combinaison linéaire des autres. Une fois qu'on s'est ramené par cette méthode à une matrice triangulaire, il suffit d'appliquer la Proposition 93.

D'après la Proposition 92, on peut également appliquer la méthode du pivot sur lignes.

Ces méthodes sont efficaces pour calculer la valeur numérique du déterminant en général. Pour certaines matrices (en particulier celles “creuses"), on pourra aussi essayer d'appliquer une des méthodes suivantes.

Développement par rapport à une ligne ou une colonne quelconque On a, pour tout i et j de $\{1,\ldots,n\}$,

$$\det(\mathbf{A}) = \sum_{k=1}^{n}(-1)^{i+k}a_{i,k}\det(\mathbf{A}_{-(i,k)}) = \sum_{k=1}^{n}(-1)^{j+k}a_{k,j}\det(\mathbf{A}_{-(k,j)}) \ . \quad (3.4)$$

Exercice : Démontrer ces relations !

Exemple En développant par rapport à la 3ème colonne on a

$$\begin{vmatrix} 1 & 2 & 0 & 5 \\ -1 & 1 & 2 & 1 \\ 4 & 0 & 0 & 6 \\ -1 & 0 & 0 & 4 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 5 \\ 4 & 0 & 6 \\ -1 & 0 & 4 \end{vmatrix} \ .$$

En développant par rapport à la 2ème colonne, il vient alors

$$\begin{vmatrix} 1 & 2 & 0 & 5 \\ -1 & 1 & 2 & 1 \\ 4 & 0 & 0 & 6 \\ -1 & 0 & 0 & 4 \end{vmatrix} = 4\begin{vmatrix} 4 & 6 \\ -1 & 4 \end{vmatrix} = 88 \ .$$

3.6 Déterminant d'un endomorphisme

Proposition 94 (Caractérisation des bases)**.** *Soit E un espace vectoriel de dimension n et $B = (\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)$ une base de E. Une famille $C = (\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_n)$ de n vecteurs de E est une base de E si et seulement si $\det_B(\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_n) \neq 0$. Dans ce cas, on a $\det_B(\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_n)\det_C(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n) = 1$.*

Preuve : Si C n'est pas une base, la famille est liée, donc $\det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n) = 0$ par la Proposition 88. Si C est une base, alors, d'après la Proposition 91,

$$1 = \det_C(\mathbf{e}'_1,\ldots,\mathbf{e}'_n) = \det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)\det_C(\mathbf{e}_1,\ldots,\mathbf{e}_n) \ .$$

□

Proposition 95. *Soient E un espace vectoriel de dimension n, $B = (\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)$ et $C = (\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_n)$ deux bases de E et $f : E \to E$, un endomorphisme de E. Alors*

$$det_C(f(\boldsymbol{e}'_1),\ldots,f(\boldsymbol{e}'_n)) = det_B(f(\boldsymbol{e}_1),\ldots,f(\boldsymbol{e}_n)) \ .$$

Preuve : D'après la Proposition 91 appliquée à la forme n-linéaire $\varphi = \det_B$ dans la base C,

$$\det_B(f(\mathbf{e}'_1),\ldots,f(\mathbf{e}'_n)) = \det_C(f(\mathbf{e}'_1),\ldots,f(\mathbf{e}'_n))\det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n) \ .$$

D'après la Proposition 91 appliquée à la forme n-linéaire $\varphi : (\mathbf{u}_1,\ldots,\mathbf{u}_n) = \det_B(f(\mathbf{u}_1),\ldots,f(\mathbf{u}_n))$ dans la base B,

$$\det_B(f(\mathbf{e}'_1),\ldots,f(\mathbf{e}'_n)) = \det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)\det_B(f(\mathbf{e}_1),\ldots,f(\mathbf{e}_n)) \ .$$

On a donc

$$\det_C(f(\mathbf{e}'_1),\ldots,f(\mathbf{e}'_n))\det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n) = \det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)\det_B(f(\mathbf{e}_1),\ldots,f(\mathbf{e}_n)) \ .$$

Comme $\det_B(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)$ est non nul d'après la Proposition 94, on a bien le résultat. □

La proposition 95 montre que $\det_B(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n))$ ne dépend pas de la base choisie. Dans la suite, on notera $\det(f)$ la valeur commune de ces déterminants. En appliquant la Proposition 91 à la forme linéaire $\varphi : (\mathbf{u}_1, \ldots, \mathbf{u}_n) = \det_B(f(\mathbf{u}_1), \ldots, f(\mathbf{u}_n))$ dans la base B, on obtient la proposition suivante.

Proposition 96 (Formule du volume)**.** *Soit E un espace vectoriel de dimension n et f un endomorphisme de E. Pour toute base B de E, on a*

$$\det_B(f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_n)) = \det(f)\det_B(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \ .$$

Proposition 97 (Critère d'inversibilité)**.** *Soit E un espace vectoriel de dimension n et f un endomorphisme de E. f est inversible si et seulement si* $\det(f) \neq 0$.

Preuve : Soit $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E. f est inversible si et seulement si $C = (f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n))$ est une base de E. D'après la Proposition 94, C est une base de E si et seulement si $\det_B(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)) \neq 0$. Comme $\det_B(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)) = \det(f)$, la preuve est terminée. □

Proposition 98 (Déterminant de la composée)**.** *Soit E un espace vectoriel de dimension n et soient f et g deux endomorphismes de E. Alors,* $\det(f \circ g) = \det(f)\det(g)$.

Preuve : Soit $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E. La formule du volume appliquée aux vecteurs $g(\mathbf{e}_1), \ldots, g(\mathbf{e}_n)$ donne

$$\begin{aligned}\det(f \circ g) &= \det_B(f \circ g(\mathbf{e}_1), \ldots, f \circ g(\mathbf{e}_n)) \\ &= \det(f)\det_B(g(\mathbf{e}_1), \ldots, g(\mathbf{e}_n)) = \det(f)\det(g) \ .\end{aligned}$$

□

Déterminant de la matrice d'un endomorphisme. Soit f un endomorphisme d'un espace vectoriel E, de matrice $\mathbf{A}$ dans une base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Alors, le déterminant de f,

$$\det(f) = \det_B(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)) = \det(\mathbf{A}) \ . \tag{3.5}$$

3.7 Déterminant d'une matrice carrée.

Les propriétés des déterminants d'endomorphismes fournissent immédiatement les propriétés suivantes de déterminants de matrices via l'identité (3.5).

— $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{BA})$.
— $\mathbf{A}$ est inversible si et seulement si $\det(\mathbf{A}) \neq 0$.
— Si $\mathbf{A}$ est inversible $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.
— Si $\mathbf{A}$ et $\mathbf{B}$ sont semblables, alors $\det(\mathbf{A}) = \det(\mathbf{B})$.

Exercice : Vérifier ces propriétés !
Commentaires : La première propriété a été établie en 1812 par Cauchy au terme de longs calculs.

Cofacteurs.

Definition 99. *On appelle cofacteur* $\mathrm{cof}_{i,j}(\boldsymbol{A})$ *la quantité*

$$\mathrm{cof}_{i,j}(\boldsymbol{A}) = (-1)^{i+j}\det(\boldsymbol{A}_{-(i,j)}) \ .$$

La matrice $\mathrm{cof}(\boldsymbol{A})$ *de taille* n*, de coefficient générique* $\mathrm{cof}_{i,j}(\boldsymbol{A})$ *est appelée la matrice des cofacteurs de* $\boldsymbol{A}$*. Elle vérifie en particulier* $\mathrm{cof}(\boldsymbol{A})^T = \mathrm{cof}(\boldsymbol{A}^T)$.

Proposition 100 (Formule de l'inverse)**.** *Soit* $\boldsymbol{A}$ *une matrice carrée d'ordre* n*. On a*

$$\boldsymbol{A}\mathrm{cof}(\boldsymbol{A})^T = \mathrm{cof}(\boldsymbol{A})^T\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}_n \ .$$

En particulier, si $\boldsymbol{A}$ *est inversible,* $\boldsymbol{A}^{-1} = (1/\det(\boldsymbol{A}))\mathrm{cof}(\boldsymbol{A})^T$.

Preuve : il suffit d'établir la première propriété, la seconde en découle directement. Pour tout couple (i,j), le coefficient d'indice (i,j) de $\mathbf{A}\mathrm{cof}(\mathbf{A})^T$ est

$$\sum_{k=1}^{n} a_{i,k}\mathrm{cof}_{j,k}(\mathbf{A}) = \sum_{k=1}^{n} a_{i,k}(-1)^{j+k}\det(\mathbf{A}_{-(j,k)}) \ .$$

Si $i = j$, ce coefficient vaut

$$\sum_{k=1}^{n}(-1)^{i+k}a_{i,k}\det(\mathbf{A}_{-(i,k)})$$

D'après l'équation (3.4), c'est donc $\det(\mathbf{A})$. Si $i \neq j$, ce coefficient est celui du déterminant de la matrice $\mathbf{A}'$ suivant sa j-ème ligne, la matrice $\mathbf{A}'$ étant obtenue en remplaçant la j-ème ligne de $\mathbf{A}$ par sa i-ème ligne. La matrice $\mathbf{A}'$ ayant deux lignes égales a un déterminant nul, ce qui conclut la preuve. □

Exercice. Calculer avec la règle précédente l'inverse de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -1 \\ 3 & -1 & 1 \\ -2 & 1 & -3 \end{bmatrix} \ .$$

Definition 101. *Un système de Cramer est un système de la forme* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$*, où* $\boldsymbol{A}$ *est une matrice carrée inversible. L'inconnue est* $\boldsymbol{x}$*.*

La solution d'un système de Cramer est égale à

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y} \ .$$

D'après la Proposition 100, on a donc

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\mathrm{cof}(\boldsymbol{A})^T\boldsymbol{y} \ .$$

La i-ème ligne de $\mathbf{x}$ est égale à

$$x_i = \frac{1}{\det(\mathbf{A})}\sum_{j=1}^{n}\mathrm{cof}_{j,i}(\mathbf{A})y_j = \frac{1}{\det(\mathbf{A})}\sum_{j=1}^{n}(-1)^{i+j}\det(\mathbf{A}_{-(j,i)})y_j \ .$$

Dans cette formule, le numérateur est le déterminant de la matrice en remplaçant dans $\mathbf{A}$ la i-ème colonne par $\boldsymbol{y}$.

En pratique, on utilise rarement ces formules pour calculer les solutions d'un système, car le nombre d'équations à résoudre croît rapidement avec la dimension. En revanche, ces formules sont fort utiles théoriquement comme nous le verrons dans les prochains chapitres.

3.8 Déterminant et rang

Definition 102 (Matrice extraite)**.** *Soit $\boldsymbol{A}$ une matrice carrée de taille n, I et J deux sous-ensembles de $\{1,\dots,n\}$. On note $\boldsymbol{A}_{I,J}$ la matrice obtenue en supprimant dans $\boldsymbol{A}$ les lignes dont les indices ne sont pas dans I et les colonnes dont les indices ne sont pas dans J. On note également $\boldsymbol{A}_{-(I,J)}$ la matrice obtenue en supprimant dans $\boldsymbol{A}$ les lignes dont les indices sont dans I et les colonnes dont les indices sont dans J.*

Proposition 103. *Une matrice $\boldsymbol{A}$ est de rang r si on peut en extraire une matrice $\boldsymbol{A}_{I,J}$ carrée, de taille r, telle que $\det(\boldsymbol{A}_{I,J}) \neq 0$ et si toute matrice extraite $\boldsymbol{A}_{I',J'}$ carrée de taille $r+1$ a un déterminant nul.*

Preuve : Supposons $\mathbf{A}$ de taille n et notons $\mathbf{u}_1,\dots,\mathbf{u}_n$ ses vecteurs colonnes. Par l'algorithme de pivot, on peut déterminer une sous famille $\mathbf{u}_{j_1},\dots,\mathbf{u}_{j_r}$ de r vecteurs indépendants. Cet algorithme construit avec r lignes $i_1,\dots,i_r$ des vecteurs $\mathbf{u}_{j_1},\dots,\mathbf{u}_{j_r}$ une famille triangulaire $w_1,\dots,w_r$ de vecteurs non nuls. Si $I=\{i_1,\dots,i_r\}$ et $J=\{j_1,\dots,j_r\}$, le déterminant de $\mathbf{A}_{I,J}$ est égal au produit des coefficient diagonaux de la matrice de colonnes $w_1,\dots,w_r$, il est non nul.

Soient maintenant I' et J' deux sous-ensembles de $\{1,\dots,n\}$ de taille $r+1$. La famille $(\mathbf{u}_i)_{i\in I'}$ est liée, il existe donc $(w_i)_{i\in I'} \neq 0$ tel que

$$\sum_{i\in I'} w_i\mathbf{u}_i = 0 \ .$$

En particulier, les $r+1$ lignes d'indices dans J' de ce système montrent que les colonnes de $\mathbf{A}_{I',J'}$ sont liées, donc que $\det(\mathbf{A}_{I',J'})=0$.

Pour la réciproque, soit $\mathbf{A}$ une matrice vérifiant les hypothèses de la proposition. On peut extraire de $\mathbf{A}$ une sous-matrice $\mathbf{A}_{I,J}$ de taille r inversible. Les vecteurs $(\mathbf{u}_i)_{i\in I}$ sont donc linéairement indépendants et $\mathbf{A}$ est donc de rang au moins r. Soit alors $k\notin I$. On applique alors l'algorithme du pivot à $(\mathbf{u}_i)_{i\in I\cup\{k\}}$. On obtient une famille triangulaire $\mathbf{v}_1,\dots,\mathbf{v}_r$ et un vecteur $\mathbf{v}_{r+1}$. $\mathbf{u}_k$ est linéairement indépendant de $(\mathbf{u}_i)_{i\in I}$ si $\mathbf{v}_{r+1}$, dont les r premières coordonnées sont nulles, a une coordonnée s non nulle. Dans ce cas, on aurait $\det(\mathbf{A}_{I\cup\{k\},J\cup\{s\}})\neq 0$. Par hypothèse, $\det(\mathbf{A}_{I\cup\{k\},J\cup\{s\}})=0$, donc $\mathbf{u}_k$ est linéairement lié à $(\mathbf{u}_i)_{i\in I}$. Comme c'est vrai pour tout $k\notin I$, $\mathbf{A}$ est de rang r. □

3.9 Exercices

Exercice 1

1. Calculer les déterminants des matrices suivantes.

$$A=\begin{bmatrix}1 & 3\\ -7 & 5\end{bmatrix},\qquad B=\begin{bmatrix}2 & 3 & 3\\ 5 & 15 & 7\\ 0 & -2 & 0\end{bmatrix},\qquad C=\begin{bmatrix}2 & -1 & 3 & 4\\ 2 & 0 & 4 & -5\\ -2 & 4 & 3 & 1\\ 0 & -3 & 1 & -1\end{bmatrix}\ .$$

2. Déterminer A^{-1}, B^{-1} en utilisant les formules de Cramer. (Calculer C^{-1} si vous voulez vérifier l'affirmation "le nombre de calculs augmente rapidement avec la dimension" du cours).

3. Calculer le déterminant des matrices suivantes.

$$D = \begin{bmatrix} 1 & \cos(a) & \cos(2a) \\ 1 & \cos(b) & \cos(2b) \\ 1 & \cos(c) & \cos(2c) \end{bmatrix}, \qquad E = \begin{bmatrix} \sin(a) & \sin(2a) & \sin(3a) \\ \sin(b) & \sin(2b) & \sin(3b) \\ \sin(c) & \sin(2c) & \sin(3c) \end{bmatrix}.$$

Exercice 2

[Jacobi 1827] Une matrice $\mathbf{A}$ de coefficient générique $a_{i,j}$ est dite antisymétrique si, pour tout i et j de $\{1, \ldots, n\}$, $a_{i,j} = -a_{j,i}$. Dans tout l'exercice, $\mathbf{A}$ désigne une matrice antisymétrique, de coefficient générique $a_{i,j}$.

1. Que dire de $a_{i,i}$?
2. Montrer que, si n est impair, $\det(\mathbf{A}) = 0$.
3. Calculer $\det(\mathbf{A})$ lorsque $n \in \{2, 4\}$. Montrer que $\det(\mathbf{A})$ est le carré d'un polynôme en les coefficients de la matrice. (Ce résultat s'étend à tout n de la forme $n = 2p$).
4. Calculer $\det(\mathbf{A})$ lorsque n est pair et que les coefficients au dessus de la diagonale sont tous nuls, sauf ceux de la forme $a_{i,i+1}$, $i = 1, \ldots, n-1$.

Exercice 3

Exprimer $\det(\mathrm{cof}(\mathbf{A}))$ en fonction de $\det(\mathbf{A})$.

Exercice 4

Dans l'exercice, les coefficients non explicités sont nul, par convention.

1. Calculer, pour $n \geqslant 2$,

$$D_n = \begin{vmatrix} a & b & & \\ & \ddots & \ddots & \\ & & a & b \\ b & & & a \end{vmatrix}.$$

2. Pour tout $n \geqslant 1$, on pose

$$D_n = \begin{vmatrix} z & x & \ldots & x \\ y & z & & \\ \vdots & & \ddots & \\ y & & & z \end{vmatrix}.$$

Déterminer une relation de récurrence entre D_n et D_{n-1} et en déduire la valeur de D_n.

3. Pour tout $n \geqslant 1$, on pose

$$D_n = \begin{vmatrix} z & x & & \\ y & z & \ddots & \\ & \ddots & \ddots & x \\ & & y & z \end{vmatrix}.$$

Trouver une relation de récurrence entre D_n, D_{n-1} et D_{n-2} et en déduire la valeur de D_n.

Exercice 5

1. Soit E un K-espace vectoriel de dimension finie n muni d'une base $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et soit f un endomorphisme de E. Soit
$$\varphi : E^n \to K, \qquad (\mathbf{u}_1, \dots, \mathbf{u}_n) \mapsto \sum_{i=1}^{n} \det_B(\mathbf{u}_1, \dots, f(\mathbf{u}_i), \dots, \mathbf{u}_n) \ .$$
Montrer que φ est une forme multilinéaire alternée et en déduire que
$$\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n) = \mathrm{Tr}(f)\det_B(\mathbf{u}_1, \dots, \mathbf{u}_n) \ .$$

2. Soit E un K-espace vectoriel de dimension finie n muni d'une base B et soit $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ une famille de n vecteurs de E. Montrer la formule suivante (Catalan 1846)
$$\det_B(\mathbf{u}_1 + \dots \mathbf{u}_n, \mathbf{u}_1 - \mathbf{u}_2, \dots, \mathbf{u}_{n-1} - \mathbf{u}_n) = n(-1)^{n-1}\det_B(\mathbf{u}_1, \dots, \mathbf{u}_n) \ .$$
En déduire
$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{vmatrix} \ .$$

Exercice 6

Soit K un corps, n un entier, $x_1, \dots, x_n$ des éléments de K. La matrice de Vandermonde associée est la matrice
$$M(x_1, \dots, x_n) = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix} \ .$$
Le déterminant de Vandermonde associé $V(x_1, \dots, x_n) = \det(M(x_1, \dots, x_n))$. Calculer $V(x_1, \dots, x_n)$.
Remarque : Vandermonde ne semble pas avoir travaillé sur les déterminant qui portent son nom.

Exercice 7

Donner, à l'aide du déterminant, une équation de l'hyperplan engendré par une famille libre $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ d'un K-espace vectoriel E de dimension n.

Exercice 8

Soit $\mathbf{A}$ la matrice carrée de taille n de coefficient générique $a_{i,j}$. On suppose que,
$$\forall i \in \{1, \dots, n\}, \qquad |a_{i,i}| > \sum_{j \in \{1,\dots,n\} \setminus \{i\}} |a_{i,j}| \ .$$

1. Montrer que $\mathbf{A}$ est inversible.
2. Si on suppose de plus que, pour tout $i \in \{1, \dots, n\}$, $a_{i,i} > 0$. Montrer qu'alors $\det(\mathbf{A}) > 0$.

Exercice 9

Soit $\mathbf{A}$ la matrice carrée de taille n de colonnes $\mathbf{A}_1, \dots, \mathbf{A}_n$ et $\mathbf{B}$ la matrice carrée de taille n de colonnes $\mathbf{B}_1, \dots, \mathbf{B}_n$ définies par

$$\forall i \in \{1, \dots, n\}, \qquad \mathbf{B}_i = \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \mathbf{A}_j \ .$$

Exprimer $\det(\mathbf{B})$ en fonction de $\det(\mathbf{A})$.

Exercice 10

Soit $\mathbf{A} \in M_n(\mathbb{R})$ une matrice dont les coefficients vérifient

$$\forall i, j \in \{1, \dots, n\}, a_{i,j} \geqslant 0, \qquad \forall i \in \{1, \dots, n\}, \ \sum_{j=1}^{n} a_{i,j} \leqslant 1 \ .$$

Montrer que $|\det(\mathbf{A})| \leqslant 1$.

Exercice 11

Soit $a, \lambda_1, \dots, \lambda_n \in \mathbb{C}$. Calculer le déterminant de la matrice suivante.

$$\mathbf{A} = \begin{bmatrix} a + \lambda_1 & a & \dots & a \\ a & a + \lambda_2 & \dots & a \\ \vdots & & \ddots & \vdots \\ a & \dots & & a + \lambda_n \end{bmatrix} .$$

Exercice 12

Soient $a_1, \dots, a_n$ et $b_1, \dots, b_n$ des complexes. Calculer le déterminant de la matrice $\mathbf{A}$ dont le coefficient générique $a_{i,j}$ est donné par

$$a_{i,j} = \begin{cases} a_i + b_i & \text{si } i = j \ , \\ b_j & \text{sinon } . \end{cases}$$

Exercice 13

Soient a, b, c des complexes. Calculer le déterminant de la matrice $\mathbf{A} \in M_n(\mathbb{C})$ définie par

$$\mathbf{A} = \begin{bmatrix} a & b & \dots & & b \\ c & a & \dots & & b \\ \vdots & & \ddots & & \vdots \\ c & & \dots & a & b \\ c & & \dots & c & a \end{bmatrix} .$$

Exercice 14

1. Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices de $M_n(\mathbb{R})$. Montrer que
$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & \mathbf{A} \end{pmatrix} \geqslant 0 \ .$$
2. Si $\mathbf{A}$ et $\mathbf{B}$ commutent ($\mathbf{AB} = \mathbf{BA}$) montrer que $\det(\mathbf{A}^2 + \mathbf{B}^2) \geqslant 0$.
3. Trouver un contre-exemple à la question 2 si $\mathbf{A}$ et $\mathbf{B}$ ne commutent pas.
4. Soient $\mathbf{C}$ et $\mathbf{D}$ deux autres matrices de $M_n(\mathbb{R})$ et supposons que $\mathbf{AC} = \mathbf{CA}$. Montrer que
$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det(\mathbf{AD} - \mathbf{CB}) \ .$$

Exercice 15

Soit $\mathbf{A}$ une matrice inversible carrée de taille n et $\mathbf{B}$ une matrice carrée de taille n. Montrer qu'il existe $\varepsilon > 0$ tel que, pour tout $x \in]-\varepsilon, \varepsilon[$, $\mathbf{A} + x\mathbf{B}$ est inversible.

Exercice 16

Soit $\mathbf{A}$ la matrice définie par
$$\mathbf{A} = \begin{bmatrix} 3 & -2 & -3 \\ -2 & 6 & 6 \\ 2 & -2 & -2 \end{bmatrix} \ .$$

1. Déterminer les valeurs de x pour lesquelles $\det(\mathbf{A} - x\mathbf{I}_3) = 0$.
2. Déterminer une matrice $\mathbf{P}$ telle que $\mathbf{PAP}^{-1}$ est diagonale.

Exercice 17

Déterminer les endomorphismes de $\mathbb{R}^n$ tels que $f(\mathbb{Z}^n) = \mathbb{Z}^n$.

Exercice 18

Soit $\mathbf{A}$ et $\mathbf{B}$ deux matrices qui commutent, montrer que les comatrices de $\mathbf{A}$ et $\mathbf{B}$ commutent.

Exercice 19

1. Déterminer le rang de $\mathbf{B} = \mathrm{cof}(\mathbf{A})^T$ en fonction de celui de $\mathbf{A}$.
2. Supposons que le rang de $\mathbf{A}$ est $n-1$ et soit $\mathbf{C}$ une matrice telle que $\mathbf{AC} = \mathbf{CA} = 0$. Montrer qu'il existe $\lambda > 0$ telle que $\mathbf{C} = \lambda\mathbf{B}$.

Chapitre 4

Réduction des endomorphismes

Dans tous le chapitre, K est un corps, E est un K-espace vectoriel de dimension finie n et f est un endomorphisme de E. On recherche une base de E dans laquelle f ait une forme diagonale si possible, ou triangulaire. Soit f l'endomorphisme de E dont la matrice dans une base $B = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ s'écrit

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{bmatrix} \ .$$

Considérons la base $B' = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ où $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (1,3,2)$, $\mathbf{v}_3 = (1,2,1)$. Montrer que

$$\mathbf{A}\mathbf{v}_1 = 0, \qquad \mathbf{A}\mathbf{v}_2 = \mathbf{v}_2, \qquad \mathbf{A}\mathbf{v}_3 = 2\mathbf{v}_3 \ .$$

En déduire, que, dans la base B', f a pour matrice

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \ .$$

Notons $\mathbf{P}$ la matrice de l'identité entre (E, B') et (E, B). $\mathbf{P}^{-1}$ est la matrice de l'identité entre (E, B) et (E, B'). Il est clair que l'application $f : (E, B') \to (E, B')$ (de matrice $\mathbf{B}$) peut se "factoriser" de la façon suivante

$$(E, B') \xrightarrow[\mathbf{P}]{\text{id}} (E, B) \xrightarrow[\mathbf{A}]{f} (E, B) \xrightarrow[\mathbf{P}^{-1}]{\text{id}} (E, B') \ .$$

On a donc les relations matricielles

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}, \qquad \mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1} \ .$$

Dans ce chapitre, on s'intéresse à déterminer des conditions suffisantes sur une matrice $\mathbf{A}$ sous lesquelles $\mathbf{A}$ est semblable à une matrice diagonale ou triangulaire. On s'intéressera aussi à la détermination d'algorithmes permettant la construction d'une base B' dans laquelle $\mathbf{A}$ est diagonale.

4.1 Formalisation du problème

Soit $\mathbf{A}$ la matrice de l'endomorphisme f dans une base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. On note $a_{i,j}$ le coefficient générique de $\mathbf{A}$. On cherche une base $B' = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ dans laquelle la matrice de f est diagonale, de matrice

$$\mathbf{A}' = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} .$$

Cela signifie que, pour tout $i \in \{1, \ldots, n\}$, le vecteur $\mathbf{v}_i$ satisfait la relation $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$. Cette relation se réécrit $\mathbf{v}_i \in \ker(f - \lambda_i \mathrm{id})$. La Proposition 88 permet alors d'affirmer que $\lambda_1, \ldots, \lambda_n$ sont solutions de l'équation

$$\det(f - \lambda \mathrm{id}) = 0 .$$

Matriciellement, one peut écrire de manière équivalente que $\lambda_1, \ldots, \lambda_n$ sont solutions de l'équation

$$\begin{vmatrix} a_{1,1} - \lambda & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \ldots & a_{n,n} - \lambda \end{vmatrix} = 0 .$$

C'est une équation polinômiale de degré n (le montrer par récurrence sur n !) à partir de laquelle on va construire B'.

4.2 Vocabulaire

Definition 104. *On appelle valeur propre de f tout élément $\lambda \in K$ tel que $\det(f - \lambda id) = 0$, i.e., tel que $f - \lambda id$ est non-injective.*

On appelle valeur propre de $\boldsymbol{A}$ tout élément $\lambda \in K$ tel que $\det(\boldsymbol{A} - \lambda \boldsymbol{I}_n) = 0$, i.e., tel que $\boldsymbol{A} - \lambda \boldsymbol{I}_n$ est non-injective.

Definition 105. *On appelle vecteur propre de f (associé à la valeur propre λ) tout vecteur de $\ker(f - \lambda id)$, i.e., tout vecteur $\boldsymbol{v}$ tel que $f(\boldsymbol{v}) = \lambda \boldsymbol{v}$.*

On appelle vecteur propre de $\boldsymbol{A}$ (associé à la valeur propre λ) tout vecteur de $\ker(\boldsymbol{A} - \lambda \boldsymbol{I}_n)$, i.e., tout vecteur $\boldsymbol{v}$ tel que $\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$.

Definition 106. *f (resp. $\boldsymbol{A}$) est* diagonalisable *si et seulement il existe une base de E formée de vecteurs propres de f (resp. $\boldsymbol{A}$).*

Definition 107. *Le* spectre *de f (resp. $\boldsymbol{A}$) est l'ensemble de ses valeurs propres.*

Definition 108. *L'espace propre de f (resp. $\boldsymbol{A}$) associé à la valeur propre λ est l'espace $\ker(f - \lambda id)$ (resp. $\ker(\boldsymbol{A} - \lambda \boldsymbol{I}_n)$).*

Definition 109. *Le* polynôme caractéristique *de f (resp. $\boldsymbol{A}$) est défini par*

$$\chi(f) : \lambda \mapsto \det(f - \lambda id), \qquad (resp.\ \chi(\boldsymbol{A}) : \lambda \mapsto \det(\boldsymbol{A} - \lambda \boldsymbol{I}_n)) .$$

C'est un polynôme de degré n en λ.

Proposition 110. *Deux matrices semblables* $\mathbf{A}$ *et* $\mathbf{B}$ *ont même polynôme caractéristique.*

Preuve : Par hypothèse, il existe $\mathbf{P}$ telle que $\mathbf{A} = \mathbf{PBP}^{-1}$, donc pour tout $\lambda \in K$,
$$\mathbf{A} - \lambda \mathbf{I}_n = \mathbf{P}(\mathbf{B} - \lambda \mathbf{I}_n)\mathbf{P}^{-1} \ .$$
Les propriétés énoncées au début de la Section 3.7 permettent alors de déduire
$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = \det(\mathbf{P}) \det(\mathbf{B} - \lambda \mathbf{I}_n) \det(\mathbf{P})^{-1} = \det(\mathbf{B} - \lambda \mathbf{I}_n) \ .$$
Ainsi, on a bien $\chi(\mathbf{A}) = \chi(\mathbf{B})$. □

Proposition 111. *Soient* λ *et* μ *deux valeurs propres distinctes de* f*, alors*
$$\ker(f - \lambda id) \cap \ker(f - \mu id) = \{0\} \ .$$
Preuve : Soit $\mathbf{v} \in \ker(f - \lambda \mathrm{id}) \cap \ker(f - \mu \mathrm{id})$. Alors
$$\lambda \mathbf{v} = f(\mathbf{v}) = \mu \mathbf{v} \ .$$
Donc $(\lambda - \mu)\mathbf{v} = 0$, et comme $\lambda - \mu \neq 0$ par hypothèse, $\mathbf{v} = 0$. □

Proposition 112. *Supposons que* f *a* k *valeurs propres distinctes* $\lambda_1, \ldots, \lambda_k$ *et notons* $E_1, \ldots, E_k$ *les espaces propres associés. Si* $E = E_1 \oplus E_2 \oplus \ldots \oplus E_k$ *alors* f *est diagonalisable.*

Preuve : Soit $\mathbf{u}_{1,i}, \ldots \mathbf{u}_{n_i,i}$ une base de E_i. $\mathbf{u}_{j,i}$ est un vecteur propre de f par définition. De plus, $(\mathbf{u}_{j,i})_{j \in \{1,\ldots,n_i\}, i \in \{1,\ldots,k\}}$, réunion des bases des E_i, est une base de E car les E_i sont en somme directe. E admet donc une base de vecteurs propres de f, donc f est diagonalisable. □

4.3 Exemple

Reprenons l'exemple de la matrice $\mathbf{A}$ donnée en introduction de ce chapitre. Son polynôme caractéristique s'écrit
$$\chi(\mathbf{A}) = \begin{vmatrix} 1-\lambda & 2 & -3 \\ 1 & 4-\lambda & -5 \\ 0 & 2 & -2-\lambda \end{vmatrix} \ .$$
Vérifier que
$$\chi(\mathbf{A}) = (1 - \lambda)\lambda(\lambda - 2) \ .$$
Ce polynôme a 3 racines 0 1 et 2, donc $\mathbf{A}$ a 3 valeurs propres distinctes. Pour déterminer l'espace propre associé à la valeur propre 0, on résout le système
$$\mathbf{Ax} = 0 \ .$$
On vérifie que $(1, 1, 1)$ est vecteur propre et que l'espace propre associé à 0 est de dimension 1.
Exercice : Déterminer de même les espaces propres associés aux valeurs propres 1 et 2.

4.4 Conditions suffisantes de diagonalisabilité

Proposition 113. *Si le polynôme caractéristique de f (resp. de $\boldsymbol{A}$) a n racines distinctes alors f (resp. $\boldsymbol{A}$) est diagonalisable.*

Preuve : Soient $\lambda_1, \ldots, \lambda_n$ les racines distinctes de $\chi(f)$ et soient $\mathbf{u}_1, \ldots, \mathbf{u}_n$ des vecteurs propres associés. On veut montrer que $\mathbf{u}_1, \ldots, \mathbf{u}_n$ est une base de E. Montrons par récurrence que $\mathbf{u}_1, \ldots, \mathbf{u}_k$ est une famille libre. La propriété étant vraie au rang 1, on la suppose vraie à un rang $k \leqslant n-1$ quelconque et on veut la montrer au rang $k+1$. Soient $w_1, \ldots, w_{k+1}$ des éléments de K tels que

$$\sum_{i=1}^{k+1} w_i \mathbf{u}_i = 0 \ . \tag{4.1}$$

En appliquant f aux deux membres de cette égalité, on en déduit

$$0 = \sum_{i=1}^{k+1} w_i f(\mathbf{u}_i) = \sum_{i=1}^{k+1} w_i \lambda_i \mathbf{u}_i \ . \tag{4.2}$$

En multipliant l'égalité (4.1) par λ_{k+1} et en soustrayant l'égalité (4.2), on en déduit

$$\sum_{i=1}^{k} w_i(\lambda_{k+1} - \lambda_i)\mathbf{u}_i = 0 \ .$$

De l'hypothèse de récurrence, on tire alors que

$$\forall i \in \{1, \ldots, k\}, \qquad w_i(\lambda_{k+1} - \lambda_i) = 0 \ .$$

Comme tous les termes $\lambda_{k+1} - \lambda_i \neq 0$ par hypothèse, il vient donc

$$\forall i \in \{1, \ldots, k\}, \qquad w_i = 0 \ .$$

Finalement, en injectant ce résultat dans l'égalité (4.1), on en déduit $w_{k+1}\mathbf{u}_{k+1} = 0$, donc $w_{k+1} = 0$, ce qui conclut la preuve. □

La Proposition (113) suggère la méthode suivante pour diagonaliser l'endomorphisme f. Soit $\mathbf{A}$ la matrice de f dans la base canonique. On calcule le polynôme caractéristique $\chi(\mathbf{A})$ de $\mathbf{A}$ et on le factorise. S'il s'écrit comme un produit $\prod_{i=1}^{n}(\lambda_i - X)$, avec $\lambda_1, \ldots, \lambda_n$ toutes distinctes, alors on résout les systèmes $(\mathbf{A} - \lambda_i \mathbf{I}_N)\mathbf{x} = 0$ pour trouver un vecteur propre $\mathbf{u}_i$ associé à λ_i. La famille $B = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ obtenue ainsi est une base de E dans laquelle f est diagonal. Pour retrouver sans erreur la formule du changement de base, le plus sûr est de "factoriser" f comme nous l'avons vu en introduction de ce chapitre.

4.5 Condition nécessaire et suffisante de diagonalisabilité

Le polynôme caractéristique de f n'a parfois pas n racines simples.

Proposition 114. *Soit λ une valeur propre de f. Si* $\dim(\ker(f - \lambda id)) = r$, *alors λ est racine d'ordre au moins r de $\chi(f)$*

Démonstration. Soit $\mathbf{u}_1, \ldots \mathbf{u}_r$ une base de $\ker(f - \lambda \mathrm{id})$. Soit $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ telle que $\mathbf{u}_1, \ldots \mathbf{u}_n$ soit une base de E. Dans cette base, la matrice $\mathbf{B}$ de f s'écrit par bloc :

$$\mathbf{B} = \begin{bmatrix} \mathbf{D} & \mathbf{C} \\ 0 & \mathbf{C}' \end{bmatrix} ,$$

avec $\mathbf{D} = \lambda \mathbf{I}_r$ et 0 la matrice nulle. Il s'en suit que

$$\chi(f)(x) = \det(\mathbf{B} - x\mathbf{I}_N) = (\lambda - x)^r \det(\mathbf{C}' - x\mathbf{I}_{n-r}) \ .$$

□

Corollaire 115. *Si $\chi(f)$ n'est pas scindé dans K, f n'est pas diagonalisable. S'il existe une racine λ de $\chi(f)$ de multiplicité strictement supérieure à la dimension de* $\ker(f - \lambda id)$*, alors f n'est pas diagonalisable.*

Preuve : Soient $\lambda_1, \ldots, \lambda_k$ les racines distinctes de $\chi(f)$ et $m_1, \ldots, m_k$ leur multiplicité respective. On a d'après la Proposition 114,

$$\sum_{i=1}^{k} \dim(\ker(f - \lambda_i \mathrm{id})) \leqslant \sum_{i=1}^{k} m_i \leqslant n \ .$$

Sous chaque hypothèse du corollaire, une des inégalités est stricte, donc les sous-espaces propres de f n'engendrent pas E. □

Proposition 116. *f est diagonalisable si et seulement si les deux conditions suivantes sont satisfaites.*

1. *$\chi(f)$ est scindé dans $K[X]$.*
2. *Si on note $\lambda_1, \ldots, \lambda_k$ les racines distinctes de $\chi(f)$ et $m_1, \ldots, m_k$ leur multiplicité respective, alors* $\dim(\ker(f - \lambda_i id)) = m_i$*, pour tout $i \in \{1, \ldots, k\}$.*

Exercice : Montrer que la Proposition 113 se déduit de la Proposition 116.

Preuve : Montrons par récurrence que les sous-espaces $E_1, \ldots, E_l$ où $E_i = \dim(\ker(f - \lambda_i \mathrm{id}))$ sont en somme directe. C'est vrai pour $l = 1$, supposons le pour $l < k$ quelconque. Soit $S_l = E_1 \oplus \ldots \oplus E_l$, il suffit de montrer que $S_l \cap E_{l+1} = \{0\}$. Soit $\mathbf{u} \in S_l \cap E_{l+1} = \{0\}$. On a $f(\mathbf{u}) = \lambda_{l+1}\mathbf{u}$. De plus, $\mathbf{u} = \mathbf{v}_1 + \ldots + \mathbf{v}_l$, où chaque $\mathbf{v}_i \in E_i$. Donc $f(\mathbf{u}) = \sum_{i=1}^{l} \lambda_i \mathbf{v}_i = \lambda_{l+1} \sum_{i=1}^{l} \mathbf{v}_i$. Ainsi,

$$\sum_{i=1}^{l} (\lambda_i - \lambda_{l+1})\mathbf{v}_i = 0 \ .$$

D'après l'hypothèse de récurrence, chaque $(\lambda_i - \lambda_{l+1})\mathbf{v}_i = 0$ et comme $\lambda_i - \lambda_{l+1} \neq 0$, chaque $\mathbf{v}_i = 0$.

Comme les sous-espaces E_i sont en somme directe, on a $\dim(E_1 \oplus \ldots \oplus E_k) = \sum_{i=1}^{k} m_i$. Comme de plus $\chi(f)$ est scindé, $\sum_{i=1}^{k} m_i = n$. Donc

$$E = E_1 \oplus \ldots \oplus E_k \ .$$

E est somme directe des sous-espaces propres de f, donc f est diagonalisable. □

La Proposition 116 permet d'étendre la méthode vue après la Proposition (113) pour diagonaliser l'endomorphisme f. Soit $\mathbf{A}$ la matrice de f dans la base canonique. On calcule le polynôme caractéristique $\chi(\mathbf{A})$ de $\mathbf{A}$ et on le factorise. S'il s'écrit comme un produit $\prod_{i=1}^{k}(\lambda_i - X)^{m_i}$, avec $\lambda_1, \ldots, \lambda_k$ toutes distinctes, alors on résout les systèmes $(\mathbf{A} - \lambda_i \mathbf{I}_N)\mathbf{x} = 0$. Si, pour tout i, on peut trouver une famille libre de m_i vecteurs propres $\mathbf{u}_{i,j}$ associé à λ_i, alors $\mathbf{A}$ est diagonalisable et la famille $B = ((\mathbf{u}_{i,j})_{j=1,\ldots,m_i})_{i=1,\ldots,k}$ obtenue ainsi est une base de E dans laquelle f est diagonale.
Exemples : Soit $\mathbf{A}$ et $\mathbf{B}$ les matrices suivantes.

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{bmatrix} \ .$$

1. Montrer que $\chi(\mathbf{A}) = \chi(\mathbf{B}) = -\lambda(\lambda - 1)^2$.
2. Montrer $\dim(\ker(\mathbf{A} - \mathbf{I}_3)) = 2$, $\dim(\ker(\mathbf{B} - \mathbf{I}_3)) = 1$.
3. Les matrices $\mathbf{A}$ et $\mathbf{B}$ sont-elles diagonalisables ?

4.6 Changement de corps de base

Considérons la matrice

$$\mathbf{A} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \ .$$

On a $\chi(\mathbf{A}) = \lambda^2 - 2\lambda\cos(\theta) + 1 = (\lambda - \cos(\theta))^2 + \sin^2(\theta)$. Ce polynôme n'a pas de racines réelles, il n'est pas scindé dans $\mathbb{R}$. Donc, tout endomorphisme f de $\mathbb{R}^2$ représenté par la matrice $\mathbf{A}$ n'est pas diagonalisable.

En revanche, le polynôme $\chi(\mathbf{A})$ a deux racines complexes, $\mathrm{e}^{i\theta}$ et $\mathrm{e}^{-i\theta}$. Donc tout endomorphisme f de $\mathbb{C}^2$ représenté par la matrice $\mathbf{A}$ est diagonalisable et sa matrice dans une base de vecteurs propres s'écrit

$$\mathbf{B} = \begin{bmatrix} \mathrm{e}^{i\theta} & 0 \\ 0 & \mathrm{e}^{-i\theta} \end{bmatrix} \ .$$

4.7 Polynômes d'endomorphisme

Dans toute cette section E est un K-espace vectoriel, f est un endomorphisme de E. Définissons par récurrence

$$f^0 = \mathrm{id}, \qquad f^k = f \circ f^{k-1}, \ \forall k \geqslant 1 \ .$$

Definition 117. *Soit $P(X) = \sum_{k=0}^{r} a_k X^k$ un polynôme à coefficient dans K. On note $P(f) = \sum_{k=0}^{r} a_k f^k$.*

La proposition suivante donne quelques propriétés élémentaires des polynômes d'endomorphisme.

Proposition 118. *$P(f)$ est un endomorphisme de E et, si P_1 et P_2 sont deux polynômes à coefficient dans K et $a \in K$, on a*

$$(P_1 + P_2)(f) = P_1(f) + P_2(f) \ ,$$
$$(P_1 P_2)(f) = P_1(f) \circ P_2(f) \ ,$$
$$(aP_1)(f) = aP_1(f) \ .$$

Exercice : Soit $\mathbf{A}$ une matrice de taille n à coefficients dans K. Donner de même une définition de $P(\mathbf{A})$ et quelques propriétés de base de cette construction.

Definition 119. *On dit que P annule f ou que P est un polynôme annulateur de f si $P(f) = 0$.*

Proposition 120. *Si E est de dimension n.*

1. *Si f est diagonalisable, il existe P scindé à racines simples tel que $P(f) = 0$.*
2. *Si P annule f, toute valeur propre de f est racine de P.*

Preuve : On montre d'abord la propriété 1. Comme f est diagonalisable, il existe une base B de E dont les éléments sont des vecteurs propres de f. Soit $\lambda_1, \ldots, \lambda_k$ les valeurs propres distinctes associées à ces vecteurs de base. Soit $P(X) = \prod_{i=1}^{k}(X - \lambda_i)$. P est un polynôme scindé à racines simples. Soit $\mathbf{u} \in B$, il existe λ_i tel que $f(\mathbf{u}) = \lambda_i \mathbf{u}$, donc

$$P(f)(\mathbf{u}) = \prod_{j=1, j\neq i}^{k} (f - \lambda_j \mathrm{id})(f - \lambda_i \mathrm{id})(\mathbf{u}) = \prod_{j=1, j\neq i}^{k} (f - \lambda_j \mathrm{id})(0) = 0 \ .$$

Donc $P(f)$ est un endomorphisme annulant tous les éléments de la base B de E, c'est donc l'endomorphisme nul.

Montrons maintenant la propriété 2. Soit λ une valeur propre de f et $\mathbf{u}$ un vecteur propre associé. On montre par récurrence que, pour tout k, $f^k(\mathbf{u}) = \lambda^k \mathbf{u}$. On en déduit par linéarité que, pour tout polynôme Q, $Q(f)(\mathbf{u}) = Q(\lambda)\mathbf{u}$. Ainsi, si $P(f) = 0$, on a donc $P(\lambda)\mathbf{u} = 0$ et comme $\mathbf{u}$ est un vecteur propre, il est non nul, donc $P(\lambda) = 0$. □

Proposition 121 (Lemme des noyaux)**.** *Soi P un polynôme annulateur de f et S et T, deux polynômes premiers entre eux tels que $P = ST$. Alors $E = \ker(S(f)) \oplus \ker(T(f))$.*

Preuve : S et T étant premiers entre eux, il existe U et V des polynômes de $K[X]$ tels que $US + VT = 1$. On a donc, pour tout $\mathbf{x} \in E$,

$$\mathbf{x} = U(f) \circ S(f)(\mathbf{x}) + V(f) \circ T(f)(\mathbf{x}) \ .$$

Si $\mathbf{x} \in \ker(S(f)) \cap \ker(T(f))$, on en déduit donc que $\mathbf{x} = 0$. Donc $\ker(S(f))$ et $\ker(T(f))$ sont en somme directe. De plus, comme P annule f,

$$T(f)(U(f) \circ S(f)) = (UP)(f) = 0, \qquad S(f)(V(f) \circ T(f)) = (VP)(f) = 0 \ .$$

Ainsi, $U(f) \circ S(f)(\mathbf{x}) \in \ker(T(f))$ et $V(f) \circ T(f)(\mathbf{x}) \in \ker(S(f))$, donc

$$E = \ker(S(f)) \oplus \ker(T(f)) \ .$$

□

Le corollaire suivant étend le lemme des noyaux à une factorisation plus générale de P.

Corollaire 122. *Si $P = P_1 \ldots P_k$ avec $P_1, \ldots, P_k$ premiers entre eux deux à deux, alors*

$$E = \ker(P_1(f)) \oplus \ldots \ker(P_k(f)) \ .$$

Exercice : Démontrer ce corollaire.

Muni de cette propriété, on peut énoncer une condition suffisante de diagonalisabilité.

Proposition 123. *S'il existe P scindé à racines simples tel que $P(f) = 0$, alors f est diagonalisable.*

Preuve : P étant scindé à racines simples, il s'écrit $P(X) = a(X - \lambda_1) \dots (X - \lambda_k)$. On peut supposer sans perte de généralité que $a = 1$. D'après le Corollaire 122, on a donc

$$E = \ker(f - \lambda_1 \mathrm{id}) \oplus \dots \oplus \ker(f - \lambda_k \mathrm{id}) \ .$$

Ceci implique que E est diagonalisable. □

Ensemble, les propositions 120 et 123 montrent que l'existence d'un polynôme scindé à racine simples annulateur de f est une condition nécessaire et suffisante de diagonalisabilité de f.

4.8 Triangularisation

Si f est triangularisable, sa matrice est semblable à une matrice triangulaire supérieure. Notons $\lambda_1, \dots, \lambda_n$ les coefficients diagonaux de cette matrice. On a donc

$$\chi(f)(X) = (\lambda_1 - X) \dots (\lambda_n - X) \ .$$

$\chi(f)$ est donc scindé dans $K[X]$. On montre dans cette section que, réciproquement, si $\chi(f)$ est scindé, f est triangularisable.

Proposition 124. *Si $\chi(f)$ est scindé, f est triangularisable.*

Remarque : en particulier, si $K = \mathbb{C}$, tout endomorphisme est triangularisable.

Preuve : On raisonne par récurrence sur n. Soit $\mathbf{u}_1$ un vecteur propre associé à λ_1 et soit $\mathbf{u}_2, \dots, \mathbf{u}_n$ des vecteurs tels que $B = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ forme une base de E. Dans cette base, la matrice de f s'écrit par blocs

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & \mathbf{L} \\ 0 & \mathbf{A}' \end{bmatrix} \ .$$

Comme $\chi(\mathbf{A})(X) = (\lambda_1 - X)\chi(\mathbf{A}')(X)$, on en déduit que

$$\chi(\mathbf{A}')(X) = (\lambda_2 - X) \dots (\lambda_n - X) \ .$$

Donc $\mathbf{A}'$ est la matrice d'un endomorphisme d'un espace E' de dimension $n-1$ dont le polynôme caractéristique est scindé sur K. D'après l'hypothèse de récurrence, il existe donc $\mathbf{P}_1$ et $\mathbf{T}$ deux matrices de taille $n-1$, $\mathbf{T}$ étant triangulaire telles que

$$\mathbf{A}' = \mathbf{P}_1 \mathbf{T} \mathbf{P}_1^{-1} \ .$$

Soit alors

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{P}_1 \end{bmatrix} \ .$$

On a

$$\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{P}_1^{-1} \end{bmatrix} ,$$

et

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & \mathbf{L}\mathbf{P}_1 \\ 0 & \mathbf{P}_1^{-1}\mathbf{A}'\mathbf{P}_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{L}\mathbf{P}_1 \\ 0 & \mathbf{T} \end{bmatrix} .$$

$\mathbf{A}$ est semblable à une matrice triangulaire, donc elle est triangularisable, donc f aussi. □

En pratique, pour chercher une base dans laquelle f est triangulaire, on calcule son polynôme caractéristique et s'il est scindé, on sait que f est triangularisable. On détermine alors un vecteur propre associé à chaque valeur propre et on complète cette famille en une base de E. On calcule la matrice de f dans cette base et on recommence la procédure sur la sous-matrice d'ordre inférieur.

4.9 Le théorème de Hamilton-Cayley

Theorem 125 (Hamilton-Cayley)**.** *Si E est de dimension finie, $\chi(f)$ annule f.*

Preuve : On démontre le théorème dans le cas ou $K = \mathbb{C}$. $\chi(f)$ est scindé, $\chi(f)(X) = (\lambda_1 - X)\dots(\lambda_n - X)$. D'après la Proposition 124, f est triangularisable, il existe une base $B = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ de E telle que

$$\forall i \in \{1, \dots, n\}, \qquad f(\mathbf{u}_i) = \lambda_i \mathbf{u}_i + \sum_{j=1}^{i-1} a_{i,j}\mathbf{u}_j .$$

Montrons par récurrence forte sur i que $(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_i \mathrm{id} - f)(\mathbf{u}_j) = 0$, pour tout $j \leqslant i$. On a $f(\mathbf{u}_1) = \lambda_i \mathbf{u}_1$, donc $(\lambda_1 \mathrm{id} - f)(\mathbf{u}_1) = 0$. La propriété est vraie pour $i = 1$. Supposons qu'elle est vraie pour tous les entiers $j \leqslant i$ pour un entier i quelconque inférieur à $n - 1$. On veut montrer que $(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_{i+1} \mathrm{id} - f)(\mathbf{u}_j) = 0$ pour tout $j \leqslant i + 1$. On a par hypothèse de récurrence, pour tout $j \leqslant i$,

$$(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_{i+1} \mathrm{id} - f)(\mathbf{u}_j) = (\lambda_{i+1} \mathrm{id} - f) \circ (\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_i \mathrm{id} - f)(\mathbf{u}_j) = 0$$

De plus

$$(\lambda_{i+1} \mathrm{id} - f)(\mathbf{u}_{i+1}) = -\sum_{j=1}^{i-1} a_{i,j}\mathbf{u}_j ,$$

donc, d'après l'hypothèse de récurrence,

$$(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_{i+1} \mathrm{id} - f)(\mathbf{u}_{i+1}) = -\sum_{j=1}^{i-1} a_{i,j}(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_i \mathrm{id} - f)(\mathbf{u}_j) = 0 .$$

La propriété est donc également vraie au rang $i + 1$. On a donc établi que, pour tout $i \in \{1, \dots, n\}$,

$$(\lambda_1 \mathrm{id} - f) \circ \dots (\lambda_i \mathrm{id} - f)(\mathbf{u}_i) = 0 ,$$

ce qui implique que $\chi(f)(\mathbf{u}_i) = 0$. Ceci étant vrai pour tout i, on a bien $\chi(f) = 0$. □

Le théorème de Hamilton-Cayley (Cayley a énoncé le résultat général dans un papier de 1853, mais en ne le démontrant que pour $n = 2, 3$) peut être appliqué pour calculer l'inverse de f.

Ecrivons $\chi(f) = a_0 + \ldots + a_n X^n$. Le théorème de Hamilton-Cayley assure donc que

$$a_n f^n + \ldots + a_1 f = -a_0 \mathrm{id} \ .$$

Dans cette relation, le terme de gauche s'écrit

$$f \circ [a_n f^{n-1} + \ldots + a_1 \mathrm{id}] \ ,$$

et dans le terme de droite $a_0 = \chi(f)(0) = \det(f)$. Donc f est inversible si et seulement $a_0 \neq 0$ (voir la Proposition 97). Dans ce cas, on a

$$f^{-1} = \frac{1}{-a_0}[a_n f^{n-1} + \ldots + a_1 \mathrm{id}] \ .$$

4.10 Applications de la réduction des endomorphismes

4.10.1 Calculs de puissance

Si $\mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ est une matrice diagonale, alors pour tout entier r, on a

$$\mathbf{D}^r = \mathrm{diag}(\lambda_1^r, \ldots, \lambda_n^r) \ .$$

Si maintenant $\mathbf{A}$ est diagonalisable, on peut calculer $\mathbf{P}$ (inversible) et $\mathbf{D}$ (diagonale) telles que

$$\mathbf{A} = \mathbf{PDP}^{-1} \ .$$

On a alors, pour tout entier r, en raisonnant par récurrence,

$$\mathbf{A}^r = \mathbf{AA}^{r-1} = \mathbf{PDP}^{-1}\mathbf{PD}^{r-1}\mathbf{P}^{-1} = \mathbf{PD}^r\mathbf{P}^{-1} \ .$$

4.10.2 Systèmes différenciels linéaires du premier ordre

On cherche les fonctions $g_1, \ldots, g_n$ telle que

$$\begin{bmatrix} g_1' \\ \vdots \\ g_n' \end{bmatrix} = \mathbf{A} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} \ .$$

Supposons $\mathbf{A}$ diagonalisable, $\mathbf{A} = \mathbf{PDP}^{-1}$, on a alors

$$\mathbf{P}^{-1} \begin{bmatrix} g_1' \\ \vdots \\ g_n' \end{bmatrix} = \mathbf{DP}^{-1} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} \ .$$

On a donc en posant $\mathbf{y} = \mathbf{P}^{-1} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}$,

$$\mathbf{y}' = \mathbf{Dy} \ .$$

On a donc

$$\mathbf{y} = \begin{bmatrix} C_1 e^{\lambda_1 t} \\ \vdots \\ C_n e^{\lambda_n t} \end{bmatrix}, \qquad \text{i.e.} \qquad \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} = \mathbf{P} \begin{bmatrix} C_1 e^{\lambda_1 t} \\ \vdots \\ C_n e^{\lambda_n t} \end{bmatrix} ,$$

où $C_1, \ldots, C_n$ sont des constantes quelconques et $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de $\mathbf{A}$.

4.10.3 Systèmes linéaires d'ordre quelconque

Considérons le système linéaire

$$a_k g^{(k)} + \ldots + a_1 g = 0 \ .$$

Ce système est équivalent au système suivant

$$\mathbf{x}' = \mathbf{A}\mathbf{x}, \qquad \text{ou} \qquad \mathbf{x} = \begin{bmatrix} g \\ g' \\ \vdots \\ g^{k-1} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & \vdots & & & \\ -\frac{a_1}{a_k} & -\frac{a_2}{a_k} & \ldots & & -\frac{a_{k-1}}{a_k} \end{bmatrix}$$

Si $\mathbf{A}$ est diagonalisable, $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ avec $\mathbf{P}$ inversible et $\mathbf{D}$ diagonale. Le système se réécrit $\mathbf{y}' = \mathbf{D}\mathbf{y}$, avec $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$, donc

$$\mathbf{y} = \begin{bmatrix} C_1 e^{\lambda_1 t} \\ \vdots \\ C_n e^{\lambda_n t} \end{bmatrix} ,$$

où $C_1, \ldots, C_n$ sont des constantes quelconques et $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de $\mathbf{A}$. Finalement, g est la première coordonnée de $\mathbf{P}\mathbf{y}$.

Exemple : Appliquer la méthode générale précédente aux équations suivantes :

$$x'' - 3x' + 2x = 0, \qquad x'' - 2x' + x = 0, \qquad x'' + x = 0 \ .$$

4.11 Exercices

Exercice 1

Déterminer le polynôme caractéristique $\chi(\mathbf{A})$, les valeurs propres de $\mathbf{A}$, les sous-espaces propres de $\mathbf{A}$ et écrire les formules de changement de base pour chacune des matrices $\mathbf{A}$ suivantes :

$$\mathbf{A} = \begin{bmatrix} 5 & -3 \\ 6 & -6 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 5 & 8 & -7 \\ -2 & -3 & 2 \\ 2 & 4 & -4 \end{bmatrix} .$$

Exercice 2

Soit u_n et v_n définies par les relations de récurrence

$$u_{n+1} = 5u_n - 3v_n \ ,$$
$$v_{n+1} = 6u_n - 6v_n \ .$$

1. Ecrire le système sous la forme $\mathbf{u}_{n+1} = \mathbf{A}\mathbf{u}_n$, avec $\mathbf{u}_n = \begin{bmatrix} u_n \\ v_n \end{bmatrix}$.
2. Déterminer $\mathbf{P}$ inversible et $\mathbf{D}$ diagonale telles que $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$.
3. Déterminer $\mathbf{v}_n$, où $\mathbf{v}_n = \mathbf{P}^{-1}\mathbf{u}_n$.
4. En déduire, pour tout n, $\mathbf{u}_n$.

Exercice 3

Déterminer les suites u_n et v_n vérifiant

$$u_{n+1} = 5u_n - 3v_n + 2 \times 3^{n+1} + 1 \ ,$$
$$v_{n+1} = 6u_n - 6v_n + 4 \times 3^n + 3 \ .$$

Exercice 4

Résoudre les systèmes différentiels suivants :

$$\begin{cases} u' = 5u - 3v \ , \\ v' = 6u - 6v \ . \end{cases}$$

$$\begin{cases} u' = 5u - 3v + 6e^{3t} + 1 \ , \\ v' = 6u - 6v + 4e^{3t} + 3 \ . \end{cases}$$

Exercice 5

Soit E un K-espace vectoriel de dimension n et soient f et g deux endomorphismes de E.

1. Montrer que les valeurs propres de $f \circ g$ sont les mêmes que celles de $g \circ f$.
2. Supposons que f a n valeurs propres distinctes et que f et g commutent. Montrer que les vecteurs propres de f sont vecteurs propres de g. En déduire qu'il existe une base B dans laquelle f et g sont diagonales. Montrer finalement qu'il existe un polynôme P de degré $n-1$ tel que $g = P(f)$.
3. Supposons que f et g commutent. Montrer que les sous-espaces propres de f sont stables par g.

Exercice 6

Soient E un K-espace vectoriel de dimension n, $L(E)$, l'ev des endomorphismes de E et p un projecteur de rang r de E. On définit $F : L(E) \to L(E)$ par $F(f) = (p \circ f + f \circ p)/2$.

1. Déterminer F^2 et F^3.
2. Déterminer un polynôme de degré 3 annulant F.
3. Montrer que F est diagonalisable.
4. Déterminer les valeurs propres de F, les propres de F et leurs dimensions en fonction $\ker(p)$, $\mathrm{Im}(p)$ et r.

Exercice 7

1. Déterminer, dans l'espace $E = C^\infty(\mathbb{R}, \mathbb{R})$ les valeurs propres et les sous-espaces propres de l'endomorphisme L de E défini par $L(f) = f''$.
2. Déterminer, dans l'espace $E = C([0,1], \mathbb{R})$ les valeurs propres et les sous-espaces propres de l'endomorphisme L de E, où, pour tout $f \in E$, $L(f)$ est défini pour tout $x \in [0,1]$ par $L(f)(x) = \int_0^1 (x \wedge t) f(t) \mathrm{d}t$.
   Indication : calculer la dérivée seconde de $L(f)$.

Exercice 8

Soit E l'espace vectoriel des fonctions continues et bornées sur $\mathbb{R}$. On définit sur E l'endomorphisme L par $L(f)(x) = f(x-1)$.

1. Soit λ une valeur propre de L. Montrer que $\lambda \in \{-1, 1\}$.
2. Déterminer les valeurs propres de L et les espaces propres associés.

Exercice 9

1. Soit $\mathbf{A}$ une matrice carrée d'ordre $n > 1$ telle que $(\mathbf{A} - 2\mathbf{I})^5 = 0$ et $(\mathbf{A} - 2\mathbf{I})^2 \neq 0.$, $\mathbf{A}$ est-elle diagonalisable ?
2. Soit $\mathbf{A}$ une matrice nilpotente non-nulle. $\mathbf{A}$ est-elle diagonalisable ?
3. Soit f un endomorphisme de rang 1 dans un espace E de dimension $n \geqslant 2$. Montre que f est diagonalisable si et seulement si $\mathrm{Tr}(f) \neq 0$.

Exercice 10

Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices à coefficients réels. On suppose qu'il existe $\mathbf{P}$ à coefficients dans $\mathbb{C}$ telle que $\mathbf{A} = \mathbf{PBP}^{-1}$. Montrer qu'il existe $\mathbf{P}$ à coefficients dans $\mathbb{R}$ telle que $\mathbf{A} = \mathbf{PBP}^{-1}$.

Chapitre 5

Orthogonalité

5.1 Orthogonalité dans le plan

Rappelons qu'un base orthonormée du plan est formée de 2 vecteurs orthogonaux de norme 1.

Definition 126. *Dans le plan muni d'une base orthonormée, la norme d'un vecteur $\boldsymbol{u} = (x, y)$ est $\|\boldsymbol{u}\| = \sqrt{x^2 + y^2}$.*

Proposition 127. *Dans le plan muni d'une base orthonormée, deux vecteurs $\boldsymbol{u} = (x, y)$ et $\boldsymbol{v} = (x', y')$ sont orthogonaux si et seulement si $xx' + yy' = 0$.*

Démonstration. D'après le théorème de Pythagore, $\mathbf{u}$ et $\mathbf{v}$ sont orthogonaux si et seulement si $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$, c'est à dire si

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

En développant les carrés, on obtient le résultat. □

Definition 128. *Le produit scalaire des vecteurs $\boldsymbol{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ et $\boldsymbol{v} = \begin{bmatrix} x' \\ y' \end{bmatrix}$ est défini par $\boldsymbol{u}^T \boldsymbol{v} = xx' + yy'$.*

Remarque : $\mathbf{u}$ et $\mathbf{v}$ sont orthogonaux si et seulement si $\mathbf{u}^T \mathbf{v} = 0$. La norme de $\mathbf{u}$ vaut $\|\mathbf{u}\| = \sqrt{\mathbf{u}^T \mathbf{u}}$.

Projection orthogonale : Soit $\mathbf{u}$ un vecteur non nul. La projection orthogonale d'un vecteur $\mathbf{v}$ sur $\mathbf{u}$ est le vecteur de la forme $\lambda \mathbf{u}$ tel que $\mathbf{v} - \lambda \mathbf{u}$ est orthogonal à $\mathbf{u}$. On a donc

$$0 = \mathbf{u}^T (\mathbf{v} - \lambda \mathbf{u}) = \mathbf{u}^T \mathbf{v} - \lambda \mathbf{u}^T \mathbf{u}, \qquad \text{donc} \qquad \lambda = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\|^2} \quad .$$

Cosinus Le cosinus de l'angle θ formé par les vecteurs non nuls $\mathbf{u}$ et $\mathbf{v}$ du plan est défini par

$$\cos(\theta) = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \quad .$$

5.2 Produits scalaires

Le produit scalaire de deux vecteurs du plan vérifie des propriétés à l'origine des définitions suivantes. Dans toute cette section, E est un K espace vectoriel.

Definition 129. *Une fonction $\varphi : E \times E \to K$ est appelée forme bilinéaire si, pour tous $\boldsymbol{u}$, $\boldsymbol{u}'$, $\boldsymbol{v}$, $\boldsymbol{v}'$ de E et tout λ de K*

$$\varphi(\boldsymbol{u} + \lambda\boldsymbol{u}', \boldsymbol{v}) = \varphi(\boldsymbol{u}, \boldsymbol{v}) + \lambda\varphi(\boldsymbol{u}', \boldsymbol{v}) \ ,$$
$$\varphi(\boldsymbol{u}, \boldsymbol{v} + \lambda\boldsymbol{v}') = \varphi(\boldsymbol{u}, \boldsymbol{v}) + \lambda\varphi(\boldsymbol{u}, \boldsymbol{v}') \ .$$

Definition 130. *Une forme bilinéaire φ est dite* symétrique *si, pour tous $\boldsymbol{u}$, $\boldsymbol{v}$ de E,*

$$\varphi(\boldsymbol{u}, \boldsymbol{v}) = \varphi(\boldsymbol{v}, \boldsymbol{u}) \ .$$

Definition 131. *Une forme bilinéaire symétrique φ est dite* positive *si, pour tout $\boldsymbol{u}$ de E,*

$$\varphi(\boldsymbol{u}, \boldsymbol{u}) \geqslant 0 \ .$$

φ est définie positive *si*

$$\varphi(\boldsymbol{u}, \boldsymbol{u}) = 0 \quad \textit{implique} \quad \boldsymbol{u} = 0 \ .$$

Definition 132. *Un produit scalaire sur E est une forme bilinéaire symétrique définie positive. Le produit scalaire de deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ est noté en général $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$.*

Definition 133. *Un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire est appelé*
- *espace Euclidien s'il est de dimension finie,*
- *espace préhilbertien s'il est de dimension infinie.*

Definition 134. *Dans un espace muni d'un produit scalaire, deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ sont dit orthogonaux si $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.*

Exemples :

1. Sur $\mathbb{R}^n$, on peut définir l'application suivante

$$\varphi\left(\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right) = \sum_{i=1}^{n} u_i v_i = \mathbf{u}^T \mathbf{v} \ .$$

   φ est clairement bilinéaire, symétrique, définie et positive, c'est un produit scalaire. On l'appellera le produit scalaire *canonique* de $\mathbb{R}^n$. Il généralise la définition vue sur $\mathbb{R}^2$.
2. Soient $\mathbf{A}$, $\mathbf{B}$ deux matrices de taille n de coefficients génériques respectifs $a_{i,j}$ et $b_{i,j}$. On définit $\varphi(\mathbf{A}, \mathbf{B}) = \sum_{1 \leqslant i,j \leqslant n} a_{i,j} b_{i,j}$. Vérifier que φ définit un produit scalaire $\langle \cdot, \cdot \rangle$ et que $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{Tr}(\mathbf{A}\mathbf{B}^T)$.
3. Soit I un intervalle de $\mathbb{R}$ et $E = C(I, \mathbb{R})$. Montrer que $\varphi(f, g) = \int_I f(t)g(t)\mathrm{d}t$ est un produit scalaire sur E. Si $I = [0, 2\pi]$, montrer que les fonctions $(\cos(kt), \sin(jt), k, j \in \mathbb{N})$ sont toutes orthogonales.
   Schmidt (1905) est le premier à parler d'orthogonalité pour des fonctions.

5.3 Expressions du produit scalaire

Soit E un espace Euclidien de dimension n et $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E. Soient $\mathbf{u}$, $\mathbf{v}$ deux vecteurs de E. Ces vecteurs se décomposent sur B, on a $\mathbf{u} = \sum_{i=1}^n u_i \mathbf{e}_i$, $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$. Par bilinéarité du produit scalaire, on a

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^n u_i \mathbf{e}_i, \sum_{j=1}^n v_j \mathbf{e}_j \right\rangle = \sum_{1 \leqslant i,j \leqslant n} u_i v_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad . \tag{5.1}$$

En particulier, on a donc

$$\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{1 \leqslant i,j \leqslant n} u_i u_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad .$$

Par symétrie du produit scalaire, on a $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle \mathbf{e}_j, \mathbf{e}_i \rangle$, donc

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n u_i v_i \langle \mathbf{e}_i, \mathbf{e}_i \rangle + \sum_{1 \leqslant i < j \leqslant n} (u_i v_j + u_j v_i) \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad ,$$

$$\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{i=1}^n u_i^2 \langle \mathbf{e}_i, \mathbf{e}_i \rangle + 2 \sum_{1 \leqslant i < j \leqslant n} u_i u_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad .$$

En posant $a_{i,j} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$, on a ainsi

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n a_{i,i} u_i v_i + \sum_{1 \leqslant i < j \leqslant n} a_{i,j} (u_i v_j + u_j v_i) \quad ,$$

$$\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{i=1}^n a_{i,i} u_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{i,j} u_i u_j \quad .$$

5.3.1 Représentation matricielle dans un espace Euclidien

Supposons E Euclidien, de dimension n, muni d'une base $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$. Soit $\mathbf{u} = \sum_{i=1}^n u_i \mathbf{e}_i$, $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$, deux vecteurs de E et soient $\mathbf{u}_c$, $\mathbf{v}_c$ les vecteurs (colonnes) de K^n associés à $\mathbf{u}$ et $\mathbf{v}$ définis respectivement par

$$\mathbf{u}_c = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} , \qquad \mathbf{v}_c = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \quad .$$

Introduisons la matrice $\mathbf{\Phi}$ de coefficient générique $a_{i,j} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$. Cette matrice sera appelée dans la suite matrice associée au produit scalaire $\langle \cdot, \cdot \rangle$.

Proposition 135. *La matrice $\mathbf{\Phi}$, associée au produit scalaire $\langle \cdot, \cdot \rangle$ est symétrique et inversible. De plus, on a, avec les notations précédentes*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}_c^T \mathbf{\Phi} \boldsymbol{v}_c \quad .$$

Preuve : $\mathbf{\Phi}$ est clairement symétrique car le produit scalaire l'est. On vérifie que

$$\mathbf{u}_c^T \mathbf{\Phi} \mathbf{v}_c = \sum_{1 \leqslant i,j \leqslant n} u_i v_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad .$$

D'après (5.1), on a donc
$$\mathbf{u}_c^T \mathbf{\Phi} \mathbf{v}_c = \langle \mathbf{u}, \mathbf{v} \rangle \quad .$$
Si maintenant $\mathbf{u}_c$ est un vecteur colonne de K^n tel que $\mathbf{\Phi}\mathbf{u}_c = 0$, on a donc $\mathbf{u}_c^T \mathbf{\Phi} \mathbf{u}_c = 0$, donc $\langle \mathbf{u}, \mathbf{u} \rangle = 0$. Comme $\langle \cdot, \cdot \rangle$ est un produit scalaire, ceci implique $\mathbf{u} = 0$, et donc $\mathbf{u}_c = 0$. Ainsi, le noyau de $\mathbf{\Phi}$ est réduit à 0, donc $\mathbf{\Phi}$ est inversible. □

Exemple La matrice associée au produit scalaire canonique sur $\mathbb{R}^n$ est $\mathbf{I}_n$.

5.3.2 Changement de base

Soient $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $B' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ deux bases de l'espace Euclidien E. Soient $\mathbf{u}$ et $\mathbf{v}$ deux vecteurs de E, soient $\mathbf{u}_c$, $\mathbf{v}_c$ les vecteurs colonnes associés dans la base B et soient $\mathbf{u}'_c$, $\mathbf{v}'_c$ les vecteurs colonnes associés dans la base B'. Soient $\mathbf{\Phi}$ (resp. $\mathbf{\Phi}'$) la matrice associée au produit scalaire de E dans la base B (resp. B'). On a donc
$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}_c^T \mathbf{\Phi} \mathbf{v}_c = (\mathbf{u}'_c)^T \mathbf{\Phi}' \mathbf{v}'_c \quad .$$
Soit $\mathbf{P}$ la matrice de l'identité de E, de la base B' à la base B. On a donc $\mathbf{u}_c = \mathbf{P}\mathbf{u}'_c$ et $\mathbf{v}_c = \mathbf{P}\mathbf{v}'_c$, donc
$$(\mathbf{u}'_c)^T \mathbf{\Phi}' \mathbf{v}'_c = \mathbf{u}_c^T \mathbf{\Phi} \mathbf{v}_c = (\mathbf{u}'_c)^T \mathbf{P}^T \mathbf{\Phi} \mathbf{P} \mathbf{v}'_c \quad .$$
Cette relation étant valable pour tous les vecteurs colonnes $\mathbf{u}'_c$ et $\mathbf{v}'_c$, on a donc
$$\mathbf{\Phi}' = \mathbf{P}^T \mathbf{\Phi} \mathbf{P} \quad .$$

5.3.3 Reconnaître un produit scalaire

On se donne une application définie par
$$\varphi(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} a_{i,i} u_i v_i + \sum_{1 \leqslant i < j \leqslant n} a_{i,j} (u_i v_j + u_j v_i) \quad .$$
On veut savoir si φ définit un produit scalaire. φ est clairement bilinéaire et symétrique, la difficulté est de savoir si elle est définie positive. La méthode suivante est due à Gauss.

- S'il existe un i tel que $a_{i,i} \leqslant 0$, alors $\varphi(\mathbf{e}_i, \mathbf{e}_i) \leqslant 0$, donc φ n'est pas définie positive. On suppose donc que tous les $a_{i,i} > 0$.
- On écrit
$$f_1(\mathbf{u}) = u_1 + \sum_{j=2}^{n} \frac{a_{1,j}}{a_{1,1}} u_j \quad .$$
On vérifie simplement que
$$\varphi_1(\mathbf{u}, \mathbf{u}) = \varphi(\mathbf{u}, \mathbf{u}) - a_{1,1} f_1^2(\mathbf{u}) \quad ,$$
est une fonction ne dépendant pas de u_1. On itère alors le processus avec φ_1 et u_2, de manière à construire une fonction $\varphi_2(\mathbf{u}, \mathbf{u})$ ne dépendant

ni de u_1, ni de u_2. On construit ainsi itérativement une suite de forme linéaire $f_i(\mathbf{u}) = u_i + \ldots$ telle que

$$\varphi(\mathbf{u}, \mathbf{u}) = \sum_{i=1}^{p} \alpha_i f_i(\mathbf{u})^2 \ .$$

— Si $p = n$ et que, pour tout i, $\alpha_i > 0$, alors $\varphi(\mathbf{u}, \mathbf{u})$ est équivalent au système triangulaire supérieur

$$\begin{cases} f_1(\mathbf{u}) = 0 \ , \\ \vdots \\ f_n(\mathbf{u}) = 0 \ . \end{cases}$$

Ce système n'admet que 0 comme solution, donc φ est une forme bilinéaire définie positive.

— Si $p < n$. Le système précédent n'est pas de rang plein, il admet donc des solutions non nulles, donc φ n'est pas définie positive.

— Si $p = n$ et qu'il existe un $\alpha_i < 0$. Il existe $\mathbf{u}$ tel que $f_j(\mathbf{u}) = 0$ quel que soit $j \neq i$ et $f_i(\mathbf{u}) \neq 0$. Un tel $\mathbf{u}$ vérifie $\varphi(\mathbf{u}, \mathbf{u}) < 0$, donc φ n'est pas définie positive.

Exemples Appliquer la méthode de Gauss pour vérifier si les formes suivantes sont définies positives.

$$\varphi(\mathbf{u}, \mathbf{u}) = u_1^2 + 3u_2^2 + 5u_3^2 \ ,$$
$$\varphi(\mathbf{u}, \mathbf{u}) = u_1^2 - 3u_2^2 + 5u_3^2 \ ,$$
$$\varphi(\mathbf{u}, \mathbf{u}) = u_1^2 + 3u_2^2 \ ,$$
$$\varphi(\mathbf{u}, \mathbf{u}) = 3u_1u_2 + u_1u_3 \ ,$$
$$\varphi(\mathbf{u}, \mathbf{u}) = 2u_1^2 + u_2^2 + 2u_3^2 - 4u_1u_2 - 2u_2u_3 + 4u_1u_3 \ .$$

5.3.4 Norme et angle

On note E un espace vectoriel sur $\mathbb{R}$ et $\langle \cdot, \cdot \rangle$ un produit scalaire sur E.

Definition 136. *La norme d'un vecteur $\boldsymbol{u} \in E$ est $\|\boldsymbol{u}\| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$.*

Proposition 137 (Inégalité de Cauchy-Schwarz)**.** *Pour tous $\boldsymbol{u}$, $\boldsymbol{v}$ de E,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leqslant \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 \ .$$

On a égalité si et seulement si $\boldsymbol{u}$ et $\boldsymbol{v}$ sont colinéaires.

Preuve. Si $\mathbf{v} = 0$, le résultat est évident. Supposons donc $\mathbf{v} \neq 0$. On note, pour tout $x \in \mathbb{R}$,

$$P(\mathbf{x}) = \|\mathbf{u} + x\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2x \langle \mathbf{u}, \mathbf{v} \rangle + x^2 \|\mathbf{v}\|^2 \ .$$

Ce polynôme du second degré est à valeurs positives ou nulles, il a donc au plus une racine, donc son discriminant est négatif ou nul. Ce discriminant vaut

$$4(\langle \mathbf{u}, \mathbf{v} \rangle^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2) \ .$$

S'ils sont colinéaires, on a clairement égalité. Si $\mathbf{u}$ et $\mathbf{v}$ ne sont pas colinéaires, $P(x)$ ne s'annule pas, donc son discriminant est strictement négatif. □

Remarque : Bunyakovski a travaillé avec Cauchy sur cette inégalité et a publié, 25 ans avec Schwarz les résultats sur la forme fonctionnelle de l'inégalité. L'inégalité de Cauchy-Schwarz ne porte donc pas le nom de tous ces auteurs.
Exercice : démontrer les relations suivantes

$$\sum_{i=1}^{n} x_i y_i \leqslant \sqrt{\Big(\sum_{i=1}^{n} x_i^2\Big)\Big(\sum_{i=1}^{n} y_i^2\Big)} \ ,$$

$$\int_I f(t)g(t)\mathrm{d}t \leqslant \sqrt{\Big(\int_I f^2(t)\mathrm{d}t\Big)\Big(\int_I g^2(t)\mathrm{d}t\Big)} \ .$$

Proposition 138. *Pour tout $\boldsymbol{u} \in E$, $\|\boldsymbol{u}\| \geqslant 0$.*
$\|\boldsymbol{u}\| = 0$ si et seulement si $\boldsymbol{u} = 0$.
Pour tout $\boldsymbol{u} \in E$ et tout $x \in \mathbb{R}$, $\|x\boldsymbol{u}\| = |x|\|\boldsymbol{u}\|$.
Pour tous $\boldsymbol{u}$ et $\boldsymbol{v}$ de E, $\|\boldsymbol{u}+\boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. (Inégalité triangulaire.)
$\|\boldsymbol{u}+\boldsymbol{v}\| = \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ si et seulement si $\boldsymbol{u}$ et $\boldsymbol{v}$ sont colinéaires.
$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = (1/2)(\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}\|^2 - \|\boldsymbol{v}\|^2)$.

Preuve : La première propriété vient de la positivité du produit scalaire. La seconde propriété vient du fait que le produit scalaire est une forme définie positive. La troisième propriété est une conséquence de la définition de la norme. Pour la quatrième propriété, on écrit d'une part

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \|\mathbf{v}\|^2 \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \ . \end{aligned} \tag{5.2}$$

D'autre part, on a

$$(\|\mathbf{u}\| + \|\mathbf{v}\|)^2 = \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \ .$$

On a donc

$$\|\mathbf{u}+\mathbf{v}\| \leqslant \|\mathbf{u}\| + \|\mathbf{v}\| \qquad \text{ssi} \qquad \langle \mathbf{u}, \mathbf{v}\rangle \leqslant \|\mathbf{u}\|\|\mathbf{v}\| \ .$$

Par l'inégalité de Cauchy-Schwarz, $\langle \mathbf{u}, \mathbf{v}\rangle \leqslant \|\mathbf{u}\|\|\mathbf{v}\|$, donc $\|\mathbf{u}+\mathbf{v}\| \leqslant \|\mathbf{u}\| + \|\mathbf{v}\|$. La cinquième propriété vie du cas d'égalité dans l'inégalité de Cauchy-Schwarz. La sixième propriété est une réécriture de l'équation (5.2). □

Remarque : L'équation (5.2) donne le théorème de Pythagore. $\mathbf{u}$ et $\mathbf{v}$ sont orthogonaux (i.e. $\langle \mathbf{u}, \mathbf{v}\rangle = 0$) si et seulement si $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$. Au passage, rappelons que les mathématiciens Babyloniens connaissaient déjà ce théorème 1200 ans avant Pythagore, qui n'a quant à lui probablement jamais fait de mathématiques.

Definition 139. *Dans un espace E muni d'un produit scalaire, on appelle angle (non-orienté) entre deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ le réel $\theta \in [0, \pi]$ tel que*

$$\cos(\theta) = \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} \ .$$

Cette expression du cosinus apparaît dans les travaux de Jordan en 1873 (dans le cas du produit scalaire canonique sur $\mathbb{R}^n$).

5.4 Bases orthogonales, orthonormées

Dans cette section, E est un espace vectoriel muni d'un produit scalaire.

Definition 140. *Une famille $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ est dite orthogonale si, pour tous $i \neq j$ dans $\{1, \ldots, k\}$, $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0$.*

Proposition 141. *Une famille orthogonale $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ sans vecteur nul est une famille libre.*

Preuve : Soient $a_1, \ldots, a_k$ dans K tels que $\sum_{i=1}^{k} a_i \mathbf{e}_i = 0$. Pour tout $j \in \{1, \ldots, k\}$, on a donc

$$0 = \left\langle \mathbf{e}_j, \sum_{i=1}^{k} a_i \mathbf{e}_i \right\rangle = a_j \langle \mathbf{e}_j, \mathbf{e}_j \rangle \ .$$

Comme $\mathbf{e}_j \neq 0$, $\langle \mathbf{e}_j, \mathbf{e}_j \rangle \neq 0$, donc $a_j = 0$. Comme c'est vrai pour tout j, la famille $\mathbf{e}_1, \ldots, \mathbf{e}_k$ est libre. □

On suppose maintenant de plus que E est de dimension finie n.

Definition 142. *Une base $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ de E est dite orthogonale si elle forme une famille orthogonale. Elle est dite orthonormée si elle est orthogonale et vérifie $\|\boldsymbol{e}_i\| = 1$ pour tout $i \in \{1, \ldots, n\}$.*

Proposition 143. *Soit $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base orthonormée de E. Soient $\boldsymbol{u}$, $\boldsymbol{v}$ deux vecteurs de E et soient $\boldsymbol{u}_c$, $\boldsymbol{v}_c$ les vecteurs colonnes associés à leur décomposition dans la base B. On a*

1. *$u_i = \langle \boldsymbol{u}, \boldsymbol{e}_i \rangle$ pour tout $i \in \{1, \ldots, n\}$.*
2. *La matrice $\boldsymbol{\Phi}$ associée au produit scalaire est $\boldsymbol{I}_n$.*
3. *$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} u_i v_i = \boldsymbol{u}_c^T \boldsymbol{v}_c$.*

Preuve :
1. $\langle \mathbf{u}, \mathbf{e}_i \rangle = \sum_{j=1}^{n} u_j \langle \mathbf{e}_j, \mathbf{e}_i \rangle = u_i$.
2. Le coefficient générique de $\boldsymbol{\Phi}$ est $\langle \mathbf{e}_i, \mathbf{e}_j \rangle$. Il vaut 1 si $i = j$ et 0 sinon.
3. C'est une conséquence de la Proposition 135 et du point 2.

□

Procédé d'orthonormalisation de Gram-Schmidt

Proposition 144. *Soit $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E. Il existe une base $B' = (\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n)$ de E, orthonormée, telle que, pour tout $i \in \{1, \ldots, n\}$, les espaces engendrés par les vecteurs $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i)$ et $(\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_i)$ sont égaux.*

Preuve : On construit les vecteurs $\mathbf{e}'_i$ par récurrence. Le vecteur $\mathbf{e}_1$ est non-nul, donc on peut poser $\mathbf{e}'_1 = (1/\|\mathbf{e}_1\|)\mathbf{e}_1$. Le vecteur $\mathbf{e}'_1$ est de norme 1 et engendre le même espace vectoriel que $\mathbf{e}_1$. Supposons maintenant qu'on ait construit une famille orthonormée $(\mathbf{e}'_1, \ldots, \mathbf{e}'_i)$ engendrant le même espace vectoriel que $(\mathbf{e}_1, \ldots, \mathbf{e}_i)$. Le vecteur $\mathbf{e}_{i+1}$ n'appartient pas à cet espace. On pose

$$\mathbf{e}''_{i+1} = \mathbf{e}_{i+1} - \sum_{j=1}^{i} \left\langle \mathbf{e}_{i+1}, \mathbf{e}'_j \right\rangle \mathbf{e}'_j \ .$$

Le vecteur $\sum_{j=1}^{i} \left\langle \mathbf{e}_{i+1}, \mathbf{e}'_j \right\rangle \mathbf{e}'_j$ appartient à l'espace vectoriel engendré par $(\mathbf{e}_1, \ldots, \mathbf{e}_i)$, donc $\mathbf{e}''_{i+1} \neq 0$, on peut donc définir

$$\mathbf{e}'_{i+1} = \frac{1}{\|\mathbf{e}''_{i+1}\|} \mathbf{e}''_{i+1} \ .$$

Il est clair que $(\mathbf{e}'_1, \ldots, \mathbf{e}'_{i+1})$ et $(\mathbf{e}'_1, \ldots, \mathbf{e}'_i, \mathbf{e}''_{i+1})$ engendrent le même espace. De plus, on vérifie sans mal que cet espace est celui engendré par $(\mathbf{e}'_1, \ldots, \mathbf{e}'_i, \mathbf{e}_{i+1})$, donc celui engendré par $(\mathbf{e}_1, \ldots, \mathbf{e}_{i+1})$ d'après l'hypothèse de récurrence. Il suffit donc de montrer que $(\mathbf{e}'_1, \ldots, \mathbf{e}'_{i+1})$ est une famille orthonormée. Tous les vecteurs qui la composent sont de norme 1 d'après l'hypothèse de récurrence et la construction de $\mathbf{e}'_{i+1}$. Il suffit donc de montrer que c'est une famille orthogonale et pour cela, d'après l'hypothèse de récurrence, il suffit de montrer que, pour tout $j \in \{1, \ldots, i\}$, $\left\langle \mathbf{e}'_{i+1}, \mathbf{e}'_j \right\rangle = 0$. Comme $\mathbf{e}'_{i+1}$ et $\mathbf{e}''_{i+1}$ sont colinéaires, il suffit en fait de montrer que, pour tout $k \in \{1, \ldots, i\}$,

$$\left\langle \mathbf{e}''_{i+1}, \mathbf{e}'_k \right\rangle = 0 \ .$$

Par hypothèse de récurrence,

$$\left\langle \sum_{j=1}^{i} \left\langle \mathbf{e}_{i+1}, \mathbf{e}'_j \right\rangle \mathbf{e}'_j, \mathbf{e}'_k \right\rangle = \left\langle \mathbf{e}_{i+1}, \mathbf{e}'_k \right\rangle \ ,$$

donc

$$\left\langle \mathbf{e}''_{i+1}, \mathbf{e}'_k \right\rangle = 0 \ .$$

□

On déduit de ce procédé que toute famille orthogonale sans vecteur nul de E peut être complétée en une base orthogonale de E.

5.5 Orthogonalité de sous-espaces.

Dans toute cette section, E désigne un espace Euclidien.

Definition 145. *Soit F un sous-ensemble de E. On note*

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

L'ensemble $F^{\perp}$ est appelé l'orthogonal de F.

Exercice : Montrer que $F^{\perp}$ est un sous-espace vectoriel de E (même si F n'est pas un espace vectoriel).
Exercice : Si F est l'espace vectoriel engendré par $\mathbf{u}_1, \ldots, \mathbf{u}_k$, montrer que $\mathbf{x} \in F^{\perp}$ si et seulement si il est orthogonal à chacun des vecteurs $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

Proposition 146. *Soit F un sous-espace vectoriel de E.*

1. $E = F \oplus F^{\perp}$.
2. $\dim(F) + \dim(F^{\perp}) = \dim(E)$.
3. *Si G est un sous-ensemble de F, alors $F^{\perp} \subset G^{\perp}$.*
4. $(F^{\perp})^{\perp} = F$.

Preuve : Soit $\mathbf{e}_1, \ldots, \mathbf{e}_d$ une base orthonormée de F. Complétons la en une base orthonormée $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de E. Soit $\mathbf{x} \in F^\perp$, on décompose $\mathbf{x}$ sur cette base :

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \ .$$

Comme $\mathbf{x} \in F^\perp$, pour tout $i \in \{1, \ldots, d\}$,

$$0 = \langle \mathbf{x}, \mathbf{e}_i \rangle = x_i \ .$$

Donc $\mathbf{x}$ appartient à l'espace engendré V par $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$. Autrement dit $F^\perp \subset V$. Réciproquement, tout vecteur de V est orthogonal à $\mathbf{e}_i$, pour tout $i \in \{1, \ldots, d\}$. Ainsi $V \subset F^\perp$ et donc $V = F^\perp$. On déduit de cette égalité les propriétés 1 et 2.

Soient F et G deux sous-ensembles de E, avec $G \subset F$. Si $\mathbf{x} \in F^\perp$, il est orthogonal à tous les vecteurs de F, donc de G, donc $\mathbf{x} \in G^\perp$.

Si $\mathbf{x} \in F$, il est orthogonal à tous les vecteurs de $F^\perp$, donc il appartient à $(F^\perp)^\perp$. De plus, par la propriété 2, $\dim((F^\perp)^\perp) = \dim(F)$, donc $(F^\perp)^\perp = F$. □

5.5.1 Formes linéaires

Pour tout vecteur $\mathbf{u} \in E$, on note $f_\mathbf{u} : \mathbf{x} \in E \mapsto f_\mathbf{u}(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \in K$. L'application $f_\mathbf{u}$ est une forme linéaire, c'est à dire une application linéaire de E dans K. L'ensemble des formes linéaires est un espace vectoriel appelé espace dual de E. On le note E^*.

5.5.2 Hyperplan orthogonal

L'orthogonal d'un vecteur $\mathbf{u}$ non nul de E est le noyau de $f_\mathbf{u}$. L'orthogonal d'un sous-ensemble F de E est l'intersection $\cap_{\mathbf{u} \in F} \ker(f_\mathbf{u})$.

5.5.3 Isomorphisme E et E^*

L'application $\Phi : \mathbf{u} \in E \mapsto f_\mathbf{u} \in E^*$ est un isomorphisme d'espaces vectoriels. On vérifie facilement que cette application est linéaire. Soit $\mathbf{u}$ telle que $f_\mathbf{u} = 0$. On a alors $0 = f_\mathbf{u}(\mathbf{u}) = \|\mathbf{u}\|^2$, donc $\mathbf{u} = 0$. L'application Φ est donc injective. Comme E et E^* ont la même dimension Φ est donc un isomorphisme. Si E est muni d'une base B, on peut munir E^* de la base duale $B^* = (\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*)$. L'application Φ vérifie, pour tout i et j, $\Phi(\mathbf{e}_i)(\mathbf{e}_j) = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$, donc $\Phi(\mathbf{e}_i) = \sum_{j=1}^n \langle \mathbf{e}_i, \mathbf{e}_j \rangle \, \mathbf{e}_j^*$. Il s'en suit que la matrice de Φ dans les base B et B^* est $\boldsymbol{\Phi}$ introduite à la Section 5.3.1.

5.6 Projection orthogonale

Definition 147. *Soit p un endomorphisme de E. p est appelée projection si $p^2 = p$.*

Proposition 148. *Si p est une projection, $E = \ker(p) \oplus Im(p)$ et $\forall \boldsymbol{x} \in Im(p)$, $p(\boldsymbol{x}) = \boldsymbol{x}$.*

Preuve : Soit $\mathbf{x} \in E$, $\mathbf{x} = (\mathbf{x} - p(\mathbf{x})) + p(\mathbf{x})$. On a $p(\mathbf{x}) \in \mathrm{Im}(p)$ et $p(\mathbf{x} - p(\mathbf{x})) = p(\mathbf{x}) - p^2(\mathbf{x}) = 0$ donc $E = \ker(p) + \mathrm{Im}(p)$. De plus, cette somme est directe par le théorème du rang. □

Proposition 149. *Soit U et V deux sous-espaces vectoriel de E tels que $E = U \oplus V$. Tout élément $\boldsymbol{x} \in E$ se décompose donc de manière unique $\boldsymbol{x} = \boldsymbol{x}_U + \boldsymbol{x}_V$, avec $\boldsymbol{x}_U \in U$, $\boldsymbol{x}_V \in V$. L'application $\pi : \boldsymbol{x} \mapsto \boldsymbol{x}_U$ est une projection appelée projection sur U parallèlement à V, son image est U et son noyau est V.*

Preuve : π est une application linéaire vérifiant clairement $\pi^2 = \pi$, c'est donc une projection. On vérifie sans peine que son image est U et son noyau est V. □

Definition 150. *Soit F un sous-espace vectoriel de E. La projection orthogonale p_F sur F est la projection sur F parallèlement à $F^\perp$.*

On a, pour tout $\mathbf{x} \in E$, $\mathbf{x} - p_F(\mathbf{x}) \in F^\perp$. On vérifie sans peine que $\mathrm{id} - p_F$ est la projection orthogonale sur $F^\perp$. En fait, on a plus généralement, si p est la projection sur U parallèlement à V, $\mathrm{id} - p$ est la projection sur V parallèlement à U.

Proposition 151. *L'application p_F est contractante, c'est à dire que, pour tout $\boldsymbol{x} \in E$, $\|p_F(\boldsymbol{x})\| \leqslant \|\boldsymbol{x}\|$. On a égalité si et seulement si $\boldsymbol{x} \in F$.*

Preuve : Comme $p_F(\mathbf{x})$ et $\mathbf{x} - p_F(\mathbf{x})$ sont orthogonaux, le théorème de Pythagore assure que

$$\|\mathbf{x}\|^2 = \|p_F(\mathbf{x})\|^2 + \|\mathbf{x} - p_F(\mathbf{x})\|^2 \ .$$

La proposition s'en déduit. □

Exemple. Si F est une droite, engendrée par un vecteur $\mathbf{u}$. Montrer que

$$\forall \mathbf{x} \in E, \qquad p_F(\mathbf{x}) = \frac{\langle \mathbf{u}, \mathbf{x}\rangle}{\|\mathbf{u}\|^2}\mathbf{u}, \qquad p_{F^\perp}(\mathbf{x}) = \mathbf{x} - \frac{\langle \mathbf{u}, \mathbf{x}\rangle}{\|\mathbf{u}\|^2}\mathbf{u} \ .$$

Matrice Soit $\mathbf{e}_1, \ldots, \mathbf{e}_d$ une base de F et $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ une base de $F^\perp$. Dans la base $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E, la matrice de p_F s'écrit par blocs

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{I}_d & 0 \\ 0 & 0 \end{bmatrix} \ .$$

Definition 152. *La distance de $\boldsymbol{x} \in E$ à un sous-ensemble V de E est définie par*

$$d(\boldsymbol{x}, V) = \inf_{\boldsymbol{u} \in V} \|\boldsymbol{x} - \boldsymbol{u}\| \ .$$

Proposition 153. *Soit F un sous-espace vectoriel de E.*

$$d(\boldsymbol{x}, F) = \|\boldsymbol{x} - p_F(\boldsymbol{x})\| \ .$$

Démonstration. Soit $\mathbf{u} \in F$. On a $p_F(\mathbf{x}) - \mathbf{u}$ et $\mathbf{x} - p_F(\mathbf{x})$ sont orthogonaux, donc, d'après le théorème de Pythagore

$$\|\mathbf{x} - \mathbf{u}\|^2 = \|\mathbf{x} - p_F(\mathbf{x})\|^2 + \|p_F(\mathbf{x}) - \mathbf{u}\|^2 \ .$$

Il s'en suit que

$$\|\mathbf{x} - \mathbf{u}\| \geqslant \|\mathbf{x} - p_F(\mathbf{x})\| \ .$$

□

Exercice. Calculer $p_F(\mathbf{x}_2)$, où E est l'ensemble des fonctions continues de $[0,1]$ à valeurs dans $\mathbb{R}$, $\mathbf{x}_i : t \mapsto t^i$ et F est l'espace vectoriel engendré par $\mathbf{x}_0$ et $\mathbf{x}_1$.

5.7 Transformations orthogonales

Dans cette section, E et E' désignent deux espaces Euclidiens de même dimension et $f : E \to E'$ est une application linéaire.

Definition 154. *On dit que f est une transformation orthogonale si*

$$\forall \boldsymbol{u}, \boldsymbol{v} \in E, \qquad \langle f(\boldsymbol{u}), f(\boldsymbol{v})\rangle_{E'} = \langle \boldsymbol{u}, \boldsymbol{v}\rangle_E \quad .$$

Proposition 155. *f est orthogonale si et seulement si f est une isométrie, c'est à dire ssi*

$$\forall \boldsymbol{u} \in E, \qquad \|f(\boldsymbol{u})\|_{E'} = \|\boldsymbol{u}\|_E \quad .$$

Preuve : Si f est orthogonale,

$$\forall \mathbf{u} \in E, \qquad \|f(\mathbf{u})\|_{E'}^2 = \langle f(\mathbf{u}), f(\mathbf{u})\rangle_{E'} = \langle \mathbf{u}, \mathbf{u}\rangle_E = \|\mathbf{u}\|_E^2 \quad .$$

Réciproquement, si f est une isométrie, par la sixième propriété de la Proposition 138,

$$\begin{aligned}
\langle f(\mathbf{u}), f(\mathbf{v})\rangle_{E'} &= \frac{1}{2}\left(\|f(\mathbf{u})+f(\mathbf{v})\|_{E'}^2 - \|f(\mathbf{u})\|_{E'}^2 - \|f(\mathbf{v})\|_{E'}^2\right)\\
&= \frac{1}{2}\left(\|f(\mathbf{u}+\mathbf{v})\|_{E'}^2 - \|f(\mathbf{u})\|_{E'}^2 - \|f(\mathbf{v})\|_{E'}^2\right)\\
&= \frac{1}{2}\left(\|\mathbf{u}+\mathbf{v}\|_{E}^2 - \|\mathbf{u}\|_{E}^2 - \|\mathbf{v}\|_{E}^2\right) = \langle \mathbf{u}, \mathbf{v}\rangle_E \quad .
\end{aligned}$$

□

Proposition 156. *L'ensemble des isométries $f : E \to E$ muni de la composition est un groupe qu'on appelle groupe orthogonal de E et qu'on note $O(E)$.*

Preuve : Si f et g sont des isométries et si $\mathbf{x} \in E$, $\|f \circ g(\mathbf{x})\| = \|g(\mathbf{x})\|$ car f est une isométrie, et donc $\|f \circ g(\mathbf{x})\| = \|\mathbf{x}\|$ car g est une isométrie. Donc la composition est bien une opération interne à $O(E)$.

L'identité est un élément neutre pour la composition, c'est une isométrie.

Si $f \in O(E)$ et si $\mathbf{x} \in \ker(f)$, $0 = \|f(\mathbf{x})\| = \|\mathbf{x}\|$, donc $\mathbf{x} = 0$, donc f est inversible. De plus, pour tout $\mathbf{x} \in E$, on a $\|\mathbf{x}\| = \|f(f^{-1}(\mathbf{x}))\| = \|f^{-1}(\mathbf{x})\|$ car f est une isométrie, donc f^{-1} est aussi une isométrie. □

Proposition 157. *Soit f est un endomorphisme de E de matrice $\boldsymbol{A}$ dans une base B.*

Si f est une isométrie, $\boldsymbol{A}$ est inversible, vérifie $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$ et $\det(\boldsymbol{A}) \in \{-1, 1\}$.

Si $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$, alors f est orthogonale.

Preuve : Si f est une isométrie, f est inversible donc $\mathbf{A}$ aussi. Comme f est une transformation orthogonale, pour tout $\mathbf{u}$ et $\mathbf{v}$ et $\mathbf{u}_c$, $\mathbf{v}_c$ les vecteurs colonnes associés à $\mathbf{u}$ et $\mathbf{v}$ dans une base orthonormale de E,

$$\mathbf{u}_c^T \mathbf{A}^T \mathbf{A} \mathbf{v}_c = (\mathbf{A}\mathbf{u}_c)^T (\mathbf{A}\mathbf{v}_c) = \langle f(\mathbf{u}), f(\mathbf{v}) \rangle = \langle u, v \rangle = \mathbf{u}_c^T \mathbf{v}_c \ .$$

Comme c'est vrai pour tous les vecteurs $\mathbf{u}_c$ et $\mathbf{v}_c$, on en déduit

$$\mathbf{A}^T \mathbf{A} = \mathbf{I} \ .$$

En prenant le déterminant dans cette égalité, on trouve alors $\det(\mathbf{A})^2 = 1$.

Réciproquement, si $\mathbf{A}^T \mathbf{A} = \mathbf{I}$,

$$\langle f(\mathbf{u}), f(\mathbf{v}) \rangle = (\mathbf{A}\mathbf{u}_c)^T (\mathbf{A}\mathbf{v}_c) = \mathbf{u}_c^T \mathbf{A}^T \mathbf{A} \mathbf{v}_c = \mathbf{u}_c^T \mathbf{v}_c = \langle u, v \rangle \ .$$

Donc f est orthogonale. □

Proposition 158. *Soit f un endomorphisme de E et $B = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base orthonormée de E.*

1. *Si f est orthogonale, alors $(f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n))$ est une base orthonormée.*
2. *Si $(f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n))$ est une base orthonormée alors f est orthogonale.*

Preuve : Si f est orthogonale, $(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n))$ est une famille orthogonale, dont les vecteurs sont de norme 1. donc une base orthonormée de E.

Soit $\mathbf{u}$ un vecteur de E. On a $\mathbf{u} = \sum_{i=1}^n u_i \mathbf{e}_i$ et, comme B est une base orthonormée,

$$\|\mathbf{u}\|^2 = \sum_{i=1}^n u_i^2 \ .$$

De plus, on a

$$f(\mathbf{u}) = \sum_{i=1}^n u_i f(\mathbf{e}_i) \ .$$

Si $(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n))$ est une base orthonormée, on a donc

$$\|f(\mathbf{u})\|^2 = \sum_{i=1}^n u_i^2 = \|\mathbf{u}\|^2 \ .$$

Donc f est orthogonale. □

Corollaire 159. *Soit B et B' deux bases orthonormées de E. La matrice de l'identité de E muni de la base B' dans E muni de la base B est orthogonale.*

Proposition 160. *Tout espace Euclidien de dimension n est isométrique à $\mathbb{R}^n$ muni du produit scalaire usuel.*

Proposition 161. *Les valeurs propres d'une transformation orthogonale sont dans $\{-1, 1\}$.*

5.8 Endomorphisme adjoint et autoadjoint

5.8.1 Adjoint d'un endomorphisme

Soit E un espace Euclidien et f un endomorphisme de E.

Proposition 162. *Il existe une unique application de $f^* : E \to E$, appelée adjoint de f, telle que,*

$$\forall \boldsymbol{u}, \boldsymbol{v} \in E, \qquad \langle f(\boldsymbol{u}), \boldsymbol{v}\rangle = \langle \boldsymbol{u}, f^*(\boldsymbol{v})\rangle \ .$$

f^* est un endomorphisme de E et, si B est une base de E et $\boldsymbol{A}$ est la matrice de f dans la base B, alors f^* a pour matrice $\boldsymbol{A}^T$ dans cette base.

Démonstration. Soit B une base orthonormée de E, $\mathbf{A}$ la matrice de f dans cette base et soit f' l'endomorphisme de E de matrice $\mathbf{A}^T$ dans la base B. On a, pour tout $\mathbf{u}$ et $\mathbf{v}$ de E,

$$\langle f(\mathbf{u}), \mathbf{v}\rangle = (\mathbf{A}\mathbf{u}_c)^T\mathbf{v}_c = \mathbf{u}_c^T(\mathbf{A}^T\mathbf{v}_c) = \langle \mathbf{u}, f'(\mathbf{v})\rangle \ .$$

On en déduit que f' est un endomorphisme satisfaisant la propriété définissant f^*, en outre, sa matrice est bien $\mathbf{A}^T$. Soit f'' une application vérifiant

$$\forall \mathbf{u}, \mathbf{v} \in E, \qquad \langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f''(\mathbf{v})\rangle \ .$$

On a alors,

$$\forall \mathbf{u}, \mathbf{v} \in E, \qquad 0 = \langle \mathbf{u}, f''(\mathbf{v}) - f'(\mathbf{v})\rangle \ .$$

On en déduit que,

$$\forall \mathbf{v} \in E, \qquad f''(\mathbf{v}) = f'(\mathbf{v}) \ .$$

□

Soit E un espace Euclidien, f et g deux endomorphismes de E. Les propriétés suivantes sont une conséquence du fait que la matrice de l'adjoint dans une base orthonormée est la transposée.

Proposition 163.
1. *$(f^*)^* = f$.*
2. *$(f+g)^* = f^* + g^*$.*
3. *$(\lambda f)^* = \lambda f^*$.*
4. *$(g \circ f)^* = f^* \circ g^*$.*

Proposition 164. *Soit B une base quelconque de E et $\boldsymbol{A}$ la matrice de f dans la base B. Soit $\boldsymbol{\Phi}$ la matrice du produit scalaire dans la base B. La matrice de f^* dans la base B est $\boldsymbol{\Phi}^{-1}\boldsymbol{A}^T\boldsymbol{\Phi}$.*

Preuve : Soit $\mathbf{u}$, $\mathbf{v}$ deux vecteurs de E et $\mathbf{u}_c$, $\mathbf{v}_c$ les vecteurs colonnes de leurs coordonnées respectives dans la base B. On a

$$\langle f(\mathbf{u}), \mathbf{v}\rangle = (\mathbf{A}\mathbf{u}_c)^T\mathbf{\Phi}\mathbf{v}_c = \mathbf{u}_c\mathbf{\Phi}(\mathbf{\Phi}^{-1}\mathbf{A}^T\mathbf{\Phi}\mathbf{v}_c) = \langle \mathbf{u}, f^*(\mathbf{v})\rangle \ .$$

On en déduit que la matrice de f^* est $\mathbf{\Phi}^{-1}\mathbf{A}^T\mathbf{\Phi}$. □

5.8.2 Endomorphisme autoadjoint

Soit E un espace Euclidien.

Definition 165. *Un endomorphisme f de E est dit autoadjoint ou symétrique si $f^* = f$.*

Proposition 166. *Un endomorphisme est autoadjoint si et seulement si sa matrice dans une base orthonormée est symétrique.*

Théorème 167. *Soit f un endomorphisme autoadjoint de l'espace Euclidien $E = \mathbb{R}^n$.*

1. *Les valeurs propres de f sont réelles.*
2. *Les sous-espaces propres de E sont orthogonaux.*
3. *f est diagonalisable.*

Preuve : **Preuve de 1.** Soit $\psi : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ l'application définie par $\psi(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^n u_i \bar{v}_i$. L'application ψ vérifie les propriétés suivantes.

1. $\psi(\mathbf{u} + \lambda\mathbf{u}', \mathbf{v}) = \psi(\mathbf{u}, \mathbf{v}) + \lambda\psi(\mathbf{u}, \mathbf{v})$.
2. $\psi(\mathbf{u}, \mathbf{v} + \lambda\mathbf{v}') = \psi(\mathbf{u}, \mathbf{v}) + \bar{\lambda}\psi(\mathbf{u}, \mathbf{v}')$.
3. $\psi(\mathbf{u}, \mathbf{u}) \geqslant 0$ et $\psi(\mathbf{u}, \mathbf{u}) = 0$ si et seulement si $\mathbf{u} = 0$.

Soit $\mathbf{A}$ la matrice de f dans la base canonique de $\mathbb{R}^n$. Soit g l'endomorphisme de $\mathbb{C}^n$ de matrice $\mathbf{A}$ dans la base canonique de $\mathbb{C}^n$. Le polynôme caractéristique de $\mathbf{A}$ est à coefficients réels, donc si λ est une racine dans $\mathbb{C}$ de $\chi(\mathbf{A})$, alors $\bar{\lambda}$ est aussi racine de $\chi(\mathbf{A})$. Comme $\mathbf{A}$ est symétrique et à coefficients réels, pour tous $\mathbf{u}$ et $\mathbf{v}$ de $\mathbb{C}^n$, on a

$$\psi(g(\mathbf{u}), \mathbf{v}) = \sum_{i=1}^n (\mathbf{A}\mathbf{u})_i \bar{v}_i = \mathbf{u}_c^T \mathbf{A} \bar{\mathbf{v}}_c = \mathbf{u}_c^T \overline{\mathbf{A}\mathbf{v}}_c = \psi(\mathbf{u}, g(\mathbf{v})) \ .$$

Soit λ une valeur propre complexe de g et $\mathbf{u}$ un vecteur propre de $\mathbb{C}^n$ associé. On a

$$\lambda\psi(\mathbf{u}, \mathbf{u}) = \psi(g(\mathbf{u}), \mathbf{u}) = \psi(\mathbf{u}, g(\mathbf{u})) = \psi(\mathbf{u}, \lambda\mathbf{u}) = \bar{\lambda}\psi(\mathbf{u}, \mathbf{u}) \ .$$

Comme $\psi(\mathbf{u}, \mathbf{u})$ est non nul, $\lambda = \bar{\lambda}$, donc $\lambda \in \mathbb{R}$.
Preuve de 2. Soient λ et μ deux valeurs propres distinctes de f et soit $\mathbf{u}$, $\mathbf{v}$ des vecteurs propres de f associés respectivement à λ et μ. Alors

$$\lambda \langle \mathbf{u}, \mathbf{v} \rangle = \langle \lambda\mathbf{u}, \mathbf{v} \rangle = \langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f(\mathbf{v}) \rangle = \langle \mathbf{u}, \mu\mathbf{v} \rangle = \mu \langle \mathbf{u}, \mathbf{v} \rangle \ .$$

On en déduit $(\lambda - \mu) \langle \mathbf{u}, \mathbf{v} \rangle = 0$ et, comme $\lambda - \mu \neq 0$, on a $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
Preuve de 3. Soit F la somme des espaces propres de f. f est diagonalisable si et seulement si $F = E$, c'est à dire, si $F^\perp = \{0\}$. Par définition, f n'a pas de vecteurs propres dans $F^\perp$.

De plus, $\mathbf{v} \in F^\perp$ si et seulement si $\langle \mathbf{v}, \mathbf{u} \rangle = 0$ pour tout vecteur propre $\mathbf{u}$ de f. Soit $\mathbf{v} \in F^\perp$ et $\mathbf{u}$ un vecteur propre de f. On a

$$\langle f(\mathbf{v}), \mathbf{u} \rangle = \langle \mathbf{v}, f(\mathbf{u}) \rangle = \lambda \langle \mathbf{v}, \mathbf{u} \rangle = 0 \ .$$

Donc si $\mathbf{v} \in F^\perp$, $f(\mathbf{v}) \in F^\perp$. L'endomorphisme $g : F^\perp \to F^\perp$ défini par $g(\mathbf{v}) = f(\mathbf{v})$, pour tout $\mathbf{v} \in F^\perp$ vérifie,

$$\forall \mathbf{u}, \mathbf{v} \in F^\perp, \qquad \langle g(\mathbf{u}), \mathbf{v}\rangle = \langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f(\mathbf{v})\rangle = \langle \mathbf{u}, g(\mathbf{v})\rangle \ .$$

Donc g est symétrique. Il s'en suit que ses valeurs propres (complexes) sont réelles. Comme il existe toujours une valeur propre complexe si $F^\perp \neq \{0\}$, il existe une valeur propre réelle de g si $F^\perp \neq \{0\}$. Si $F^\perp \neq \{0\}$, il existe donc $\mathbf{u} \in F^\perp \setminus \{0\}$ et $\lambda \in \mathbb{R}$ tels que $f(\mathbf{u}) = g(\mathbf{u}) = \lambda\mathbf{u}$. Donc $\mathbf{u}$ est un vecteur propre de f, ce qui est absurde. On en déduit que $F^\perp = \{0\}$. □

Corollaire 168. *Si $\boldsymbol{A}$ est une matrice symétrique réelle, il existe une matrice $\boldsymbol{P}$ orthogonale telle que $\boldsymbol{PAP}^T$ est diagonale.*

Preuve : Il suffit de prendre pour $\mathbf{P}$ la matrice de passage d'une base B orthonormée de E vers une base B' orthonormée de vecteurs propres de l'endomorphisme f de matrice $\mathbf{A}$ dans la base B. □

5.9 Exercices

Exercice 1

Les applications suivantes déterminent elles des produits scalaires sur $\mathbb{R}[X]$?

$$\varphi_1(P,Q) = P(0)Q(0)+P(1)Q(1), \qquad \varphi_2(P,Q) = \int_0^1 (P(t)Q(t)+P'(t)Q'(t))\mathrm{d}t \ .$$

Les applications suivantes déterminent elles des produits scalaires sur $\mathbb{R}^3$?

$$\varphi_1(\mathbf{u},\mathbf{v}) = u_1v_1+2u_2v_2+4u_3v_3, \qquad \varphi_2(\mathbf{u},\mathbf{v}) = u_1v_1+2u_1v_2+2u_2v_1+u_2v_2+4u_3v_3 \ .$$

Déterminer les valeurs de a pour lesquelles la forme bilinéaire $\varphi(\mathbf{u},\mathbf{v})$ suivante définit un produit scalaire.

$$\varphi(\mathbf{u},\mathbf{v}) = u_1v_1 + 6u_2v_2 + au_3v_3 - 2u_1v_2 - 2u_2v_1 - 3u_1v_3 - 3u_3v_1 \ .$$

Lorsque $a = 28$, donner la matrice $\mathbf{\Phi}$ du produit scalaire associé à la forme bilinéaire précédente. Déterminer une base B' dans laquelle la matrice de φ est diagonale. Donner un lien entre les deux matrices de φ.

Exercice 2

Montrer que les fonctions c_k et s_k $k \in \mathbb{N}^*$ forment une famille libre de $E = C([0,2\pi],\mathbb{R})$, où

$$c_k(t) = \cos(kt), \qquad s_k(t) = \sin(kt) \ .$$

Montrer que les fonctions e_k $k \in \mathbb{Z}$ forment une famille libre de $E = C([0,2\pi],\mathbb{C})$, où

$$e_k(t) = \mathrm{e}^{ikt} \ .$$

Exercice 3

Soit E un espace Euclidien et F_1, F_2, deux sous-espaces de E. Montrer que

$$(F_1+F_2)^\perp = F_1^\perp \cap F_2^\perp, \qquad (F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp \ .$$

Exercice 4

Montrer que, pour tout $n \geqslant 1$,

$$\sum_{k=1}^{n} k\sqrt{k} \leqslant n(n+1)\frac{\sqrt{2n+1}}{2\sqrt{3}} \ .$$

Exercice 5

Soit $E = \mathbb{R}_n[X]$ muni du produit scalaire $\langle P, Q\rangle = \int_0^1 P(t)Q(t)\mathrm{d}t$. Soit F l'espace engendré par les polynômes $X, X^2, \ldots, X^n$. Le but de l'exercice est de déterminer la distance d du polynôme 1 à l'espace F.

Soit

$$S(X) = \frac{(X-1)\ldots(X-n)}{(X+1)\ldots(X+n+1)} \ .$$

La décomposition de S en éléments simples s'écrit

$$S(X) = \frac{a_0}{X+1} + \ldots + \frac{a_n}{X+n+1} \ .$$

1. Calculer a_0.
2. Montrer que le polynôme $P(X) = a_0 + \ldots + a_n X^n$ est orthogonal à F. En déduire $F^\perp$.
3. Montrer que $d^2 = \|P\|^2/a_0^2$. En déduire d.

Exercice 6

Pour tout $x > 0$, notons E_x l'espace vectoriel des fonctions de classe C^1 sur $[0, x]$, nulles en 0.

1. Montrer que $\langle f, g\rangle_x = \int_0^x f'(t)g'(t)\mathrm{d}t$ définit un produit scalaire sur E_x.
2. Soit $f \in E_1$, montrer que $\sup_{x\in[0,1]} |f(x)| \leqslant \sqrt{\int_0^1 f'(t)^2\mathrm{d}t}$.

Exercice 7

Soit E l'espace des fonctions $f : [0,1] \to \mathbb{R}$ telles que $\int_0^1 f^2(t)/t\mathrm{d}t$ converge.

1. Montrer que E est un espace vectoriel.
2. Montrer $\varphi(f, g) = \int_0^1 f(t)g(t)/t\mathrm{d}t$ définit un produit scalaire sur E.

Exercice 8

Soit ℓ^2 l'espace des suites $(u_n)_{n\geqslant 0}$ de nombres réels tels que $\sum_{n=0}^{+\infty} u_n^2 < \infty$ est un espace préhilbertien pour un produit scalaire qu'on précisera.

Exercice 9

Déterminer le polynôme unitaire de degré 4 tel que $\int_{-1}^{1} P^2(t)\mathrm{d}t$ est minimal.

Exercice 10

Soit $n \geqslant 3$ un entier et soit $\mathbf{A}$ une matrice symétrique réelle telle que $\mathbf{A}^n = \mathbf{I}$. Montrer que $\mathbf{A}^2 = \mathbf{I}$.

Exercice 11

Soit $\mathbf{A}$ une matrice carrée de taille n associée à un produit scalaire.

1. Montrer que les valeurs propres de $\mathbf{A}$ sont strictement positives.
2. Montrer que $\det(\mathbf{A}) \leqslant (\mathrm{Tr}(\mathbf{A})/n)^n$.

Exercice 12

Soit $E = \mathbb{R}_n[X]$ muni du produit scalaire $\langle P, Q\rangle = \int_0^1 P(t)Q(t)\mathrm{d}t$. Soit $L : E \to E$ défini par

$$L(P)(x) = \int_0^1 (x+t)^n P(t)\mathrm{d}t \quad .$$

Montrer que L est autoadjoint.

Exercice 13

Soient f et g deux endomorphismes de $\mathbb{R}^n$, $B = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $B' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ deux bases orthonormées de $\mathbb{R}^n$. Soit f^* l'adjoint de f.

1. Montrer que $\sum_{i=1}^n \left\langle f(\mathbf{e}_i), g(\mathbf{e}'_j)\right\rangle^2 = \|f^* \circ g(\mathbf{e}'_j)\|^2$.
2. Montrer que $\sum_{1\leqslant i,j\leqslant n} \left\langle f(\mathbf{e}_i), g(\mathbf{e}'_j)\right\rangle^2$ ne dépend pas des bases B et B'.

Exercice 14

Soit E un espace Euclidien et $S = (\mathbf{u}_1, \ldots, \mathbf{u}_p)$ une famille de vecteurs de E. Le déterminant de Gram de S, noté $G(S)$ est le déterminant de la matrice carrée de coefficient générique $\langle \mathbf{u}_i, \mathbf{u}_j\rangle$.

1. Déterminer $G(S)$ lorsque S est une famille orthogonale.
2. Montrer que $G(S)$ ne change pas si on ajoute à un vecteur de S une combinaison linéaire des autres.
3. On construit à partir de S une famille orthogonale S' en appliquant le procédé de Gram-Schmidt. Montrer que $G(S) = G(S')$.
4. Soit V l'espace engendré par S, $\mathbf{u}$ un vecteur de E et $T = (\mathbf{u}_1, \ldots, \mathbf{u}_p, \mathbf{u})$. Montrer que
$$d(\mathbf{u}, V)^2 = \frac{G(T)}{G(S)} \quad .$$

Chapitre 6

Le problème des moindres carrés

Ce chapitre ne contient pas nouvelle théorie. Il présente plutôt des extensions classiques des résultats du cours qui sont particulièrement utiles en statistiques et en analyse numérique.

Soit $\mathbf{y} \in \mathbb{R}^n$ et $\mathbf{X}$ une matrice de type (n,p). L'idée générale est de chercher à résoudre l'équation $\mathbf{Xb} = \mathbf{y}$. Ceci n'étant pas toujours possible, on cherche de façon générale les vecteurs $\mathbf{b} \in \mathbb{R}^p$ tels que $Q(\mathbf{b}) = \|\mathbf{y} - \mathbf{Xb}\|^2$ soit minimale.

Definition 169. *Tout vecteur $\boldsymbol{b} \in \mathbb{R}^p$ tel que,*

$$\forall \boldsymbol{b}' \in \mathbb{R}^p, \qquad Q(\boldsymbol{b}) \leqslant Q(\boldsymbol{b}') \ ,$$

est appelé solution du problème des moindres carrés.

6.1 Les équations normales

On appelle gradient de la fonction Q le vecteur des dérivées partielles de Q par rapport à chaque variable b_i :

$$\nabla Q(\mathbf{b}) = \begin{bmatrix} \frac{\partial Q}{\partial b_1}(\mathbf{b}) \\ \vdots \\ \frac{\partial Q}{\partial b_p}(\mathbf{b}) \end{bmatrix} \ .$$

Tout vecteur $\mathbf{b}$ solution des moindres carrés vérifie $\nabla Q(\mathbf{b}) = 0$. Ce résultat n'est pas étonnant, il généralise une identité bien connue en analyse réelle, la dérivée d'une fonction s'annule en un minimiseur. Il sera prouvé formellement à la Proposition 175. Pour calculer le gradient de Q, la proposition suivante va être utile.

Proposition 170. *Soit $\boldsymbol{a}$ un vecteur de $\mathbb{R}^p$ et $\boldsymbol{A}$ une matrice carrée de taille p. Soit $\ell(\boldsymbol{b}) = \boldsymbol{a}^T\boldsymbol{b}$ et $q(\boldsymbol{b}) = \boldsymbol{b}^T\boldsymbol{A}\boldsymbol{b}$. Alors*

$$\nabla \ell(\boldsymbol{b}) = \boldsymbol{a}, \qquad \nabla q(\boldsymbol{b}) = (\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{b} \ .$$

Preuve. On a $\ell(\mathbf{b}) = a_1 b_1 + \ldots + a_n b_n$ de sorte que $\frac{\partial \ell}{\partial b_i}(\mathbf{b}) = a_i$. Ceci étant vrai pour tout, le premier résultat est démontré.

Soit maintenant $\mathbf{h} \in \mathbb{R}^p$, on a

$$q(\mathbf{b}+\mathbf{h}) = q(\mathbf{b}) + \mathbf{h}^T\mathbf{A}\mathbf{b} + \mathbf{b}^T\mathbf{A}\mathbf{h} + q(\mathbf{h}) = q(\mathbf{b}) + \mathbf{h}^T(\mathbf{A}+\mathbf{A}^T)\mathbf{b} + q(\mathbf{h}) \ .$$

Soit $B = (\mathbf{e}_1, \ldots, \mathbf{e}_p)$ la base canonique de $\mathbb{R}^p$ et soit $\mathbf{h} = h\mathbf{e}_i$. On a

$$\frac{\partial Q}{\partial b_i}(\mathbf{b}) = \lim_{h\to 0} \frac{q(\mathbf{b}+h\mathbf{e}_i) - q(\mathbf{b})}{h} = \mathbf{e}_i^T(\mathbf{A}+\mathbf{A}^T)\mathbf{b} \ .$$

Ceci démontre le second point. □

Nous pouvons maintenant établir le corollaire suivant.

Corollaire 171. *On a $\nabla Q(\boldsymbol{b}) = 0$ si et seulement si $\boldsymbol{b}$ est solution des équations normales :*

$$\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}^T\boldsymbol{y} \ .$$

Preuve. Pour tout vecteur $\mathbf{x} \in \mathbb{R}^p$, on a

$$\|\mathbf{x}\|^2 = \sum_{i=1}^p x_i^2 = \mathbf{x}^T\mathbf{x} = \mathrm{Tr}(\mathbf{x}\mathbf{x}^T) \ .$$

On a donc

$$\begin{aligned} Q(\mathbf{b}) &= (\mathbf{y}-\mathbf{X}\mathbf{b})^T(\mathbf{y}-\mathbf{X}\mathbf{b}) = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}\mathbf{b} - \mathbf{b}^T\mathbf{X}^T\mathbf{y} + \mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{b} \\ &= \mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{X}\mathbf{b} + \mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{b} \ . \end{aligned}$$

D'après la Proposition 170, on a donc $\nabla Q(\mathbf{b}) = -2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\mathbf{b}$, de sorte que $\nabla Q(\mathbf{b}) = 0$ est équivalent à $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$. □

6.2 La géométrie des moindres carrés

D'après le corollaire (171), pour montrer qu'il existe une solution à $\nabla Q(\mathbf{b}) = 0$, il suffit de montrer que $\mathbf{X}^T\mathbf{y}$ est dans l'image $\mathrm{Im}(\mathbf{X}^T\mathbf{X})$.

Lemme 172. *On a* $\ker(\boldsymbol{X}) = \ker(\boldsymbol{X}^T\boldsymbol{X})$.

Preuve. Exercice. □

Lemme 173. *Soit $\boldsymbol{X}$ une matrice de type (n,p), alors* $\ker(\boldsymbol{X})$ *et* $\mathrm{Im}(\boldsymbol{X}^T)$ *sont des sous-espaces orthogonaux.*

Preuve. Soit $\mathbf{u} \in \ker(\mathbf{X})$ et $\mathbf{v} \in \mathbb{R}^n$, on a $\mathbf{u}^T\mathbf{X}^T\mathbf{v} = (\mathbf{X}\mathbf{u})^T\mathbf{v} = 0$. □

Lemme 174. *On a* $\mathrm{Im}(\boldsymbol{X}^T) = \mathrm{Im}(\boldsymbol{X}^T\boldsymbol{X})$.

Preuve. Clairement $\mathrm{Im}(\mathbf{X}^T\mathbf{X}) \subset \mathrm{Im}(\mathbf{X}^T)$. D'autre part, si $\mathbf{x} \in \mathrm{Im}(\mathbf{X}^T)$, d'après le lemme 173, $\mathbf{x} \in \ker(\mathbf{X})^\perp$, donc

$$\begin{aligned} \dim(\mathrm{Im}(\mathbf{X}^T)) &\leqslant \dim(\ker(\mathbf{X})^\perp) = n - \dim(\ker(\mathbf{X})) \\ &= n - \dim(\ker(\mathbf{X}^T\mathbf{X})) = \dim(\mathrm{Im}(\mathbf{X}^T\mathbf{X})) \ . \end{aligned}$$

La première égalité vient de la proposition 146, la seconde du lemme 172 et la troisième du théorème du rang. Ceci termine la preuve. □

Le lemme 174 assure qu'il existe toujours une solution aux équations normales. D'après le résultat suivant, il existe donc toujours une solution au problème des moindres carrés.

Proposition 175. *Les solutions des équations normales sont les solutions du problème des moindres carrés.*

Preuve. Soit $\mathbf{b}$ une solution des équations normales et soit $\mathbf{b}' \in \mathbb{R}^p$. On a

$$Q(\mathbf{b}') = Q(\mathbf{b}) + 2(\mathbf{b}'-\mathbf{b})^T\mathbf{X}^T(\mathbf{y}-\mathbf{X}\mathbf{b}) + \|X(\mathbf{b}'-\mathbf{b})\|^2 = Q(\mathbf{b}) + \|X(\mathbf{b}'-\mathbf{b})\|^2 \ .$$

Ainsi, $\mathbf{b}$ est solution du problème des moindres carrés.

Inversement, si $\mathbf{b}$ minimise Q, on a nécessairement $(\mathbf{b}'-\mathbf{b})^T\mathbf{X}^T(\mathbf{y}-\mathbf{X}\mathbf{b}) = 0$ pour tout $\mathbf{b}' \in \mathbb{R}^p$, donc $\mathbf{X}^T(\mathbf{y}-\mathbf{X}\mathbf{b}) = 0$, donc $\mathbf{b}$ est solution des équations normales. □

Dans la preuve précédente, nous avons établi le corollaire suivant qui est fondamental pour comprendre la géométrie du problème des moindres carrés et dont nous discuterons les implications.

Corollaire 176. *Si $\boldsymbol{b}$ et $\boldsymbol{b}'$ sont deux solutions des équations normales, alors $\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}\boldsymbol{b}'$.*

Preuve. Par la preuve précédente, si $\mathbf{b}$ et $\mathbf{b}'$ minimisent Q, $Q(\mathbf{b}) = Q(\mathbf{b}')$ et $\|X(\mathbf{b}-\mathbf{b}')\|^2 = 0$, d'où le résultat. □

Ainsi, la différence entre deux solutions du problème des moindres carrés est un vecteur de $\ker(\mathbf{X}) = \ker(\mathbf{X}^T\mathbf{X})$. Le point le plus proche de $\mathbf{y}$ dans $\mathrm{Im}(\mathbf{X})$ est le vecteur $\mathbf{X}\mathbf{b}$, projection orthogonale de $\mathbf{y}$ sur $\mathrm{Im}(\mathbf{X})$, la différence $\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b}$ est dans l'orthogonale de $\mathrm{Im}(\mathbf{X})$ et vérifie en effet, d'après les équations normales

$$\mathbf{X}^T\mathbf{e} = \mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{X}\mathbf{b} = 0 \ .$$

D'après le théorème de Pythagore, on a aussi

$$\|\mathbf{y}\|^2 = \|\mathbf{X}\mathbf{b}\|^2 + \|\mathbf{e}\|^2 \ .$$

On aurait donc pu également construire la matrice de projection sur $\mathrm{Im}(\mathbf{X})$. Nous détaillons maintenant cette construction.

Proposition 177. *$\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}\boldsymbol{B}$ si et seulement si $\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}$.*

Preuve. Si $\mathbf{X}\mathbf{A} = \mathbf{X}\mathbf{B}$, en multipliant cette égalité à gauche par $\mathbf{X}^T$, on obtient que $\mathbf{X}^T\mathbf{X}\mathbf{A} = \mathbf{X}^T\mathbf{X}\mathbf{B}$.

Réciproquement, supposons que $\mathbf{X}^T\mathbf{X}\mathbf{A} = \mathbf{X}^T\mathbf{X}\mathbf{B}$. Alors, les colonnes de $\mathbf{X}(\mathbf{A}-\mathbf{B})$ sont dans $\ker(\mathbf{X}^T)$. Elles sont également clairement dans $\mathrm{Im}(\mathbf{X})$, et comme ces espaces sont orthogonaux d'après le lemme 173, les colonnes des $\mathbf{X}(\mathbf{A}-\mathbf{B})$ sont nulles. □

Definition 178. *Soit $\boldsymbol{A}$ une matrice de type (n,p). On appelle inverse généralisée de $\boldsymbol{A}$ toute matrice $\boldsymbol{A}^g$ telle que $\boldsymbol{A}\boldsymbol{A}^g\boldsymbol{A} = \boldsymbol{A}$.*

Proposition 179. *Toute matrice carrée $\boldsymbol{A}$ de taille n diagonalisable admet une inverse généralisée.*

Preuve. On écrit $\mathbf{A} = \mathbf{PDP}^{-1}$ avec $\mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$ et tous les $\lambda_i \neq 0$. Posons $\mathbf{D}^g = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_r^{-1}, 0, \ldots, 0)$. On a

$$\mathbf{DD}^g\mathbf{D} = \mathbf{D} \ ,$$

donc, en posant $\mathbf{A}^g = \mathbf{PD}^g\mathbf{P}^{-1}$, on en déduit

$$\mathbf{AA}^g\mathbf{A} = \mathbf{PDP}^{-1}\mathbf{PD}^g\mathbf{P}^{-1}\mathbf{PDP}^{-1} = \mathbf{A} \ .$$

□

Proposition 180. *Toute matrice $\boldsymbol{X}$ de type (n,p) admet une inverse généralisée. Si $(\boldsymbol{X}^T\boldsymbol{X})^g$ désigne une inverse généralisée de la matrice symétrique (donc diagonalisable) $\boldsymbol{X}^T\boldsymbol{X}$, alors $\boldsymbol{X}^g = (\boldsymbol{X}^T\boldsymbol{X})^g\boldsymbol{X}^T$ est une inverse généralisée de $\boldsymbol{X}$.*

Preuve. Calculons

$$\tilde{\mathbf{X}} = \mathbf{XX}^g\mathbf{X} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{X} \ .$$

Clairement, $\mathbf{X}^T\tilde{\mathbf{X}} = \mathbf{X}^T\mathbf{X}$ par définition de $(\mathbf{X}^T\mathbf{X})^g$. On déduit alors de la proposition 177 avec $\mathbf{B} = \mathbf{I}$ et $\mathbf{A} = (\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{X}$ que

$$\tilde{\mathbf{X}} = \mathbf{XI} = \mathbf{X} \ .$$

□

Theorem 181. *La matrice $\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^g\boldsymbol{X}^T$ est la matrice de la projection orthogonale sur $\mathrm{Im}(\boldsymbol{X})$. En particulier, on a donc*

1. $\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{P}_{\boldsymbol{X}}$.
2. $\mathrm{Im}(\boldsymbol{P}_{\boldsymbol{X}}) = \mathrm{Im}(\boldsymbol{X})$.
3. *$\boldsymbol{P}_{\boldsymbol{X}}$ ne dépend pas de l'inverse généralisée $(\boldsymbol{X}^T\boldsymbol{X})^g$ de $\boldsymbol{X}^T\boldsymbol{X}$.*
4. $\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{P}_{\boldsymbol{X}}^T$.

Preuve. D'après la proposition 180, on a $\mathbf{P}_\mathbf{X} = \mathbf{XX}^g$, donc

$$\mathbf{P}_\mathbf{X}\mathbf{P}_\mathbf{X} = \mathbf{XX}^g\mathbf{XX}^g = \mathbf{XX}^g = \mathbf{P}_\mathbf{X} \ .$$

Clairement $\mathrm{Im}(\mathbf{P}_\mathbf{X}) \subset \mathrm{Im}(\mathbf{X})$. De plus, si $\mathbf{Xu} \in \mathrm{Im}(\mathbf{X})$, on a

$$\mathbf{Xu} = \mathbf{XX}^g\mathbf{Xu} \in \mathrm{Im}(\mathbf{P}_\mathbf{X}) \ .$$

Soient $\mathbf{G}_1$, $\mathbf{G}_2$ deux inverses généralisées de $\mathbf{X}^T\mathbf{X}$. on a donc

$$\mathbf{X}^T\mathbf{XG}_1\mathbf{X}^T\mathbf{X} = \mathbf{X}^T\mathbf{XG}_2\mathbf{X}^T\mathbf{X}$$

D'après la proposition 177, on en déduit

$$\mathbf{XG}_1\mathbf{X}^T\mathbf{X} = \mathbf{XG}_2\mathbf{X}^T\mathbf{X} \ .$$

En transposant, on obtient alors

$$\mathbf{X}^T\mathbf{XG}_1^T\mathbf{X}^T = \mathbf{X}^T\mathbf{XG}_2^T\mathbf{X}^T \ .$$

En appliquant à nouveau la proposition 177, on en déduit

$$\mathbf{X}\mathbf{G}_1^T\mathbf{X}^T = \mathbf{X}\mathbf{G}_2^T\mathbf{X}^T \ .$$

En transposant à nouveau, on obtient le résultat 3.

Enfin, $\mathbf{X}^T\mathbf{X}$ étant symétrique réelle, il existe $\mathbf{P}$ orthogonale telle que $\mathbf{X}^T\mathbf{X} = \mathbf{P}\mathbf{D}\mathbf{P}^T$. On a $\mathbf{G} = \mathbf{P}\mathbf{D}^g\mathbf{P}^T$ avec $\mathbf{D}^g$ une inverse généralisée diagonale de $\mathbf{D}$ est une inverse généralisée symétrique de $\mathbf{X}^T\mathbf{X}$. D'après le point 3. on a donc $\mathbf{P}_\mathbf{X} = \mathbf{X}\mathbf{G}\mathbf{X}^T$ est symétrique. □

Proposition 182. *La matrice* $\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{X}}$ *est la matrice de la projection orthogonale sur* $\ker(\boldsymbol{X}^T)$. *En particulier,* $\boldsymbol{X}^T(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{X}}) = 0$.

Démonstration. On a vu au chapitre précédent que $\mathbf{I}-\mathbf{P}_\mathbf{X}$ est la matrice de la projection orthogonale sur $\mathrm{Im}(\mathbf{X})^\perp$ et au lemme 173 que $\mathrm{Im}(\mathbf{X})^\perp = \ker(\mathbf{X}^T)$. Le point 1 s'en déduit immédiatement.

Pour le point 2, on a pour tout $\mathbf{u} \in \mathbb{R}^p$ et $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{u}^T\mathbf{X}^T(\mathbf{I}-\mathbf{P}_\mathbf{X})\mathbf{v} = 0$, d'où le résultat. □

On peut maintenant réécrire que toute solution du problème des moindres carrés $\mathbf{b}$ vérifie $\mathbf{X}\mathbf{b} = \mathbf{P}_\mathbf{X}\mathbf{y}$. Le résidu $\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b} = (\mathbf{I}-\mathbf{P}_\mathbf{X})\mathbf{y}$ est donc la projection de $\mathbf{y}$ sur son complémentaire orthogonal.

On a ainsi, pour tout $\mathbf{b}' \in \mathbb{R}^p$, d'après la relation de Pythagore

$$Q(\mathbf{b}') = \|\mathbf{y} - \mathbf{P}_\mathbf{X}\mathbf{y} + \mathbf{P}_\mathbf{X}\mathbf{y} - \mathbf{X}\mathbf{b}'\|^2 = \|\mathbf{y} - \mathbf{P}_\mathbf{X}\mathbf{y}\|^2 + \|\mathbf{P}_\mathbf{X}\mathbf{y} - \mathbf{X}\mathbf{b}'\|^2 \ .$$

Dans cette décomposition, le premier terme ne change pas en modifiant $\mathbf{b}'$ tandis que le second, positif, peut être annulé car $\mathrm{Im}(\mathbf{P}_\mathbf{X}) = \mathrm{Im}(\mathbf{X})$. On a le résultat suivant.

Proposition 183. *Les solutions des équations normales sont les solutions de* $\boldsymbol{X}\boldsymbol{b} = \boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y}$.

Preuve. Si $\mathbf{b}$ vérifie $\mathbf{X}\mathbf{b} = \mathbf{P}_\mathbf{X}\mathbf{y}$, on a par la proposition 182, $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{P}_\mathbf{X}\mathbf{y} = \mathbf{X}^T\mathbf{y}$. Réciproquement, si $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$, on a

$$\mathbf{X}^T\mathbf{P}_\mathbf{X}\mathbf{y} = \mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{y} = \mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{X}\mathbf{b} \ ,$$

donc par la proposition 177 appliquée avec $\mathbf{A} = \mathbf{b}$ et $\mathbf{B} = (\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{y}$,

$$\mathbf{X}\mathbf{b} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^g\mathbf{X}^T\mathbf{y} = \mathbf{P}_\mathbf{X}\mathbf{y} \ .$$

□

Concluons ce paragraphe par un théorème très utile en statistiques.

Théorème 184. *Si* $\mathrm{Im}(\boldsymbol{W}) \subset \mathrm{Im}(\boldsymbol{X})$, *alors* $\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{W}}$ *est la projection orthogonale sur* $\mathrm{Im}((\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{W}})\boldsymbol{X})$.

Preuve. C'est clairement une projection orthogonale, il suffit donc de vérifier les images. Soit $\mathbf{u} \in \mathbb{R}^p$, on a $\mathbf{X}\mathbf{u} = \mathbf{P}_\mathbf{X}\mathbf{X}\mathbf{u}$ car $\mathbf{X}\mathbf{u} \in \mathrm{Im}(\mathbf{X}) = \mathrm{Im}(\mathbf{P}_\mathbf{X})$, donc

$$(\mathbf{I}-\mathbf{P}_\mathbf{W})\mathbf{X}\mathbf{u} = (\mathbf{P}_\mathbf{X}-\mathbf{P}_\mathbf{W})(\mathbf{X}\mathbf{u}) \in \mathrm{Im}(\mathbf{P}_\mathbf{X}-\mathbf{P}_\mathbf{W}) \ .$$

Réciproquement, soit $\mathbf{u} \in \mathbb{R}^n$ et $\mathbf{v} = (\mathbf{P_X} - \mathbf{P_W})\mathbf{u}$. Comme $\mathbf{P_X u} \in \mathrm{Im}(\mathbf{P_X}) = \mathrm{Im}(\mathbf{X})$, il existe $\mathbf{w} \in \mathbb{R}^n$ tel que $\mathbf{P_X u} = \mathbf{Xw}$. Or, comme $\mathrm{Im}(\mathbf{W}) \subset \mathrm{Im}(\mathbf{X})$, tout vecteur orthogonal à $\mathrm{Im}(\mathbf{X})$ est orthogonal à $\mathrm{Im}(\mathbf{W})$, donc $(\mathbf{I} - \mathbf{P_X})\mathbf{u}$ est orthogonal à $\mathrm{Im}(\mathbf{W})$, donc $\mathbf{P_W}(\mathbf{I} - \mathbf{P_X})\mathbf{u} = 0$. Ainsi

$$\mathbf{v} = (\mathbf{P_X} - \mathbf{P_W})\mathbf{u} = \mathbf{Xw} - \mathbf{P_W P_X u} = \mathbf{Xw} - \mathbf{P_W X w} = (\mathbf{I} - \mathbf{P_W})\mathbf{Xw} \ .$$

Ainsi, $\mathbf{v} \in \mathrm{Im}((\mathbf{I} - \mathbf{P_W})\mathbf{X})$. □

6.3 Décomposition QR

Le principal résultat de cette section est le suivant.

Théorème 185. *Soit $\boldsymbol{X}$ une matrice de type (n,p) à coefficients réels, de rang d. Il existe une matrice $\boldsymbol{Q}$ une matrice telle que $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_d$ et $\boldsymbol{R}$ une matrice triangulaire supérieure (i.e. de coefficient générique $r_{i,j}$ tel que $r_{i,j} = 0$, $\forall i > j$) telles que $\boldsymbol{X} = \boldsymbol{QR}$.*

Nous présentons deux preuves de ce résultat et discutons une application au problème des moindres carrés.

6.3.1 La méthode de Gram-Schmidt

Soit $\mathbf{X}_{:,1}, \dots, \mathbf{X}_{:,p}$ les vecteurs colonnes de $\mathbf{X}$. On applique l'algorithme d'orthonormalisation de Gram-Schmidt à la famille $(\mathbf{X}_{:,1}, \dots, \mathbf{X}_{:,p})$. On produit récursivement une famille $\mathbf{Q}_{:,1}, \dots, \mathbf{Q}_{:,d}$ orthonormée telle que,

$$\forall j \in \{1, \dots, p\}, \qquad \mathbf{X}_{:,j} = \sum_{i=1}^{d_j} r_{i,j}\mathbf{Q}_{:,i} \ , \tag{6.1}$$

ou la suite d'entiers d_j vérifie $d_1 \in \{0,1\}$ et $0 \leqslant d_{j+1} - d_j \leqslant 1$, de sorte que $d_j \leqslant j$ pour tout i. Définissons alors $r_{i,j} = 0$ si $i \in \{d_j + 1, \dots, d\}$ de sorte que la matrice $\mathbf{R}$ de type (d,p), de coefficient générique $r_{i,j}$, est triangulaire supérieure. Définissons également la matrice $\mathbf{Q}$ de type (n,d) dont les colonnes sont les vecteurs $\mathbf{Q}_{:,1}, \dots, \mathbf{Q}_{:,d}$. Comme la famille $\mathbf{Q}_{:,1}, \dots, \mathbf{Q}_{:,d}$ est orthonormée, on a $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_d$. Enfin, les relations (6.1) impliquent $\mathbf{X} = \mathbf{QR}$.

6.3.2 Matrices de Householder

Definition 186. *Soit $\boldsymbol{u}$ un vecteur non nul de $\mathbb{R}^n$, la matrice de Householder associée au vecteur $\boldsymbol{u}$ est la matrice $\boldsymbol{U} = \boldsymbol{I} - 2(\boldsymbol{u}\boldsymbol{u}^T)/(\boldsymbol{u}^T\boldsymbol{u})$.*

Remarque 187. *On montre sans peine que la matrice $(\boldsymbol{u}\boldsymbol{u}^T)/(\boldsymbol{u}^T\boldsymbol{u})$ est la matrice de la projection orthogonale sur* $\mathrm{Vect}(\boldsymbol{u})$ *(vérifier le !). Géométriquement, la matrice de Householder est donc la matrice de l'application laissant invariant les vecteurs de $\boldsymbol{u}^\perp$ et envoyant les vecteurs colinéaires $\boldsymbol{v}$ à $\boldsymbol{u}$ sur leur opposé $-\boldsymbol{v}$. On l'appelle parfois la symétrie orthogonale par rapport à $\boldsymbol{u}^\perp$.*

Proposition 188. *Une matrice de Householder $\boldsymbol{U}$ est symétrique et orthogonale.*

Preuve. Exercice. □

Nous pouvons maintenant démontrer le théorème 184. On procède par récurrence et considérons pour cela la propriété suivante.

$P(k)$: Il existe des matrices $\mathbf{U}_1^*, \ldots, \mathbf{U}_k^*$ de Householder ou identité, carrées et de tailles respectives $n, \ldots, n-k+1$ telles que

$$\mathbf{U}_k \ldots \mathbf{U}_1 \mathbf{X} = \begin{bmatrix} \mathbf{T}_k & \mathbf{A} \\ 0 & \mathbf{M}_{k+1} \end{bmatrix} \ .$$

Dans cette décomposition, $\mathbf{T}_k$ est une matrice triangulaire supérieure de taille k, $\mathbf{M}_{k+1}$ est une matrice carrée de type $(n-k, p-k)$ et

$$\mathbf{U}_k = \begin{bmatrix} \mathbf{I}_{k-1} & 0 \\ 0 & \mathbf{U}_k^* \end{bmatrix}$$

Montrons $P(1)$, l'hérédité est alors immédiate. Considérons $\mathbf{e}_1$ le premier vecteur de la base canonique et supposons $\mathbf{X}_{:,1}$ n'est pas colinéaire à $\mathbf{e}_1$, sinon, la propriété est vraie. Soit $\mathbf{u} = \mathbf{X}_{:,1} - \|\mathbf{X}_{:,1}\|\mathbf{e}_1$ et $\mathbf{U}_1$ la matrice de Householder associée au vecteur $\mathbf{u}$. On a

$$\mathbf{U}_1\mathbf{X} = \begin{bmatrix} \mathbf{U}_1\mathbf{X}_{:,1} & \ldots & \mathbf{U}_1\mathbf{X}_{:,p} \end{bmatrix} \ .$$

Ainsi, il suffit de calculer $\mathbf{U}_1\mathbf{X}_{:,1}$. Soit $\mathbf{x} = \mathbf{X}_{:,1}/\|\mathbf{X}_{:,1}\|$. On a

$$\mathbf{U}_1\mathbf{X}_{:,1} = \|\mathbf{X}_{:,1}\| \left(\mathbf{x} - \frac{2(\mathbf{x}-\mathbf{e}_1)^T\mathbf{x}}{(\mathbf{x}-\mathbf{e}_1)^T(\mathbf{x}-\mathbf{e}_1)}(\mathbf{x}-\mathbf{e}_1) \right) = \|\mathbf{X}_{:,1}\|\mathbf{e}_1 \ .$$

Ceci démontre $P(1)$. L'hérédité se démontre en utilisant les mêmes idées. (Faites le en exercice !)

Le théorème 184 est alors une conséquence directe de $P(n \wedge p)$.

6.3.3 Application à la résolution de systèmes linéaires

Supposons qu'on souhaite déterminer $\mathbf{b}$ tel que

$$\mathbf{X}\mathbf{b} = \mathbf{c}$$

On a d'après le théorème 184 $\mathbf{X} = \mathbf{Q}\mathbf{R}$, avec $\mathbf{Q}$ telle que $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_d$ et $\mathbf{R}$ triangulaire supérieure. Donc en multipliant à gauche par $\mathbf{Q}^T$, le système à résoudre est équivalent à

$$\mathbf{R}\mathbf{b} = \mathbf{Q}^T\mathbf{c} \ .$$

C'est un système linéaire triangulaire supérieur qui se résout sans difficulté.

6.4 Décomposition en valeurs singulières et pseudo-inverse de Moore-Penrose

6.4.1 Décomposition en valeurs singulières

Théorème 189. *Pour toute matrice $\boldsymbol{X}$ de type (n,p), il existe $\boldsymbol{U}$ orthogonale carrée de taille n, $\boldsymbol{V}$ orthogonale carrée de taille p et $\boldsymbol{D}$ de type (n,p), à coefficients nuls hors diagonale et positifs sur la diagonale telles que $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$.*

Les valeurs de la diagonale de $\boldsymbol{D}$ sont appelées les valeurs singulières de $\boldsymbol{X}$ et la décomposition $\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$ est appelée décomposition de $\mathbf{X}$ en valeurs singulières*.*

Preuve. Soit r le rang de $\mathbf{X}$. La matrice $\mathbf{X}^T\mathbf{X}$ étant symétrique à coefficients réels de taille p, le théorème spectral assure qu'elle est diagonalisable, donc qu'il existe une matrice orthogonale $\mathbf{P}$ carrée de taille p et une matrice diagonale carrée de taille r à coefficients strictement positifs $\mathbf{D}_1$ telles que

$$\mathbf{P}^T(\mathbf{X}^T\mathbf{X})\mathbf{P} = \begin{bmatrix} \mathbf{D}_1 & 0 \\ 0 & 0 \end{bmatrix}$$

Soit alors $\mathbf{P}_1$ la matrice des r premières colonnes de $\mathbf{P}$ et $\mathbf{P}_2$ la matrice des $p-r$ dernières colonnes de $\mathbf{P}$, on a $\mathbf{P}_1^T(\mathbf{X}^T\mathbf{X})\mathbf{P}_1 = \mathbf{D}_1$ et $\mathbf{X}\mathbf{P}_2 = 0$ (car les vecteurs colonnes de cette matrice sont de norme nulle). Soit alors $\mathbf{U}_1 = \mathbf{D}_1^{-1/2}\mathbf{P}_1^T\mathbf{X}^T$, on a

$$\mathbf{U}_1\mathbf{X}\mathbf{P}_1 = \mathbf{D}_1^{-1/2}\mathbf{P}_1^T\mathbf{X}^T\mathbf{X}\mathbf{P}_1 = \mathbf{D}_1^{1/2} \ .$$

De plus,

$$\mathbf{U}_1\mathbf{U}_1^T = \mathbf{D}_1^{-1/2}\mathbf{P}_1^T\mathbf{X}^T\mathbf{X}\mathbf{P}_1\mathbf{D}_1^{-1/2} = \mathbf{I}_r \ .$$

La seconde équation montre que les lignes de $\mathbf{U}_1$ forment un système orthonormé de r vecteurs de $\mathbb{R}^n$ qu'on peut compléter en une base orthonormée de $\mathbb{R}^n$. La matrice $\mathbf{U}$ dont les r premières lignes sont formées par $\mathbf{U}_1$ et les $n-r$ dernières sont données par la matrice $\mathbf{U}_2$ des vecteurs complémentaires de la base orthonormée est donc orthogonale. De plus, on a donc

$$\mathbf{U}\mathbf{X}\mathbf{P} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{X}\mathbf{P}_1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1^{1/2} & 0 \\ \mathbf{U}_2\mathbf{X}\mathbf{P}_1 & 0 \end{bmatrix}$$

Enfin, comme $0 = \mathbf{U}_2\mathbf{U}_1^T = \mathbf{U}_2\mathbf{X}\mathbf{P}\mathbf{D}_1^{-1/2}$, on a, en multipliant à droite par $\mathbf{D}_1^{1/2}$, $\mathbf{U}_2\mathbf{X}\mathbf{P}_1 = 0$, ce qui conclut la preuve. □

6.4.2 Pseudo-inverse de Moore-Penrose.

Lorsque le système $\mathbf{X}\mathbf{b} = \mathbf{c}$ admet plusieurs solutions, une idée pour en choisir une est de prendre celle de plus petite norme. Dans cette section, nous allons montrer qu'une telle solution est unique et allons utiliser pour cela décomposition en valeurs singulières de $\mathbf{X}$.

Proposition 190. *Soit $\boldsymbol{X}$ une matrice de type (n,p) et $\boldsymbol{c} \in \mathrm{Im}(\boldsymbol{X})$. Soit $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$ la décomposition de $\boldsymbol{X}$ en valeurs singulières. Alors*

1. *$\boldsymbol{X}^g = \boldsymbol{V}\boldsymbol{D}^g\boldsymbol{U}^T$, où $\boldsymbol{D}^g$ est la matrice obtenue en inversant les coefficients non nuls de $\boldsymbol{D}$ (et en laissant nuls les autres) est une inverse généralisée de $\boldsymbol{X}$.*
2. *$\boldsymbol{b}_0 = \boldsymbol{X}^g\boldsymbol{c}$ vérifie $\boldsymbol{X}\boldsymbol{b}_0 = \boldsymbol{c}$.*
3. *Pour tout $\boldsymbol{b}$ tel que $\boldsymbol{X}\boldsymbol{b} = \boldsymbol{c}$, on a $\|\boldsymbol{b}_0\| \leqslant \|\boldsymbol{b}\|$ avec égalité si et seulement si $\boldsymbol{b} = \boldsymbol{b}_0$.*

$\boldsymbol{X}^g$ est appelée inverse de Moore-Penrose de $\boldsymbol{X}$.

Preuve. On a

$$\mathbf{X}\mathbf{X}^g\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V}\mathbf{D}^g\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{X} \ .$$

De plus, $\mathbf{X}\mathbf{b}_0 = \mathbf{X}\mathbf{X}^g\mathbf{c} = \mathbf{P}_{\mathbf{X}}\mathbf{c} = \mathbf{c}$ car $\mathbf{c} \in \mathrm{Im}(\mathbf{X})$.

L'item 3. est le plus délicat. On a $\mathbf{x} \in \ker(\mathbf{X})$ si et seulement si $\mathbf{D}\mathbf{V}^T\mathbf{x} = 0$, i.e., si et seulement si $\mathbf{D}^g\mathbf{V}^T\mathbf{x} = 0$, i.e., si et seulement si $\mathbf{x} \in \ker((\mathbf{X}^g)^\perp)$. Donc

$\ker(\mathbf{X}) = \ker((\mathbf{X}^g)^{\perp})$. Soit maintenant $\mathbf{b}$ tel que $\mathbf{X}\mathbf{b} = \mathbf{c}$, on a $\mathbf{X}(\mathbf{b} - \mathbf{b}_0) = 0$, donc $\mathbf{b} - \mathbf{b}_0 \in \ker((\mathbf{X}^g)^{\perp})$. Ainsi,

$$\|\mathbf{b}\|^2 = \|(\mathbf{b}-\mathbf{b}_0)+\mathbf{X}^g\mathbf{c}\|^2 = \|\mathbf{b}_0\|^2+\|\mathbf{b}-\mathbf{b}_0\|^2+2\mathbf{c}^T(\mathbf{X}^g)^T(\mathbf{b}-\mathbf{b}_0) = \|\mathbf{b}_0\|^2+\|\mathbf{b}-\mathbf{b}_0\|^2 \ .$$

$\square$