

Université Paris-Dauphine  
DE MI2E, 2ème année

Algèbre linéaire 3 :  
normes, produits scalaires,  
espaces euclidiens, formes quadratiques.

Cours 2017/2018

Stéphane Mischler

(adapté des notes de cours de Olivier Glass)

Le polycopié qui suit peut avoir des différences notables avec le cours dispensé en amphitheâtre (qui seul fixe le programme de l'examen). Il comporte des passages qui ne seront pas traités en amphitheâtre, et a contrario dans ce dernier pourront être donnés des compléments ne figurant pas dans ces notes, les preuves être plus détaillées, etc.

Certains points plus délicats (en particulier certaines démonstrations) sont signalés par un symbole de virages dangereux. Ils peuvent être passés en première lecture (mais pas nécessairement en seconde...)

## Table des matières

<b>1</b>	<b>Espaces euclidiens et espaces vectoriels normés</b>	<b>1</b>
1.1	Rappel : espaces vectoriels . . . . .	1
1.2	Rappel : Famille génératrice, libre, bases . . . . .	3
1.3	Rappel : espaces vectoriels normés . . . . .	6
1.4	Définition d'un espace euclidien . . . . .	9
1.5	Exemples d'espaces euclidiens . . . . .	11
1.5.1	Produit scalaire canonique de $\mathbb{R}^n$ . . . . .	11
1.5.2	Autres exemples d'espaces munis d'un produit scalaire : les fonctions continues . . . . .	12
1.6	Premières propriétés des espaces euclidiens . . . . .	12
1.6.1	Identité du parallélogramme . . . . .	12
1.6.2	Angle non-orienté des deux vecteurs . . . . .	13
1.6.3	Théorème de Pythagore . . . . .	13
1.6.4	Orthogonalité et liberté . . . . .	14
1.7	Bases orthonormées, procédé d'orthonormalisation de Schmidt . . . . .	14
1.7.1	Motivation . . . . .	14
1.7.2	Bases orthonormées . . . . .	14
1.7.3	Procédé d'orthonormalisation de Schmidt . . . . .	15
<b>2</b>	<b>Espaces euclidiens, dualité et applications linéaires</b>	<b>17</b>
2.1	Rappel : matrice d'application linéaire . . . . .	17
2.2	Structure des espaces euclidiens par les bases orthonormées. Equivalence avec $\mathbb{R}^n$ . . . . .	19
2.3	Projections . . . . .	20
2.3.1	Définition . . . . .	20
2.3.2	Application au problème des moindres carrés . . . . .	22

2.4	Dualité . . . . .	24
2.4.1	Rappel : dual d'un espace vectoriel . . . . .	24
2.4.2	Structure du dual d'un espace euclidien . . . . .	26
2.5	Changement de base orthonormée et groupe orthogonal . . . . .	27
2.5.1	Groupe orthogonal . . . . .	27
2.5.2	Endomorphismes orthogonaux . . . . .	28
<b>3</b>	<b>Formes bilinéaires et quadratiques</b>	<b>30</b>
3.1	Motivation . . . . .	30
3.2	Formes bilinéaires . . . . .	30
3.3	Formes quadratiques . . . . .	32
3.4	Méthode de Gauss . . . . .	34
3.5	Signature d'une forme quadratique . . . . .	37
<b>4</b>	<b>Théorème de diagonalisation</b>	<b>41</b>
4.1	Motivation . . . . .	41
4.2	Changement de base pour les matrices de forme bilinéaire . . . . .	41
4.3	Vecteurs et valeurs propres . . . . .	42
4.4	Théorème de diagonalisation . . . . .	43
4.5	Critères de positivité des formes quadratiques . . . . .	45
4.6	Deux applications de la diagonalisation . . . . .	46
<b>5</b>	<b>Espace vectoriel normé</b>	<b>49</b>
5.1	Evn et sa structure "topologique" . . . . .	49
5.2	Norme d'opérateurs (ou norme subordonnée) . . . . .	51
5.3	La fonction exponentielle de matrices . . . . .	53
<b>6</b>	<b>Applications</b>	<b>55</b>
6.1	Coniques . . . . .	55
6.2	Interprétation géométrique des groupes $O(2)$ et $O(3)$ . . . . .	56
6.2.1	Le groupe $O(2)$ . . . . .	56
6.2.2	Le groupe $O(3)$ . . . . .	57
6.3	Une deuxième preuve du théorème de diagonalisation . . . . .	59
6.4	Matrice carrée symétrique positive et rayon spectral . . . . .	60

# 1 Espaces euclidiens et espaces vectoriels normés

Les trois premiers paragraphes de ce chapitre sont composés de rappels. Les preuves y seront donc en général omises.

## 1.1 Rappel : espaces vectoriels

**Définition 1.** Soit  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ , que l'on appellera le "corps de base". On appelle espace vectoriel (souvent abrégé e.v.) sur  $\mathbb{K}$  un ensemble  $E$  doté des lois de composition suivantes :

- une loi interne, c'est-à-dire une opération  $E \times E \rightarrow E$  qui à  $(x, y) \in E \times E$  associe  $x + y$ , dite addition, telle que  $(E, +)$  soit un groupe commutatif, c'est-à-dire que
  - elle est associative :  $\forall x, y, z \in E, (x + y) + z = x + (y + z)$ ,
  - elle est commutative :  $\forall x, y \in E, x + y = y + x$ ,
  - elle admet un élément neutre (noté  $0_E$  ou plus simplement  $0$ ) :  $\forall x \in E, x + 0_E = 0_E + x = x$ ,
  - tout élément admet un inverse (dit opposé) :  $\forall x \in E, \exists y \in E, x + y = y + x = 0_E$  ; cet élément est noté  $-x$ ,
- une loi externe sur  $E$  par les éléments de  $\mathbb{K}$ , c'est-à-dire une opération  $\mathbb{K} \times E \rightarrow E$  qui à  $(\lambda, y) \in \mathbb{K} \times E$  associe  $\lambda \cdot x$ , dite multiplication par un scalaire de  $\mathbb{K}$ , telle que :

$$\begin{aligned}\forall \lambda, \mu \in \mathbb{K}, \forall x \in E, \quad \lambda \cdot (\mu \cdot x) &= (\lambda\mu) \cdot x, \\ \forall \lambda, \mu \in \mathbb{K}, \forall x \in E, \quad (\lambda + \mu) \cdot x &= \lambda \cdot x + \mu \cdot x, \\ \forall x, y \in E, \quad \forall \lambda \in \mathbb{K}, \quad \lambda \cdot (x + y) &= \lambda \cdot x + \lambda \cdot y, \\ \forall x \in E, \quad 1 \cdot x &= x \text{ (où } 1 \text{ est l'élément neutre pour la multiplication dans } \mathbb{K}\text{)}.\end{aligned}$$

On notera dans toute la suite " $\lambda x$ " plutôt que " $\lambda \cdot x$ ". On appellera les éléments de  $\mathbb{K}$  des **scalaires** et les éléments de  $E$  des **vecteurs**.

### Exemples.

- $\mathbb{R}^n$  muni de l'addition et de la multiplication par un scalaire "coordonnée par coordonnée"

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &= (a_1 + b_1, \dots, a_n + b_n), \\ \lambda(a_1, \dots, a_n) &= (\lambda a_1, \dots, \lambda a_n).\end{aligned}$$

est un espace vectoriel sur  $\mathbb{R}$ .

- Le corps des complexes  $E = \mathbb{C}$  est un espace vectoriel réel (sur  $\mathbb{K} = \mathbb{R}$ ) ou complexe (sur  $\mathbb{K} = \mathbb{C}$ ).
- L'espace  $\mathbb{C}^n$  avec les mêmes opérations d'addition et de multiplication que ci-dessus (voir  $\mathbb{R}^n$ ) est un espace vectoriel sur  $\mathbb{R}$  ou sur  $\mathbb{C}$ .

- Soit  $X$  un ensemble quelconque (non vide), soit  $E$  un espace vectoriel sur  $\mathbb{K}$ . On introduit alors  $\mathcal{A}(X; E)$  (noté encore  $X^E$ ) l'ensemble des applications de  $X$  dans  $E$ . Muni des opérations standard d'addition et multiplication par un scalaire de fonctions (on fait les opérations "point par point"), ceci est aussi un espace vectoriel sur  $\mathbb{K}$ .
- L'espace des matrices à  $n$  lignes et  $p$  colonnes  $M_{n,p}(\mathbb{K})$  à entrées dans  $\mathbb{K}$  est un  $\mathbb{K}$ -espace vectoriel.
- L'espace  $\mathbb{K}[X]$  des polynômes à coefficients dans  $\mathbb{K}$  est un  $\mathbb{K}$ -espace vectoriel.

**Proposition 1.** *Pour tout  $\lambda \in \mathbb{K}$  et  $x \in E$ , on a :*

- $\lambda 0 = 0$  ( $0$  est ici l'élément neutre de  $E$  pour l'addition i.e. "le vecteur nul"),
- $0x = 0$  ( $0$  est ici l'élément neutre de  $\mathbb{K}$  pour l'addition i.e. le  $0$  de  $\mathbb{R}$  ou  $\mathbb{C}$ ),
- si  $\lambda x = 0$  alors soit  $\lambda = 0$ , soit  $x = 0$ ,
- $(-\lambda)x = \lambda(-x) = -(\lambda x)$ .

**Démonstration.** Pour la première utiliser  $0 + 0 = 0$  (dans  $E$ ), pour la deuxième utiliser  $0 + 0 = 0$  (dans  $\mathbb{K}$ ), pour la troisième multiplier par  $\lambda^{-1}$  et pour la dernière vérifier les propriétés de l'inverse.  $\square$

**Définition 2.** *Soient  $E$  un espace vectoriel sur  $\mathbb{K}$  et  $F \subset E$  un sous ensemble de  $E$ . On dit que  $F$  est un **sous-espace vectoriel** (souvent abrégé : s.e.v.) si et seulement si c'est une partie non vide et stable par les deux lois de  $E$ . Autrement dit  $F$  est un s.e.v. si et seulement si :*

- $F \neq \emptyset$
- $\forall x, y \in F, \lambda \in \mathbb{K}, x + \lambda y \in F$

Un des grands intérêts de la notion de sous-espace vectoriel est qu'elle permet de montrer facilement qu'un certain ensemble est muni d'une structure d'espace vectoriel. En effet on a le résultat suivant bien connu.

**Proposition 2.** *Soient  $E$  un espace vectoriel et  $F \subset E$  un sous-espace vectoriel de  $E$ . Alors  $F$  muni des lois interne et externe de  $E$  devient lui-même un espace vectoriel sur  $\mathbb{K}$ .*

**Exemples.**

- Une droite passant par l'origine d'un espace vectoriel  $E$  : pour tout  $x \in E \setminus \{0\}$  :  $\{\lambda x, \lambda \in \mathbb{K}\}$  est un sous espace vectoriel de  $E$ .
- Les sous-espaces engendrés (voir plus bas) sont des s.e.v.
- L'espace des fonctions continues de  $\mathbb{R}$  dans  $\mathbb{R}$  est un sous-espace vectoriel de l'espace  $\mathcal{A}(\mathbb{R}; \mathbb{R})$ . C'est donc lui-même un espace vectoriel.
- L'espace  $\mathbb{K}_n[X]$  des polynômes à coefficients dans  $\mathbb{K}$  et de degré au plus  $n$  est un sous-espace vectoriel de  $\mathbb{K}[X]$ .
- En revanche le groupe  $GL(n; \mathbb{K})$  des matrices carrées de taille  $n$  inversibles, n'est pas un sous-espace vectoriel de l'espace  $M_n(\mathbb{K})$  des matrices carrées de taille  $n$ .

**Proposition 3.** Soient  $F$  et  $G$  deux s.e.v de  $E$ . Alors :

- $F \cap G$  est un s.e.v. de  $E$ ,
- $F + G := \{x + y \text{ pour } x \in F \text{ et } y \in G\}$  est un s.e.v. de  $E$ ,
- mais  $F \cup G$  n'est pas en général un s.e.v de  $E$  (faire un dessin dans  $\mathbb{R}^2$  !)

**Définition 3.** On dit que la somme de deux sous-espaces  $F$  et  $G$  d'un espace vectoriel  $E$  est **directe** et l'on note  $F + G = F \oplus G$  lorsque tout  $x$  de  $F + G$  s'écrit de manière unique comme somme d'un élément de  $F$  et d'un élément de  $G$ . Il est équivalent que  $F \cap G = \{0\}$ . Deux sous-espaces  $F$  et  $G$  sont dits **supplémentaires** lorsque leur somme est directe et que  $E = F + G$

**Définition 4.** Soient  $E$  et  $F$  deux  $\mathbb{K}$ -espaces vectoriels. Une application  $T : E \rightarrow F$  est dite **linéaire** (ou  $\mathbb{K}$ -linéaire) si elle satisfait aux relations :

- $\forall x, y \in E, T(x + y) = T(x) + T(y)$ ,
- $\forall \lambda \in \mathbb{K}, \forall x \in E, T(\lambda x) = \lambda T(x)$ .

En abrégé, ces deux conditions peuvent s'écrire sous la forme :

$$\forall \lambda, \mu \in \mathbb{K}, \forall x, y \in E, T(\lambda x + \mu y) = \lambda T(x) + \mu T(y). \quad (1)$$

On notera  $L(E, F)$  l'ensemble des applications linéaires de  $E$  dans  $F$ . C'est un sous-espace vectoriel de  $\mathcal{A}(E, F)$ . Lorsque  $E = F$ , on parle d'**endomorphisme** (ou morphisme d'espaces vectoriels) de  $E$  et on note  $End(E) := L(E, E)$ .

**Exemples.** L'application de  $\mathbb{R}^3$  dans  $\mathbb{R}^2$  qui à  $(x, y, z)$  associe  $(x, y + 2z)$  est linéaire. Pas l'application de  $\mathbb{R}^2$  dans  $\mathbb{R}$  qui à  $(x, y)$  associe  $xy$ .

**Proposition 4.** Soient  $T$  une application linéaire de  $E$  dans  $F$ . Alors les parties

$$\begin{aligned} Ker(T) &:= \{x \in E / T(x) = 0\}, \\ Im(T) &:= \{y \in F / \exists x \in E, y = T(x)\}, \end{aligned}$$

appelées respectivement **noyau** et **image** de  $T$ , sont des sous-espaces vectoriels de  $E$  et  $F$  respectivement.

## 1.2 Rappel : Famille génératrice, libre, bases

**Définition 5.** Soit  $E$  un e.v. sur  $\mathbb{K}$ , et soient  $p$  vecteurs de  $E : x_1, \dots, x_p \in E$  et  $p$  scalaires  $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ . On appelle **combinaison linéaire** des vecteurs  $x_i$  de coefficients  $\lambda_i$  ( $i = 1, \dots, p$ ) le vecteur

$$\lambda_1 x_1 + \dots + \lambda_p x_p.$$

**Définition 6.** Soient  $E$  un e.v. sur  $\mathbb{K}$ , et  $x_1, \dots, x_p \in E$  des vecteurs de  $E$ . On appelle **espace vectoriel engendré** par cette famille et on note  $\text{Vect}\{x_1, \dots, x_p\}$ , l'ensemble de toutes les combinaisons linéaires des vecteurs  $x_1, \dots, x_p$  :

$$\text{Vect}\{x_1, \dots, x_p\} = \left\{ \sum_{i=1}^p \lambda_i x_i, (\lambda_i)_{i=1, \dots, p} \in \mathbb{K}^p \right\}.$$

**Proposition 5.** L'espace vectoriel engendré par la famille  $x_1, \dots, x_p \in E$  est le plus petit s.e.v. de  $E$  contenant les vecteurs  $x_1, \dots, x_p$ .

**Démonstration.** On procède en deux étapes : d'abord on observe que l'espace vectoriel engendré par la famille  $x_1, \dots, x_p$  tel que défini plus haut est bien un s.e.v. de  $E$ ; ensuite qu'il est contenu dans tout s.e.v. de  $E$  qui contient  $x_1, \dots, x_p$ .  $\square$

**Définition 7.** On dit que  $\{x_1, \dots, x_p\}$  est un **système de générateurs** ou encore une **famille génératrice** de  $E$  si tout élément de  $E$  peut se mettre sous la forme d'une combinaison linéaire de  $x_1, \dots, x_p$ , i.e. si

$$\text{Vect}\{x_1, \dots, x_p\} = E.$$

On dira également que  $x_1, \dots, x_p$  engendrent  $E$ .

**Interprétation en termes de système linéaire dans  $\mathbb{R}^n$ .** Si  $x_1, \dots, x_p$  sont des éléments de  $\mathbb{R}^n$ , dire qu'ils forment une famille génératrice signifie que le système linéaire de  $n$  équations à  $p$  inconnues  $\lambda_1, \dots, \lambda_p$  :

$$\lambda_1 x_1 + \dots + \lambda_p x_p = u,$$

admet toujours **au moins** une solution quel que soit  $u$  de  $\mathbb{R}^n$ .

**Exemple.** La famille  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  est un système de générateurs de  $\mathbb{R}^3$  car tout élément  $(a, b, c) \in \mathbb{R}^3$  est combinaison linéaire des éléments de la famille :

$$(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1).$$

**Définition 8.** Soient  $E$  un e.v. sur  $\mathbb{K}$ , et  $x_1, \dots, x_p \in E$  des vecteurs de  $E$ . On dit que la famille  $\{x_1, \dots, x_p\}$  est **liée** (ou que les vecteurs  $x_1, \dots, x_p$  sont **linéairement dépendants**), s'il existe une combinaison linéaire de ces vecteurs, à coefficients non tous nuls qui soit égale à zéro, i.e.

$$\exists \lambda_1, \dots, \lambda_p \in \mathbb{K} \text{ tels que } \sum_{i=1}^p \lambda_i x_i = 0, \text{ et } \exists i, \lambda_i \neq 0.$$

**Exemple.** Soient  $E$  un e.v. et  $x_1, x_2, x_3 \in E$ . Alors la famille  $\{x_1, x_2, x_3, 7x_1 + 2x_2\}$  est linéairement dépendante.

**Définition 9.** Soient  $E$  un e.v. sur  $\mathbb{K}$ , et  $x_1, \dots, x_p \in E$  des vecteurs de  $E$ . On dit que la famille  $\{x_1, \dots, x_p\}$  est **libre** (ou encore que les vecteurs  $x_1, \dots, x_p$  sont **linéairement indépendants**), si cette famille n'est pas liée, autrement dit, si la seule combinaison linéaire des vecteurs  $x_1, \dots, x_p$  qui soit nulle, est celle dont les coefficients sont tous nuls. Ce que l'on peut encore exprimer sous la forme : si  $\sum_{i=1}^p \lambda_i x_i = 0$  alors  $\lambda_1 = \dots = \lambda_p = 0$ .

**Interprétation en termes de système linéaire dans  $\mathbb{R}^n$ .** Si  $x_1, \dots, x_p$  sont des éléments de  $\mathbb{R}^n$ , dire qu'ils forment une famille libre signifie que le système linéaire de  $n$  équations à  $p$  inconnues  $\lambda_1, \dots, \lambda_p$  :

$$\lambda_1 x_1 + \dots + \lambda_p x_p = u,$$

admet toujours **au plus** une solution quel que soit  $u$  de  $\mathbb{R}^n$ . En effet, s'il y en a deux différentes, en faisant la différence entre les deux équations, on trouve une combinaison nulle des vecteurs  $x_1, \dots, x_p$ , à coefficients non tous nuls.

**Définition 10.** Soient  $E$  un e.v. sur  $\mathbb{K}$  et  $X \subset E$  un ensemble de vecteurs de  $E$ . On dit que l'ensemble  $X$  est une **base** de  $E$  s'il forme à la fois une famille libre et génératrice.

**Théorème 1.** Soient  $E$  un e.v. sur  $\mathbb{K}$  et  $X \subset E$  un ensemble de vecteurs de  $E$ . L'ensemble  $X$  est une base de  $E$  si et seulement si tout vecteur de  $E$  s'écrit de façon unique comme combinaison linéaire d'éléments de  $X$ .

**Démonstration.** Condition nécessaire. Si  $X$  est une base et s'il y a deux façons d'écrire un vecteur  $y$  comme combinaisons linéaires des éléments de  $X$ , on fait la différence entre ces deux écritures et on déduit que  $X$  n'est pas libre. Par ailleurs, comme  $X$  est génératrice, on déduit qu'il y a au moins une écriture.

Condition suffisante. Si la propriété de droite est satisfaite, alors  $X$  est de toute évidence d'une famille génératrice; l'unicité de l'écriture donne le fait que la famille est libre (car l'unique écriture de 0 comme combinaison des éléments de  $X$  est la combinaison à coefficients nuls).  $\square$

**Définition 11.** Soient  $E$  un e.v. sur  $\mathbb{K}$ , et  $X = \{x_1, \dots, x_n\}$  une base de  $E$ . Pour tout  $x \in E$  les coefficients  $\lambda_i$  de l'écriture

$$x = \sum_{i=1}^n \lambda_i x_i$$

s'appellent les coordonnées de  $x$  dans la base  $X$ .

**Théorème 2.** Soit  $E$  un e.v. sur  $\mathbb{K}$ . On suppose que  $E$  est engendré par un nombre fini de vecteurs. Alors toutes les bases de  $E$  ont le même nombre d'éléments. Cet nombre s'appelle **la dimension** de l'espace vectoriel  $E$ .

Nous ne refaisons pas la preuve de ce théorème fondamental, mais rappelons qu'un lemme central (parfois appelé lemme de Steinitz) est le suivant.

**Lemme 1.** Soient  $E$  un espace vectoriel sur  $\mathbb{K}$ , et  $\{x_1, \dots, x_p\}$  une famille génératrice de  $E$ . Alors toute famille d'au moins  $p + 1$  vecteurs est liée.

**Exemples.** L'espace  $\mathbb{R}^n$  est de dimension  $n$ . L'espace  $\mathbb{R}_n[X]$  est de dimension  $n + 1$  puisque  $1, X, X^2, \dots, X^n$  en forme une base. L'espace  $\mathbb{R}[X]$  est de dimension infinie, puisqu'il contient une famille infinie de vecteurs libres (les  $X^k$ ).

Terminons ce paragraphe par un théorème très utile en algèbre linéaire et deux de ses conséquences.

**Théorème 3** (de la base incomplète). Soient  $E$  un espace vectoriel sur  $\mathbb{K}$  de dimension finie, et  $A = \{x_1, \dots, x_p\} \subset B = \{x_1, \dots, x_n\}$  (où  $p \leq n$ ) deux familles de vecteurs de  $E$ . On suppose que  $A$  est une famille libre et que  $B$  est une famille génératrice. Alors il existe une base de  $E$  contenant  $A$  et incluse dans  $B$ . Autrement dit, on peut piocher parmi  $\{x_{p+1}, \dots, x_n\}$  pour compléter  $A = \{x_1, \dots, x_p\}$  en une base.

Une conséquence bien connue et importante du théorème de la base incomplète est le suivant.

**Théorème 4** (du rang). Soit  $T \in L(E, F)$ , où  $E$  est de dimension finie. Alors on a :

$$\dim(E) = \dim(\text{Im}(T)) + \dim(\text{Ker}(T)).$$

**Corollaire 1.** Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie. Alors pour  $T \in \text{End}(E)$ , on a :

$$T \text{ injectif} \iff T \text{ surjectif} \iff T \text{ bijectif}.$$

### 1.3 Rappel : espaces vectoriels normés

Soit  $E$  un espace vectoriel sur  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ . Il est fréquemment utile de pouvoir mesurer la "taille" d'un vecteur donné ou la distance entre deux vecteurs de l'espace  $E$ . Pour cela on introduit la notion de **norme**.

**Définition 12.** Une **norme** est une application de  $E$  dans  $\mathbb{R}_+$  qui à tout vecteur  $x \in E$  associe le nombre réel positif noté  $\|x\| \geq 0$  vérifiant :

$$\|x\| = 0 \iff x = 0, \tag{2}$$

$$\forall \alpha \in \mathbb{K}, \forall x \in E : \|\alpha x\| = |\alpha| \|x\|, \tag{3}$$

$$\forall x, y \in E : \|x + y\| \leq \|x\| + \|y\| \text{ (inégalité triangulaire)}. \tag{4}$$

On appelle **espace vectoriel normé** (en abrégé, *e.v.n.*) un espace vectoriel muni d'une norme.

Un espace vectoriel normé est un cas particulier d'espace métrique avec la **distance** entre deux éléments donnée par :

$$d(x, y) = \|x - y\|.$$

On appelle également  $\|x\|$  la longueur du vecteur  $x$ .

**Interprétation géométrique.** Lorsqu'on est dans  $\mathbb{R}^2$  avec la norme

$$\|(x, y)\| = \sqrt{x^2 + y^2},$$

$\|x + y\|$ ,  $\|x\|$  et  $\|y\|$  sont les longueurs des cotés d'un triangle dans  $\mathbb{R}^2$  (mettons  $x = \overrightarrow{AB}$ ,  $y = \overrightarrow{BC}$ , et donc  $x + y = \overrightarrow{AC}$ ), d'où l'appellation d'inégalité triangulaire.

**Remarque 1.** On peut construire beaucoup de normes différentes sur un même espace vectoriel (il y a plusieurs façons possibles de mesurer les tailles de vecteurs et les distances). En ce cas, on peut distinguer les normes en leur ajoutant un indice. Par exemple on pourra noter une norme  $x \mapsto \|x\|_1$ . Par ailleurs, pour alléger la notation, on écrit souvent "la norme  $\|\cdot\|_1$ ".

**Exemples.** On peut munir  $\mathbb{R}^n$  de plusieurs normes différentes comme :

$$\begin{aligned} \|(x_1, \dots, x_n)\|_1 &:= |x_1| + \dots + |x_n|, \\ \|(x_1, \dots, x_n)\|_\infty &:= \max\{|x_1|, \dots, |x_n|\}. \end{aligned}$$

De même, on peut munir l'espace  $C^0([0, 1]; \mathbb{R})$  des fonctions continues sur  $[0, 1]$  à valeurs réelles des normes suivantes :

$$\begin{aligned} \|f\|_1 &:= \int_0^1 |f(x)| dx, \\ \|f\|_\infty &:= \sup_{x \in [0, 1]} |f(x)|. \end{aligned}$$

Le fait que ces deux dernières soient bien définies vient du théorème affirmant qu'une fonction continue sur un segment est bornée.

**Exercice.** Montrer que les applications définies ci-dessus sont des normes.

**Exemple.** On peut aussi définir sur  $\mathbb{R}^n$  la norme  $\|\cdot\|_p$  pour  $p \in [1, \infty[$  :

$$\|(x_1, \dots, x_n)\|_p := (|x_1|^p + \dots + |x_n|^p)^{1/p},$$



et sur  $C^0([0, 1]; \mathbb{R})$  :

$$\|f\|_p := \left( \int_0^1 |f(x)|^p dx \right)^{1/p}.$$

Montrer qu'il s'agit bien de normes pour des  $p$  généraux est faisable mais assez délicat.

Une notion importante associée à celle de norme est celle de l'équivalence de normes.

**Définition 13.** Soit  $E$  un espace vectoriel sur  $\mathbb{K} = \mathbb{R}$ , muni de deux normes  $\|\cdot\|_1$  et  $\|\cdot\|_2$ . On dit que ces deux normes sont équivalentes lorsqu'il existe deux constantes  $c$  et  $C$  **strictement positives** telles que

$$\text{pour tout } x \in E, \quad c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1.$$

Cela veut dire que les deux normes peuvent être différentes, mais qu'elles sont néanmoins "du même ordre". En particulier, si l'une des deux normes indique qu'un vecteur est très "petit" (il a une petite norme), alors l'autre aussi est "petit".

### Remarques.

1. Comme d'habitude en mathématiques, l'ordre des quantificateurs est très important. Il existe  $c > 0$  et  $C > 0$  tels que pour tout  $x \in E$  l'inégalité est satisfaite, cela veut bien sûr dire que  $c$  et  $C$  **ne doivent pas** dépendre de  $x$  ! Et d'ailleurs, s'ils pouvaient en dépendre, la notion serait essentiellement vide car toujours satisfaite.

2. L'équivalence des normes est une *relation d'équivalence*, au sens où cette relation est :

- réflexive : une norme est équivalente à elle-même,
- symétrique : la norme  $\|\cdot\|_1$  est équivalente à la norme  $\|\cdot\|_2$  si et seulement si la norme  $\|\cdot\|_2$  est équivalente à la norme  $\|\cdot\|_1$ ,
- transitive : si la norme  $\|\cdot\|_1$  est équivalente à la norme  $\|\cdot\|_2$  et que la norme  $\|\cdot\|_2$  est équivalente à la norme  $\|\cdot\|_3$ , alors la norme  $\|\cdot\|_1$  est équivalente à la norme  $\|\cdot\|_3$ .

Le théorème fondamental en termes d'équivalences de normes est le suivant. Sa démonstration est assez difficile et requiert des notions qui sortent largement du cadre de ce cours !



**Théorème 5.** Soit  $E$  un espace vectoriel sur  $\mathbb{R}$  ou  $\mathbb{C}$ . Alors :

- si  $E$  est de dimension finie, alors toutes les normes sur  $E$  sont équivalentes,
- si  $E$  est de dimension infinie, alors il existe **toujours** des normes non équivalentes.

Il est essentiel de retenir la dimension finie dans la première partie de cet énoncé !

**Exercice.** Montrer que sur  $\mathbb{R}^n$ , les normes  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  et  $\|\cdot\|_\infty$  définies plus tôt sont équivalentes (sans utiliser le théorème ci-dessus bien sûr !). Montrer que sur  $C^0([0, 1]; \mathbb{R})$ , les normes  $\|\cdot\|_1$  et  $\|\cdot\|_\infty$  ne **sont pas** équivalentes.

## 1.4 Définition d'un espace euclidien

Nous avons vu que dans des espaces normés, on peut obtenir des informations sur la façon dont la norme se comporte par rapport aux deux opérations de l'espace : addition et multiplication par des scalaires. Mais en ce qui concerne l'addition, on dispose seulement d'une inégalité. Dans ce chapitre, nous allons voir que l'on peut donner plus d'informations sur ce sujet pour certaines normes particulières : celles qui dérivent d'un “produit scalaire”, qui est le concept principal de ce chapitre. En outre, le produit scalaire permettra d'introduire d'autres notions importantes comme l'orthogonalité.

Nous allons donc commencer par définir la notion de produit scalaire, puis voir comment à tout produit scalaire on associe une norme particulière. Dans ce cours, nous allons nous restreindre au cas réel mais les notions rencontrées ici peuvent s'étendre sans difficulté au cas d'un espace vectoriel complexe.

**Définition 14.** Soit  $E$  un espace vectoriel sur  $\mathbb{K}$ . Une application  $a : E \times E \rightarrow \mathbb{K}$ ,  $(x, y) \mapsto a(x, y) \in \mathbb{K}$  est dite **bilinéaire**, lorsqu'elle est linéaire par rapport à chacune de ses variables, c'est-à-dire :

- pour  $y \in E$  fixé, l'application  $x \mapsto a(x, y)$  est linéaire,
- pour  $x \in E$  fixé, l'application  $y \mapsto a(x, y)$  est linéaire.

Autrement dit, pour tout  $x, y, z \in E$  et  $\lambda, \mu \in \mathbb{K}$  :

$$\begin{aligned}a(\lambda x + \mu y, z) &= \lambda a(x, z) + \mu a(y, z), \\a(x, \lambda y + \mu z) &= \lambda a(x, y) + \mu a(x, z).\end{aligned}$$

**Définition 15.** Soit  $E$  un espace vectoriel **réel**. Une application  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto \langle x, y \rangle$  est appelée un **produit scalaire** si elle possède les propriétés suivantes :

- elle est bilinéaire,
- elle est symétrique :

$$\forall x, y \in E, \langle x, y \rangle = \langle y, x \rangle,$$

- elle est positive :  $\forall x \in E, \langle x, x \rangle \geq 0$ ,
- elle est “définie” :  $\langle x, x \rangle = 0 \Rightarrow x = 0$ .

**Remarque 2.** On dit aussi “**non dégénérée positive**” à la place de “définie positive”. Voir la remarque 23 beaucoup plus loin !

Il est aussi usuel de noter le produit scalaire, lorsqu'il n'y a pas d'ambiguïté,  $\langle x, y \rangle = x \cdot y$ .

**Exemple.** Dans  $\mathbb{R}^2$ , l'application

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2,$$

définit un produit scalaire. Un autre exemple (exercice) :

$$\langle (x_1, x_2), (y_1, y_2) \rangle = (x_1 + x_2)(y_1 + y_2) + 2x_2 y_2.$$

**Notation.** On notera

$$\|x\| = \sqrt{\langle x, x \rangle},$$

qui est bien défini grâce à la propriété de positivité!

**Lemme 2** (Inégalité de Cauchy-Schwarz). *Dans un espace réel  $E$  muni d'un produit scalaire, on a pour tous  $x, y \in E$ ,*

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

*avec égalité si et seulement si  $x$  et  $y$  sont linéairement dépendants.*

**Démonstration.** On commence d'abord par le cas réel. On considère l'application  $\lambda \mapsto \|x + \lambda y\|^2$  qui doit être positive pour tout  $\lambda \in \mathbb{R}$  :

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

Du fait de cette positivité sur  $\mathbb{R}$ , comme il s'agit d'un polynôme de second ordre, son discriminant doit être négatif, i.e.

$$4|\langle x, y \rangle|^2 \leq 4\|x\|^2 \cdot \|y\|^2,$$

d'où l'inégalité dans le cas réel.

En outre, s'il y a égalité dans l'inégalité, cela signifie que  $|\langle x, y \rangle| = \|x\| \|y\|$ , alors le discriminant précédent est nul, et donc le polynôme a une racine réelle  $\mu$ . On voit bien que cela indique que  $x + \mu y = 0$ .  $\square$

Comme annoncé, à chaque produit scalaire on peut associer une norme particulière.

**Théorème 6.** *Dans un espace réel muni d'un produit scalaire, l'application  $x \mapsto \|x\|$  est une norme, appelée norme **induite par le produit scalaire**  $\langle \cdot, \cdot \rangle$  et aussi appelée norme **euclidienne**.*

**Démonstration.** Il s'agit de montrer que l'application ainsi définie a toutes les propriétés d'un norme : la positivité et les propriétés (2)-(4). Le seul point non évident est l'inégalité triangulaire (4). Pour cela on écrit :

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2, \end{aligned}$$

puis par utilisation de Cauchy-Schwarz on obtient

$$\begin{aligned} \|x + y\|^2 &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

$\square$

**Définition 16.** On appelle espace euclidien tout espace vectoriel **réel, de dimension finie**, muni d'un produit scalaire.

**Note.** Tout espace euclidien est en particulier un espace normé.

**Définition 17.** Deux vecteurs  $x, y \in E$  tels que  $\langle x, y \rangle = 0$  sont dits **orthogonaux**, ce que l'on note  $x \perp y$ . Un vecteur  $x$  est dit orthogonal à un s.e.v.  $F$  et l'on note  $x \perp F$  lorsque pour tout  $y \in F$ , on a  $x \perp y$ . Deux s.e.v.  $F$  et  $G$  de  $E$  sont dits orthogonaux, et l'on note  $F \perp G$  lorsque tout vecteur de  $F$  est orthogonal à tout vecteur de  $G$ .

## 1.5 Exemples d'espaces euclidiens

### 1.5.1 Produit scalaire canonique de $\mathbb{R}^n$

Dans l'espace  $\mathbb{R}^n$ , on introduit le produit scalaire **canonique**, c'est-à-dire celui qu'on utilisera de manière standard dans  $\mathbb{R}^n$  (sauf précision contraire bien entendu). Il s'agit du produit scalaire suivant :

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle_{\mathbb{R}^n} = x_1 y_1 + \dots + x_n y_n.$$

Il arrive parfois qu'on note ce produit scalaire simplement  $\langle \cdot, \cdot \rangle$  lorsque le contexte n'est pas ambigu.

**Remarque 3.** Avec ce choix du produit scalaire la longueur d'un vecteur et l'orthogonalité prennent les sens classiques :

- la norme de  $(x_1, x_2, x_3)$  (ou distance de 0 à  $(x_1, x_2, x_3)$ ) dans  $\mathbb{R}^3$  est  $\sqrt{x_1^2 + x_2^2 + x_3^2}$ ,
- $x \perp y$  est équivalent  $\langle x, y \rangle = 0$  ce qui est équivalent à  $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ .

Pour ajouter une remarque, rappelons une notation.

**Notation.** Dans toute la suite, si  $M \in M_{n,p}(\mathbb{K})$  est une matrice, nous désignerons par  ${}^t M \in M_{p,n}(\mathbb{K})$  sa **transposée** (obtenue par écriture de gauche à droite à la  $j$ -ième ligne des entrées de la  $j$ -ième colonne de  $M$  lues de haut en bas). Rappelons la formule  ${}^t(AB) = {}^t B {}^t A$ , valable dès que le produit matriciel  $AB$  a un sens.

Avec cette notation, notons que si  $U$  et  $V$  sont des vecteurs colonnes de  $\mathbb{R}^n$ , on a :

$$\langle U, V \rangle_{\mathbb{R}^n} = {}^t U V, \tag{5}$$

en identifiant (ce que nous ferons systématiquement) les matrices scalaires (à une ligne et une colonne) et le scalaire qu'elles contiennent.

## 1.5.2 Autres exemples d'espaces munis d'un produit scalaire : les fonctions continues

• Soit  $E = C([a, b]; \mathbb{R})$  le  $\mathbb{R}$ -e.v. des applications continues de  $[a, b]$  dans  $\mathbb{R}$ . On définit le produit scalaire sur  $E$  par

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

• On peut munir l'espace des matrices carrées à entrées réelles d'une structure euclidienne, comme suit. Soit  $E = M_n(\mathbb{R})$  l'espace des matrices  $n \times n$ . On peut définir un produit scalaire sur  $E$  par la formule :

$$\langle A, B \rangle = \text{Tr}({}^tAB).$$

## 1.6 Premières propriétés des espaces euclidiens

Dans cette section, nous donnons quelques propriétés importantes des espaces euclidiens.

### 1.6.1 Identité du parallélogramme

L'identité suivante est en fait valable dans tout espace muni d'un produit scalaire (même s'il est complexe ou de dimension infinie).

**Proposition 6** (Identité du parallélogramme). *Dans tout espace euclidien  $E$  on a,  $\forall x, y \in E$  :*

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

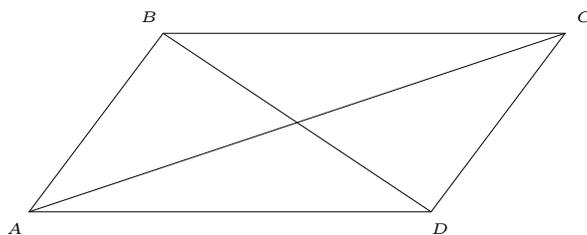


FIGURE 1 – Identité du parallélogramme

Cette égalité est dite *identité du parallélogramme*, car elle permet de voir que dans un parallélogramme, la somme des carrés des longueurs des diagonales est égale à la somme des carrés des longueurs des côtés. Ainsi, dans la Figure 1, on a la relation

$$\|\vec{AC}\|^2 + \|\vec{BD}\|^2 = \|\vec{AB}\|^2 + \|\vec{BC}\|^2 + \|\vec{CD}\|^2 + \|\vec{DA}\|^2 = 2(\|\vec{AB}\|^2 + \|\vec{BC}\|^2).$$

**Démonstration.** Il suffit de développer  $\|x + y\|^2 + \|x - y\|^2 = \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \dots$

□

**Remarque 4.** Certaines normes ne provenant pas d'un produit scalaire ne satisfont pas cette identité. C'est même un moyen simple de déterminer qu'une norme ne provient pas d'un produit scalaire. Exemple : montrer que la norme sur  $\mathbb{R}^2$  :  $\|(x, y)\|_1 := |x| + |y|$  ne provient pas d'un produit scalaire.

Une autre identité utile des espaces euclidiens : on sait déjà déterminer la norme à partir du produit scalaire, mais **pour une norme dont on sait qu'elle provient d'un produit scalaire**, on sait faire le trajet inverse :

**Proposition 7.** Dans un espace euclidien, on a,  $\forall x, y \in E$  :

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

**Démonstration.** Il suffit de développer  $\|x + y\|^2$  à l'aide du produit scalaire. □

### 1.6.2 Angle non-orienté des deux vecteurs

Soient  $u$  et  $v$  deux vecteurs non nuls d'un espace euclidien  $E$ . On sait d'après l'inégalité de Cauchy-Schwarz que  $\left| \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} \right| \leq 1$ . On peut donc trouver un angle  $\alpha \in [0, \pi]$  tel que

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

**Définition 18.** Cet  $\alpha$  est appelé **angle non orienté** entre  $u$  et  $v$ .

**Note.** On retrouve bien la notion usuelle d'angle non orienté du plan euclidien.

### 1.6.3 Théorème de Pythagore

**Théorème 7.** Soit  $E$  un espace euclidien. Alors :

- on a l'équivalence

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow \langle x, y \rangle = 0,$$

- si  $x_1, \dots, x_n$  sont des vecteurs orthogonaux deux à deux, on a :

$$\|x_1 + \dots + x_n\|^2 = \sum_{i=1}^n \|x_i\|^2. \tag{6}$$

**Démonstration.** Développer les carrés à l'aide du produit scalaire ! □

**Note.** À partir de trois vecteurs, la réciproque n'est pas toujours vraie (l'identité (6) n'implique pas que les vecteurs soient deux à deux orthogonaux). Par exemple, pour  $(1, 0)$ ,  $(0, 1)$ ,  $(1, -1)$  :

$$\begin{aligned} \|(1, 0) + (0, 1) + (1, -1)\|^2 &= \|(2, 0)\|^2 = 4, \\ \|(1, 0)\|^2 + \|(0, 1)\|^2 + \|(1, -1)\|^2 &= 1 + 1 + 2 = 4, \end{aligned}$$

donc les deux quantités sont égales, mais les vecteurs ne sont pas deux à deux orthogonaux :

$$\langle (0, 1), (1, -1) \rangle = 0 \cdot 1 + 1 \cdot (-1) = -1 \neq 0.$$

#### 1.6.4 Orthogonalité et liberté

**Proposition 8.** *Soit  $E$  un espace euclidien. Soit  $X = \{x_1, \dots, x_n\}$  une famille de vecteurs, deux à deux orthogonaux et non nuls. Alors  $X$  est une famille libre de  $E$ .*

**Démonstration.** Si une combinaison linéaire des  $x_i$  est nulle, on en prend le produit scalaire par  $x_k$  (quel que soit  $k$ ). En utilisant les relations d'orthogonalité (et le fait que  $x_k$  soit non nul), on voit que le coefficient de  $x_k$  de cette combinaison est nul.  $\square$

### 1.7 Bases orthonormées, procédé d'orthonormalisation de Schmidt

#### 1.7.1 Motivation

Soient  $E$  un espace euclidien et  $X = \{e_1, \dots, e_n\}$  une base de  $E$ . Pour  $x = \sum_{k=1}^n \alpha_k e_k$  et  $y = \sum_{j=1}^n \beta_j e_j$ , le calcul de  $\langle x, y \rangle$  donne :

$$\langle x, y \rangle = \left\langle \sum_{k=1}^n \alpha_k e_k, \sum_{j=1}^n \beta_j e_j \right\rangle = \sum_{j,k=1}^n \alpha_k \beta_j \langle e_k, e_j \rangle. \quad (7)$$

Ce calcul n'est en général pas aisé car il faut prendre en compte tous les produits  $\langle e_k, e_j \rangle$ . Mais toutes les bases ne sont donc pas identiques de ce point de vue. En particulier, celles qui ont beaucoup de facteurs  $\langle e_k, e_j \rangle$  nuls vont permettre un calcul plus rapide. Il faut néanmoins remarquer que quelle que soit la base, on aura  $\langle e_k, e_k \rangle = \|e_k\|^2 \neq 0$ .

#### 1.7.2 Bases orthonormées

**Définition 19.** *Soient  $E$  un espace euclidien et  $X = \{e_1, \dots, e_n\}$  une base de  $E$ . La base  $X$  est dite **orthogonale** lorsque*

$$\langle e_i, e_j \rangle = 0, \text{ pour tout } i \neq j.$$

La base  $X$  est dite **orthonormée** lorsque

$$\langle e_i, e_j \rangle = \delta_{ij} \text{ pour tout } i, j \leq n.$$

**Remarque 5.** *Dans une base orthonormée, le calcul (7) devient très simple :*

$$\langle x, y \rangle = \left\langle \sum_{k=1}^n \alpha_k e_k, \sum_{j=1}^n \beta_j e_j \right\rangle = \sum_{j,k=1}^n \alpha_k \beta_j \langle e_k, e_j \rangle = \sum_{k=1}^n \alpha_k \beta_k. \quad (8)$$

Ainsi, si on raisonne sur les coordonnées dans cette base, on est ramené au cas de  $\mathbb{R}^n$  avec le produit euclidien usuel. Un énoncé précis donnant cette équivalence d'un espace euclidien à  $\mathbb{R}^n$  muni du produit scalaire canonique est donné plus bas (voir Théorème 9).

**Proposition 9.** Soient  $E$  un espace euclidien et  $\{e_1, \dots, e_n\}$  une base orthonormée de  $E$ . Alors

$$\forall x \in E, \quad x = \sum_{i=1}^n \langle x, e_i \rangle e_i \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^n |\langle x, e_i \rangle|^2.$$

**Démonstration.** Il suffit d'écrire  $x$  comme combinaison linéaire des  $e_i$ , et de considérer les produits scalaires avec les  $e_i$  pour identifier les coefficients. Ensuite, c'est le théorème de Pythagore.  $\square$

**Remarque 6.** Donnons un exemple de construction d'une base orthonormée dans  $\mathbb{R}^2$ , comme une "déformation" d'une base donnée.

À partir de  $v_1 = (2, 0)$  et  $v_2 = (2, 1)$  on peut chercher une base  $\{e_1, e_2\}$  orthonormée en "normalisant" le premier :  $e_1 = \frac{v_1}{\|v_1\|} = (1, 0)$  et en cherchant le second comme une combinaison linéaire des deux :  $e_2 = \alpha e_1 + \beta v_2$ . On obtient  $e_2 = (0, 1)$ .

### 1.7.3 Procédé d'orthonormalisation de Schmidt

Un procédé pour construire une base orthonormée à partir d'une base quelconque et qui généralise la méthode de la Remarque 6 est le procédé d'orthonormalisation de Schmidt (aussi appelé procédé de Gram-Schmidt) décrit dans le théorème et la preuve suivants.

**Théorème 8.** Soient  $E$  un espace euclidien de dimension  $n$  et  $X = \{x_1, \dots, x_n\}$  une base de  $E$ . Alors il existe une unique base orthonormée  $X_S = \{e_1, \dots, e_n\}$  telle que pour  $k = 1, \dots, n$ ,

$$\begin{aligned} \text{Vect}\{e_1, \dots, e_k\} &= \text{Vect}\{x_1, \dots, x_k\}, \\ \langle x_k, e_k \rangle &> 0. \end{aligned}$$

**Corollaire 2.** Tout espace euclidien admet une base orthonormée.

**Démonstration.** On définit par récurrence les vecteurs suivants :

$$\begin{aligned} e_1 &= \frac{x_1}{\|x_1\|}, \\ e_2 &= \frac{y_2}{\|y_2\|} := \frac{x_2 - \langle x_2, e_1 \rangle e_1}{\|x_2 - \langle x_2, e_1 \rangle e_1\|}, \\ &\vdots \\ e_n &= \frac{y_n}{\|y_n\|} := \frac{x_n - \sum_{j=1}^{n-1} \langle x_n, e_j \rangle e_j}{\|x_n - \sum_{j=1}^{n-1} \langle x_n, e_j \rangle e_j\|}. \end{aligned}$$

On montre maintenant (et toujours par récurrence) différentes propriétés.

- Chaque  $y_k$  appartient à  $\text{Vect}\{x_1, \dots, x_k\}$  : cela est clair dans les formules ci-dessus, par récurrence sur  $k$ .

- Pour tout  $k$ ,  $y_k \neq 0$  : cela vient du fait que  $X$  est une base et du fait que, par hypothèse de récurrence, pour  $i = 1, \dots, k-1$ ,  $e_i \in \text{Vect}\{x_1, \dots, x_{k-1}\}$ . Cela permet en particulier de voir que  $e_k$  est bien défini (pas de division par 0 dans sa définition).
- Chaque  $x_k$  est inclus dans  $\text{Vect}\{e_1, \dots, e_k\}$  : cela vient encore des formules ci-dessus.
- Les  $e_j$ ,  $j \leq n$  sont orthogonaux deux à deux : il suffit de montrer que  $e_k$  est orthogonal à tous les  $e_j$  pour  $j = 1, \dots, k-1$ ; pour s'en convaincre on multiplie la formule précédente de  $e_k$  par  $e_j$  avec  $j < k$ .
- Enfin, pour tout  $k$ , on a  $\langle x_k, e_k \rangle > 0$  car

$$0 < \|e_k\|^2 = \langle e_k, e_k \rangle = \left\langle \frac{x_k - \sum_{j=1}^{k-1} \langle x_k, e_j \rangle e_j}{\|x_k - \sum_{j=1}^{k-1} \langle x_k, e_j \rangle e_j\|}, e_k \right\rangle = \frac{\langle x_k, e_k \rangle}{\|x_k - \sum_{j=1}^{k-1} \langle x_k, e_j \rangle e_j\|}.$$

Pour ce qui est de l'unicité : supposons avoir une base orthonormée  $(e_1, \dots, e_n)$  satisfaisant les conditions précédentes. On commence par voir que  $e_1$  est nécessairement défini comme précédemment : il appartient à  $\text{Vect}\{x_1\}$  et est donc de la forme  $\lambda x_1$ ; ensuite  $\|e_1\| = 1$  et  $\langle e_1, x_1 \rangle > 0$  montrent que nécessairement  $\lambda = 1/\|x_1\|$ . Supposons ensuite que nous avons prouvé que  $e_1, \dots, e_k$  avaient nécessairement la forme précédente; prouvons que  $e_{k+1}$  a lui aussi cette forme. D'après les hypothèses, on sait que  $e_{k+1} \in \text{Vect}\{x_1, \dots, x_{k+1}\} = \text{Vect}\{e_1, \dots, e_k, x_{k+1}\}$ . On écrit donc  $e_{k+1}$  sous la forme :

$$e_{k+1} = \sum_{i=1}^k \alpha_i e_i + \lambda x_{k+1}.$$

On prend le produit scalaire avec  $e_i$  pour  $i = 1, \dots, k$  et on trouve  $\alpha_i = -\lambda \langle x_{k+1}, e_i \rangle$ , donc

$$e_{k+1} = \lambda \left( x_{k+1} - \sum_{i=1}^k \langle x_{k+1}, e_i \rangle e_i \right).$$

Enfin les conditions  $\|e_{k+1}\| = 1$  et  $\langle e_{k+1}, x_{k+1} \rangle > 0$  déterminent  $\lambda$  et montrent que  $e_{k+1}$  a bien la forme précédente.  $\square$

## 2 Espaces euclidiens, dualité et applications linéaires

### 2.1 Rappel : matrice d'application linéaire

**Définition 20.** Soient  $E$  et  $F$  deux espaces vectoriels de dimension finie, de base respective  $\{e_1, \dots, e_n\}$  et  $\{f_1, \dots, f_m\}$ . Soit  $T : E \rightarrow F$  une application linéaire. Pour chaque  $j$  entre 1 et  $n$ , le vecteur  $T(e_j)$  de  $F$  se décompose dans la base  $\{f_1, \dots, f_m\}$  selon

$$T(e_j) = \sum_{i=1}^m a_{ij} f_i.$$

On appelle **matrice de  $T$  dans les bases  $\{e_1, \dots, e_n\}$  et  $\{f_1, \dots, f_m\}$** , le tableau à  $m$  lignes et  $n$  colonnes, dont l'entrée (c'est-à-dire l'élément) à la  $j$ -ème colonne et  $i$ -ème ligne est  $a_{ij}$ . On la note aussi parfois  $(a_{ij})_{i=1..m, j=1..n}$ .

**Convention.** Quand on désigne l'entrée  $a_{ij}$  d'une matrice, le premier indice  $i$  correspond à la ligne, le second  $j$  à la colonne. On note  $M_{d,n}(\mathbb{K})$  l'espace des matrices à  $d$  lignes et  $n$  colonnes, dont les entrées sont des éléments de  $\mathbb{K}$ .

Autrement dit, dans la  $k$ -ième colonne ( $k$  allant de 1 à  $n$ ), on range les  $d$  coordonnées de  $T(e_k)$  dans la base  $\{f_1, \dots, f_d\}$ .

**Remarque 7.** Si  $E = F$  et que  $T$  est donc un endomorphisme, on choisit souvent la même base  $\{e_1, \dots, e_n\}$  "au départ" et à l'arrivée. Dans ce cas, on parle de matrice de l'endomorphisme  $T$  dans la base  $\{e_1, \dots, e_n\}$ . Il s'agit bien sûr, dans ce cas, d'une matrice « carrée », c'est-à-dire ayant le même nombre de lignes et de colonnes.

**Exemple 1.** Soit  $E := \mathbb{R}_1[X]$  l'espace des polynômes de degré inférieur ou égal à 1. On vérifie sans peine que  $\{X, X - 1\}$  est une base de  $E$ . On considère alors l'endomorphisme  $T$  de  $E$  qui à  $P$  associe le polynôme  $Q(X) := P(X + 1)$ . On voit sans peine que  $T(X) = X + 1 = 2.X - 1.(X - 1)$  tandis que  $T(X - 1) = X = 1.X + 0.(X - 1)$ . La matrice de  $T$  dans la base  $\{X, X - 1\}$  est donc

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

**Définition 21.** Soit  $A = (a_{ij}) \in M_{p,q}(\mathbb{K})$  et  $B = (b_{jk}) \in M_{q,r}(\mathbb{K})$ . On définit le **produit matriciel** de  $A$  et  $B$ , dans cet ordre, comme la matrice  $C \in M_{p,r}(\mathbb{K})$ , dont les coefficients  $c_{ik}$  pour  $i = 1 \dots p$  et  $k = 1 \dots r$  sont

$$c_{ik} = \sum_{j=1}^q a_{ij} b_{jk}.$$

**Proposition 10.** Soient  $E, F$  et  $G$  trois espaces vectoriels de base respective  $\{e_1, \dots, e_\ell\}$ ,  $\{f_1, \dots, f_m\}$  et  $\{g_1, \dots, g_n\}$ . Soient  $S \in L(E, F)$  et  $T \in L(F, G)$ , de matrice respective  $A$  et  $B$  dans ces bases. Alors la matrice de  $T \circ S \in L(E, G)$  dans ces bases est la matrice produit  $B \cdot A$ .

**Démonstration.** Les décompositions de  $S(e_i)$  et  $T(f_j)$  s'écrivent

$$S(e_i) = \sum_{j=1}^m a_{ji} f_j \quad \text{et} \quad T(f_j) = \sum_{k=1}^n b_{kj} g_k$$

et on en déduit bien

$$T \circ S(e_i) = \sum_{j=1}^m \sum_{k=1}^n a_{ji} b_{kj} g_k = \sum_{k=1}^n (BA)_{ki} g_k,$$

pour tout  $1 \leq i \leq \ell$ . □

**Définition 22.** Soit  $E$  un espace vectoriel de dimension finie, et soient  $\{e_1, \dots, e_n\}$  et  $\{f_1, \dots, f_n\}$  deux bases de  $E$ . On appelle **matrice de passage de la base  $\{e_1, \dots, e_n\}$  vers la base  $\{f_1, \dots, f_n\}$** , la matrice carrée de taille  $n$ , dont la  $k$ -ième colonne contient les coordonnées du vecteur  $f_k$  dans la base  $\{e_1, \dots, e_n\}$ .

**Remarque 8.** On retiendra qu'on exprime les « nouveaux » vecteurs dans « l'ancienne base ».

**Exemple 2.** Soit  $E := \mathbb{R}_1[X]$  l'espace des polynômes de degré inférieur ou égal à 1. On vérifie que  $\{1, X\}$  et  $\{2X, 3X - 1\}$  sont deux bases de  $E$ . On voit que la matrice de passage de  $\{1, X\}$  vers  $\{2X, 3X - 1\}$  est

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

**Proposition 11.** Soit  $E$  un espace vectoriel de dimension finie, et soient  $\{e_1, \dots, e_n\}$  et  $\{f_1, \dots, f_n\}$  deux bases de  $E$ . Soit  $x$  un vecteur  $E$ , de coordonnées  $(x_1, \dots, x_n)$  dans la base  $\{e_1, \dots, e_n\}$  et  $(y_1, \dots, y_n)$  dans la base  $\{f_1, \dots, f_n\}$ . Soit  $P$  la matrice de passage de  $\{e_1, \dots, e_n\}$  vers  $\{f_1, \dots, f_n\}$ . Alors on a :

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

**Remarque 9.** Le produit matriciel naturel donne donc les anciennes coordonnées en fonction des nouvelles.

**Proposition 12.** Soit  $E$  un espace vectoriel de dimension finie, soient  $\{e_1, \dots, e_n\}$  et  $\{f_1, \dots, f_n\}$  deux bases de  $E$ , et soit  $P$  la matrice de passage de  $\{e_1, \dots, e_n\}$  vers  $\{f_1, \dots, f_n\}$ . Soit  $T$  un endomorphisme de  $E$ , de matrice  $A$  dans la base  $\{e_1, \dots, e_n\}$ . Alors sa matrice  $\tilde{A}$  dans la base  $\{f_1, \dots, f_n\}$  est

$$\tilde{A} = P^{-1}AP.$$

## 2.2 Structure des espaces euclidiens par les bases orthonormées. Equivalence avec $\mathbb{R}^n$ .

**Définition 23.** Soient  $E$  et  $F$  deux espaces euclidiens. On appelle **isométrie vectorielle** entre  $E$  et  $F$  (ou encore, **isomorphisme d'espaces euclidiens**), une bijection linéaire  $f : E \rightarrow F$  qui préserve le produit scalaire, c'est-à-dire telle que :

$$\forall u, v \in E, \langle f(u), f(v) \rangle_F = \langle u, v \rangle_E.$$

**Théorème 9.** Soit  $E$  un espace euclidien de dimension  $n$ . Alors  $E$  est en isométrie avec  $\mathbb{R}^n$ , c'est-à-dire qu'il existe une isométrie vectorielle entre  $E$  et  $\mathbb{R}^n$ .

**Démonstration.** On dénote par  $\{e_1, \dots, e_n\}$  une base orthonormée de  $E$ . On sait que tout  $u \in E$  s'écrit d'une manière unique

$$u = \sum_{i=1}^n U^i e_i, \tag{9}$$

puisque  $E$  est une base. On définit l'application

$$f : \begin{cases} \mathbb{R}^n & \longrightarrow E \\ U = (U^i)_{i=1}^n & \longmapsto f(U) = \sum_{i=1}^n U^i e_i. \end{cases}$$

En notant  $(\varepsilon_i)_{1 \leq i \leq n}$  la base canonique de  $\mathbb{R}^n$ , on observe que  $f(\varepsilon_i) = e_i$  pour tout  $1 \leq i \leq n$ . Cela implique bien que  $f$  est une bijection entre deux e.v. de dimension  $n$ .

Montrons que  $f$  est linéaire et préserve le produit scalaire. Soient  $U = (U^i)_{i=1}^n \in \mathbb{R}^n$  et  $V = (V^i)_{i=1}^n \in \mathbb{R}^n$ . La linéarité se montre de manière standard :

$$\begin{aligned} f(U + V) &= f((U^i)_{i=1}^n + (V^i)_{i=1}^n) = f((U^i + V^i)_{i=1}^n) \\ &= \sum_{i=1}^n (U^i + V^i) e_i = \sum_{i=1}^n U^i e_i + \sum_{i=1}^n V^i e_i = f(U) + f(V), \end{aligned}$$

$$f(\lambda U) = f((\lambda U^i)_{i=1}^n) = \sum_{i=1}^n (\lambda U^i) e_i = \lambda \sum_{i=1}^n U^i e_i = \lambda f(U).$$

Le caractère isométrique est donné par le calcul suivant :

$$\langle U, V \rangle_{\mathbb{R}^n} = \sum_{i=1}^n U^i V^i = \left\langle \sum_{i=1}^n U^i e_i, \sum_{i=1}^n V^i e_i \right\rangle = \langle f(U), f(V) \rangle_E.$$

Par conséquent l'application  $f$  qui à  $(U^i)_{i=1}^n \in \mathbb{R}^n$  associe  $u \in E$  est un isomorphisme d'espaces euclidiens.  $\square$

**Remarque 10.** L'isomorphisme est **différent** pour chaque base orthonormée  $\{e_1, \dots, e_n\}$  choisie.

## 2.3 Projections

### 2.3.1 Définition

**Définition 24.** Soient  $E$  un espace euclidien et  $A \subset E$  une partie non vide de  $E$ . On note

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

et on l'appelle l'**orthogonal** de  $A$ .

**Proposition 13.** L'orthogonal  $A^\perp$  de tout partie non vide  $A$  est un s.e.v de  $E$ . On a également  $A \perp B$  implique  $A \subset B^\perp$  pour deux sous-ensembles (non vides)  $A, B \subset E$ . En outre on a :

$$\{0\}^\perp = E, \quad E^\perp = \{0\}.$$

**Démonstration.** La propriété de s.e.v. et d'inclusion se vérifient trivialement. De même, tout vecteur est orthogonal à zéro. Pour voir que  $E^\perp = \{0\}$ , il suffit de remarquer que si  $v \in E^\perp$ , alors  $\langle v, v \rangle = 0$ .  $\square$

**Proposition 14.** Soit  $F$  un s.e.v. d'un espace euclidien  $E$ . Alors :

- $E = F \oplus F^\perp$ ,
- $F^{\perp\perp} = F$ .

Soient  $F_1$  et  $F_2$  deux s.e.v. d'un espace euclidien  $E$ . Si  $F_1 \oplus F_2 = E$  et  $F_1 \perp F_2$ , alors  $F_1 = F_2^\perp$ .

**Démonstration.** Pour le premier point : on considère une base de  $F$ ; on étend ensuite celle-ci à une base de  $E$  (on utilise ici le Théorème 3 de la base incomplète). Ensuite on orthonormalise (par le procédé de Schmidt) et on obtient donc une base orthonormée de  $E$  qui contient une base orthonormée de  $F$ , mettons  $\{e_1, \dots, e_p\}$  où les  $p$  premiers vecteurs forment une base de  $F$ . Il n'est alors pas difficile de voir que

$$F^\perp = \text{Vect} \{e_{p+1}, \dots, e_n\},$$

en utilisant la formule (8). On conclut ainsi que  $E = F \oplus F^\perp$  et que  $\dim(E) = \dim(F) + \dim(F^\perp)$ .

Pour le deuxième point, on a tout d'abord  $\dim(F^{\perp\perp}) = \dim(E) - \dim(F^\perp) = \dim(F)$  et par ailleurs on vérifie facilement que  $F \subset F^{\perp\perp}$ . On a donc l'égalité.

Pour la dernière partie, on raisonne de la même manière. On a d'une part  $\dim(F_2) = \dim(E) - \dim(F_1) = \dim(F_1^\perp)$ . Comme d'autre part  $F_2 \subset F_1^\perp$ , cette inclusion est en fait une égalité.  $\square$

**Définition 25.** Le théorème précédent nous montre que tout  $x \in E$  s'écrit d'une manière unique comme  $x = u + v$  avec  $u \in F$  et  $v \in F^\perp$ . On appelle  $u$  la **projection orthogonale** de  $x$  sur  $F$  et on note par  $p_F$  l'application  $x \mapsto p_F(x) = u$  ainsi construite.

**Exemple.** Dans  $E = \mathbb{R}^3$  la projection de  $(3, 2, 4)$  sur le plan  $x = 0$  est  $(0, 2, 4)$  alors que sa projection sur  $x + y + z = 0$  est  $(0, -1, 1)$ .

**Proposition 15.** Soit  $F$  un s.e.v. d'un espace euclidien  $E$ . Alors

1. Orthogonalité : pour tout  $x \in E$  on a :  $\forall y \in F, x - p_F(x) \perp y$  et  $x = p_F(x) + p_{F^\perp}(x)$ .  
De plus  $\|p_F(x)\| \leq \|x\|$ .
2. si  $\{e_1, \dots, e_k\}$  est une base orthonormée de  $F$  on a

$$\forall x \in E, p_F(x) = \sum_{i=1}^k \langle x, e_i \rangle e_i. \quad (10)$$

3. Si  $p$  est une projection orthogonale  $p \circ p = p$ .
4. Si  $p$  est un endomorphisme vérifiant  $p \circ p = p$  et si en plus  $Im(p) \perp Ker(p)$  alors  $p$  est la projection orthogonale sur  $Im(p)$ .
5. Caractérisation par la propriété de minimalité :

$$\forall x \in E, \|x - p_F(x)\| = \inf_{w \in F} \|x - w\|.$$

**Démonstration.** Le 1. est une conséquence simple de l'écriture  $x = u + v$  et de  $F = F^{\perp\perp}$ . L'inégalité  $\|p_F(x)\| \leq \|x\|$  découle de Pythagore.

Pour le 2., on écrit la projection comme une combinaison linéaire dans cette base, ensuite les coefficients se déterminent en utilisant  $x - p_F(x) \perp F$ , donc pour  $i = 1 \dots k$ ,  $\langle p_F(x), e_i \rangle = \langle p_F(x) - x, e_i \rangle$  :

$$p_F(x) = \sum_{i=1}^k \langle p_F(x), e_i \rangle e_i = \sum_{i=1}^k \langle x, e_i \rangle e_i + \langle p_F(x) - x, e_i \rangle e_i.$$

Pour le 3., on utilise que pour  $x \in F$ ,  $p_F(x) = x$  (ce qui est clair puisque l'écriture  $x = x + 0$  avec  $x \in F$  et  $0 \in F^\perp$  est unique).

Pour le 4., on démontre d'abord que  $Im(p) \oplus Ker(p) = E$  : tout  $x \in E$  s'écrit  $x = p(x) + x - p(x)$  et on voit que  $p(x - p(x)) = 0$  donc  $x - p(x) \in Ker(p)$ . De plus l'écriture est unique puisque si  $x = a + b$  avec  $b \in Ker(p)$  alors  $p(x) = p(a) = a$  (une autre manière de faire était de voir que les espaces  $Im(p)$  et  $Ker(p)$  sont en somme directe (du fait de leur orthogonalité) et d'utiliser le théorème du rang). On conclut par la définition de la projection orthogonale.

Pour le 5., on écrit  $\|x - w\|^2 = \|x - p_F(x) + p_F(x) - w\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - w\|^2$ , la dernière égalité venant de la propriété d'orthogonalité.  $\square$

**Remarque 11.** On reconnaît dans la formule (10) une écriture que nous avons vue dans le procédé d'orthonormalisation de Schmidt. On voit que la base obtenue par orthonormalisation peut également s'écrire :

$$e_k = \frac{x_k - P_{F_{k-1}}(x_k)}{\|x_k - P_{F_{k-1}}(x_k)\|}, \text{ où } F_{k-1} := Vect\{e_1, \dots, e_{k-1}\} = Vect\{x_1, \dots, x_{k-1}\}.$$

**Proposition 16.** Soit  $E$  un espace euclidien muni d'une base orthonormée  $(v_1, \dots, v_n)$ . Soient  $F$  un s.e.v. de  $E$  et  $(e_1, \dots, e_k)$  une base orthonormée de  $F$ . Notons  $U$  la matrice des coordonnées de  $e_1, \dots, e_k$  en colonne dans la base  $(v_1, \dots, v_n)$ , (qui est donc une matrice à  $n$  lignes et  $k$  colonnes). Soit  $x \in E$ , et soient  $X$  et  $Y$  les matrices coordonnées respectives de  $x$  et  $p_F(x)$  dans la base  $(v_1, \dots, v_n)$ . Alors on a :

$$Y = U^t U X.$$

**Démonstration.** Ce n'est en fait rien d'autre que la formule (10) :  ${}^t U X$  est le vecteur colonne (à  $k$  lignes) :

$$\begin{pmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_k \rangle \end{pmatrix},$$

et on conclut aisément. □

### 2.3.2 Application au problème des moindres carrés

Soient  $m$  points de  $\mathbb{R}$  notés  $x_i$ ,  $i = 1, \dots, m$  et  $c_i$  des valeurs associées à ces points. On s'intéresse au problème d'approcher cette ensemble de points par le graphe d'une fonction élémentaire, c'est-à-dire de trouver une fonction simple telle que  $U(x_i) \simeq c_i$ . Un bon choix peut être constitué par les fonctions affines de la forme  $U(x) = ax + b$ . Autrement dit, nous allons chercher à approcher un "nuage de points" par une droite. C'est le classique problème de la "régression linéaire", tel que décrit sur la Figure 2.

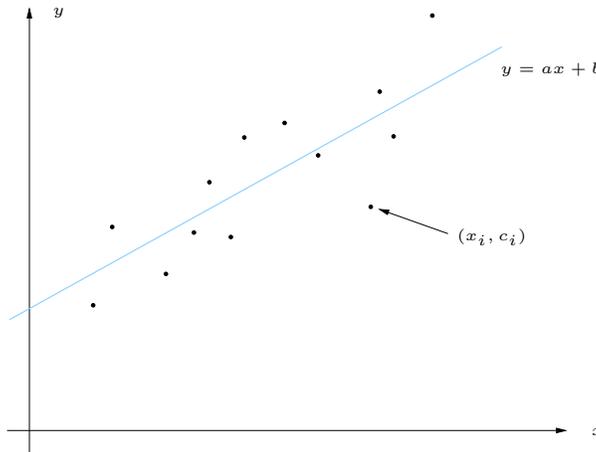


FIGURE 2 – Droite de régression linéaire

On peut formuler notre problème comme la **minimisation** de

$$S(a, b) = \sum_{i=1}^m (ax_i + b - c_i)^2.$$

Soit  $c$  le vecteur colonne d'entrées  $c_i$ , autrement dit,  $c = {}^t(c_1, \dots, c_m)$ . On considère aussi  $x = {}^t(x_1, \dots, x_m)$ ,  $u = {}^t(1, \dots, 1) \in \mathbb{R}^m$ . On observe alors que

$$S(a, b) = \|ax + bu - c\|^2.$$

On peut donc reformuler le problème comme suit : trouver la combinaison linéaire des vecteurs  $x$  et  $u$  (de coefficients  $a$  et  $b$ ) à distance minimale du vecteur  $c$ . Soit donc l'espace vectoriel  $V$  engendré par  $x$  et  $u$  dans  $\mathbb{R}^m$ . D'après ce qui précède, nous cherchons donc la **projection orthogonale** de  $c$  sur  $V$  :

$$ax + bu = p_V(c).$$

Pour effectuer le calcul on utilise la propriété d'orthogonalité :

$$(ax + bu - c) \perp x \quad \text{et} \quad (ax + bu - c) \perp u.$$

Cela donne un système de deux équations à deux inconnues à résoudre.

Ceci se généralise de plusieurs façons :

- On peut chercher la meilleure approximation parmi des fonctions non affines. Par exemple, soient  $x \mapsto g_1(x)$  et  $x \mapsto g_2(x)$ . On peut chercher la meilleure approximation comme combinaison de ces deux fonctions

$$ag_1(x_i) + bg_2(x_i) \simeq c_i,$$

au sens des moindres carrés, c'est-à-dire qui minimise

$$\sum_{i=1}^m |ag_1(x_i) + bg_2(x_i) - c_i|^2.$$

Le cas précédent correspond à  $g_1 : x \mapsto x$  et à  $g_2$  la fonction constante à 1. On introduit alors les vecteurs

$$G_1 = {}^t(g_1(x_1), \dots, g_1(x_m)) \quad \text{et} \quad G_2 = {}^t(g_2(x_1), \dots, g_2(x_m)).$$

On cherche donc le vecteur au sein de  $\text{Vect}\{G_1, G_2\}$  qui soit à distance minimale du vecteur  $c$ . Il s'agit donc encore du projeté orthogonal de  $c$  sur le sous-espace  $\text{Vect}\{G_1, G_2\}$ .

- On peut utiliser plus de deux fonctions, par exemple  $g_1, \dots, g_k$ , et chercher la meilleure approximation sous la forme

$$\alpha_1 g_1(x_i) + \dots + \alpha_k g_k(x_i) \simeq c_i.$$

On introduira alors comme précédemment les vecteurs colonnes  $G_1, \dots, G_k$  donnés par

$$G_j = \begin{pmatrix} g_j(x_1) \\ \vdots \\ g_j(x_m) \end{pmatrix} \text{ pour } j = 1, \dots, k,$$

et la meilleure approximation au sens des moindres carrés sera le projeté orthogonal de  $c$  sur le sous-espace  $\text{Vect} \{G_1, \dots, G_k\}$ .

- On peut considérer le cas d'une dépendance de plusieurs variables, par exemple quand on a un nuage tridimensionnel  $(x_i, y_i, c_i)_{i=1\dots m}$  et qu'on cherche à approcher  $c$  par une fonction des deux variables  $x, y$ . Dans le cas linéaire pour simplifier, on cherche donc la meilleure approximation

$$\alpha x_i + \beta y_i + b \simeq c_i,$$

ce qui donne au sens des moindres carrés

$$\sum_{i=1}^m |\alpha x_i + \beta y_i + b - c_i|^2.$$

On cherche encore le projeté orthogonal de  $c$  sur le sous-espace de  $\mathbb{R}^m$  :

$$\text{Vect} \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

On obtient dans ce cas précis trois équations à trois inconnues.

## 2.4 Dualité

### 2.4.1 Rappel : dual d'un espace vectoriel

**Définition 26.** Soit  $E$  un e.v. sur  $\mathbb{K}$ . On appelle *dual* de  $E$  et on note par  $E^*$  ou par  $E'$  l'ensemble des applications linéaires sur  $E$  à valeurs dans  $\mathbb{K}$ . Autrement dit,  $E^* := L(E, \mathbb{K})$ .

**Remarque 12.** On appelle **forme linéaire** une application linéaire dont l'espace d'arrivée est  $\mathbb{K}$ . On fera attention à distinguer cette expression du cas général où l'on parle « d'application » linéaire.

**Exemples.** Soit  $E = \mathbb{R}^3$ ; alors  $\mu(x, y, z) = x - y$  est une forme linéaire sur  $E$  donc  $\mu \in E^*$ . De même pour  $\nu(x, y, z) = x + y - 2z$ . En revanche,  $\ell(x, y, z) = x + 1$  et  $m(x, y, z) = x + |z|$  ne sont pas des formes linéaires.

**Remarque 13.** Pour  $f \in E^*$  et  $x \in E$  on notera également  $f(x)$  par  $\langle f, x \rangle_{E^*, E}$  ou encore par  $\langle x, f \rangle_{E, E^*}$ .

**Théorème 10.** Soit  $E$  un e.v. sur  $\mathbb{K}$ . Alors :

- le dual  $E^*$  est aussi un e.v. sur  $\mathbb{K}$ .
- si de plus  $E$  est **de dimension finie** alors  $E^*$  est aussi de dimension finie ; mieux,  $\dim(E^*) = \dim(E)$ .

**Démonstration.**

Pour le premier point, il suffit d'observer que pour tous  $f, g \in E^*$ ,  $\alpha, \beta \in \mathbb{K}$  la fonction  $\alpha f + \beta g$  définie sur  $E$  à valeurs en  $\mathbb{K}$  est encore linéaire et à valeurs dans  $\mathbb{K}$  ; donc  $\alpha f + \beta g \in E^*$ .

Pour le second point, soit  $\{e_i\}_{i=1}^n$  une base de  $E$  ; on définit les applications  $\mu_i : E \rightarrow \mathbb{K}$  de la manière suivante :

$$\mu_i \left( \sum_{k=1}^n x_k e_k \right) = x_i.$$

On vérifie qu'il s'agit bien de formes linéaires. De plus, on a que  $\mu_i(e_j) = \delta_{ij}$ , pour tous  $i, j = 1, \dots, n$ . (Rappelons que symbole de Kronecker  $\delta_{ij}$  vaut 0 si  $i \neq j$  et 1 sinon.)

Il suffit maintenant de montrer que la famille  $\{\mu_i\}_{i=1}^n$  forme une base de  $E^*$ . D'une part, du fait des valeurs des  $\mu_j$  sur les vecteurs  $e_j$ , elles forment une famille libre. D'autre part n'importe quelle forme linéaire  $f \in E^*$  s'écrit comme combinaison linéaire des  $\mu_i$  :

$$f = \sum_{i=1}^n f(e_i) \mu_i.$$

En effet, on sait en effet que les valeurs d'une application linéaire est déterminée par ses valeurs sur une base ; les deux formes linéaires coïncident ici sur les  $(e_i)$ .  $\square$

**Définition 27.** La base  $(\mu_1, \dots, \mu_n)$  de  $E^*$  ainsi déterminée est appelée **base duale** de la base  $(e_j)$ . Elle est parfois notée  $(e_1^*, \dots, e_n^*)$ .

**Définition 28.** Soient  $E$  un e.v. de dimension finie sur  $\mathbb{K}$  et  $F$  un s.e.v. de  $E$ . L'**annulateur**  $F^0$  de  $F$  est l'ensemble des formes linéaires  $f \in E^*$  s'annulant sur  $F$  :

$$F^0 = \{f \in E^* / \forall x \in F, f(x) = 0\}.$$

**Remarque 14.** L'annulateur  $F^0$  de  $F$  est un s.e.v. de  $E^*$  : en effet, toute combinaison linéaire  $\alpha f + \beta g$  avec  $f, g \in F^0$  s'annule sur  $F$ .

**Exemples.** Soit  $E = \mathbb{R}^3$  ;  $\mu(x, y, z) = x - y$  et  $\nu(x, y, z) = x + y - 2z$  sont dans l'annulateur de  $F = \text{Vect}\{(1, 1, 1)\}$  vu comme s.e.v. de  $E = \mathbb{R}^3$ .

**Théorème 11.** Soient  $E$  un e.v. de dimension finie sur  $\mathbb{K}$  et  $F$  un s.e.v. de  $E$ . Alors

$$\dim(F^0) = \dim(E) - \dim(F).$$

**Démonstration.** On introduit une base de  $F$ , mettons  $(e_1, \dots, e_k)$  (avec  $k = \dim(F)$ ). On complète cette base de  $F$  à une base de  $E$ , mettons que nous obtenons  $(e_1, \dots, e_n)$  (avec  $n = \dim(E)$ ). Maintenant, on introduit la base duale  $(e_1^*, \dots, e_n^*)$  comme décrit précédemment. On voit alors que

$$F^0 = \text{Vect} \{e_{k+1}^*, \dots, e_n^*\}.$$

En effet, il est clair que  $e_{k+1}^*, \dots, e_n^*$  sont dans  $F^0$ . Par ailleurs, un élément  $\ell$  de  $F^0$  ne peut avoir de composante selon  $e_1^*, \dots, e_k^*$ , comme on le voit en regardant  $\ell(e_1) = \dots = \ell(e_k) = 0$ . Donc  $\dim(F^0) = n - k$ .  $\square$

## 2.4.2 Structure du dual d'un espace euclidien

Soit  $E$  un espace euclidien et soit  $E^*$  le dual de  $E$  (i.e., rappelons-nous, l'espace des formes linéaires sur  $E$ ).

Pour  $E = \mathbb{R}^3$ , nous avons vu comme exemples d'éléments de  $E^*$  les applications  $\mu(x, y, z) = x + 2y - z$ , et  $\nu(x, y, z) = x - z$ . Plus généralement, toute application  $\lambda_{\alpha, \beta, \gamma}(x, y, z) = \alpha x + \beta y + \gamma z$  ( $\alpha, \beta, \gamma \in \mathbb{R}$ ) est linéaire donc  $\lambda_{\alpha, \beta, \gamma} \in E^*$ . Remarquons que l'on peut écrire  $\lambda_{\alpha, \beta, \gamma}(x, y, z) = \langle (\alpha, \beta, \gamma), (x, y, z) \rangle$ .

Plus généralement, pour tout  $e \in E$ , l'application  $d_e : E \rightarrow \mathbb{R}$  définie par  $d_e(v) = \langle v, e \rangle$  est une forme linéaire sur  $E$  donc  $d_e \in E^*$ . Nous sommes ainsi conduits à nous demander si la réciproque est également vraie, c'est-à-dire, si toutes les applications  $f \in E^*$  sont de la forme précédente. Autrement dit, pour tout  $f$  de  $E^*$ , existe-t-il  $w$  dans  $E$  tel que pour tout  $v$  de  $E$  on ait  $f(v) = \langle v, w \rangle$ ? La réponse est donnée par le résultat suivant.

**Théorème 12.** *L'application  $d : E \rightarrow E^*$  qui à tout  $v$  de  $E$  associe l'élément  $d(v)$  de  $E^*$  (qu'on notera  $d_v$ ) défini par*

$$\forall u \in E, d_v(u) = \langle u, v \rangle,$$

*est un isomorphisme (i.e. une bijection linéaire) de  $E$  sur  $E^*$ . De plus si  $F$  est un s.e.v. de  $E$  et  $F^0$  est l'annulateur de  $F$  (voir la Définition 28 page 25) on a*

$$d(F^\perp) = F^0.$$

**Démonstration.** On sait d'après le Théorème 10 (page 25) que  $E$  et  $E^*$  ont la même dimension. Comme la linéarité de  $d$  est évidente, il suffit juste de montrer l'injectivité. Mais si jamais  $d(v) = 0$  il suit qu'en particulier  $d_v(v) = \langle v, v \rangle = 0$  d'où l'injectivité.

Une autre façon de faire est de considérer une base  $\{e_1, \dots, e_n\}$  de  $E$  et pour  $f \in E^*$ ,  $v = \sum_{i=1}^n v_i e_i$  constater que  $f(v) = \sum_{i=1}^n v_i f(e_i)$  donc  $f(v) = \langle v, w \rangle$  avec  $w = \sum_{i=1}^n f(e_i) e_i \in E$  (un peu comme dans la preuve du Théorème 10).

Pour le dernier point, on écrit

$$\begin{aligned} d^{-1}(F^0) &= \{y \in E \mid d_y \in F^0\} \\ &= \{y \in E \mid d_y(v) = 0, \forall v \in F\}, \\ &= \{y \in E \mid \langle v, y \rangle = 0, \forall v \in F\} \\ &= F^\perp. \end{aligned}$$

Comme  $d$  est une bijection il suit que  $d(F^\perp) = F^0$ . □

## 2.5 Changement de base orthonormée et groupe orthogonal

### 2.5.1 Groupe orthogonal

Commençons par un lemme qui nous sera utile ici, mais également plus tard dans le cours.

**Lemme 3** (Produit scalaire et transposée). *Soit  $M \in M_n(\mathbb{R})$ . Pour tous  $U, V \in \mathbb{R}^n$  des vecteurs colonnes,*

$$\langle MU, V \rangle_{\mathbb{R}^n} = \langle U, {}^t M V \rangle_{\mathbb{R}^n}.$$

**Démonstration.** C'est un simple calcul. On voit que

$$\langle MU, V \rangle_{\mathbb{R}^n} = {}^t(MU)V = {}^tU {}^t M V = \langle U, {}^t M V \rangle_{\mathbb{R}^n},$$

en utilisant (5). □

Soit maintenant  $\{e_1, \dots, e_n\}$  une base orthonormée de l'espace euclidien  $\mathbb{R}^n$  (par exemple, la base canonique) et soit également  $M \in M_n(\mathbb{R})$  une matrice carrée. On pose  $v_i = Me_i$ . La question qui nous intéresse est la suivante : quelles sont les propriétés de  $M$  qui assurent que  $\{v_1, \dots, v_n\}$  soit une base orthonormée de  $E$  ?

Or il est facile de voir que pour une matrice carrée  $A = (A_{ij})_{i,j=1}^n$ , on a  $\langle Ae_i, e_j \rangle = A_{ji}$ . Nous avons donc démontré le résultat suivant.

**Proposition 17.** *La matrice  $M$  définit un changement de bases orthonormées si et seulement si*

$${}^t M M = I_n, \tag{11}$$

*ou, d'une manière équivalente,*

$$M^{-1} = {}^t M.$$

**Démonstration.** On calcule

$$\langle v_i, v_j \rangle = \langle Me_i, Me_j \rangle = \langle {}^t M Me_i, e_j \rangle.$$

Or pour une matrice  $A \in M_n(\mathbb{R})$  et  $u \in \mathbb{R}^n$ , on a par définition

$$(Au)_j = \sum_k A_{jk} u_k,$$

ce qui implique pour la base canonique  $(e_i)_{1 \leq i \leq n}$  que  $(Ae_i)_j = A_{ji}$ , et donc

$$\langle Ae_i, e_j \rangle = \left\langle \sum_{k=1}^n (Ae_i)_k e_k, e_j \right\rangle = \sum_{k=1}^n A_{ki} \langle e_k, e_j \rangle = A_{ji}.$$

On en déduit que pour que  $(v_i)$  satisfasse la condition d'orthonormalité, il faut que

$$\delta_{ij} = \langle v_i, v_j \rangle = \langle {}^t M M e_i, e_j \rangle = ({}^t M M)_{ji},$$

ce qui est bien dire que  $I_n = {}^t M M$ . □

**Définition 29.** Une matrice  $M \in M_n(\mathbb{R})$  qui satisfait (11) est dite **orthogonale**. L'ensemble des matrices orthogonales de dimension  $n$  est noté  $O(n)$  et est appelé **groupe orthogonal**.

**Remarque 15.** Une matrice  $M$  est orthogonale si et seulement si elle transforme la base canonique de  $\mathbb{R}^n$  (qui est orthonormée pour le produit scalaire canonique) en une base orthonormée de  $\mathbb{R}^n$ , donc si et seulement si ses vecteurs colonnes sont normés et deux à deux orthogonaux pour le produit scalaire usuel.

**Proposition 18.**  $O(n)$  est un groupe, c'est-à-dire :

- $\forall A, B \in O(n), AB \in O(n)$ ,
- $I_n \in O(n)$ ,
- $\forall A \in O(n), A^{-1} \in O(n)$ .

**Démonstration.** On peut le montrer soit par un calcul direct en utilisant (11), soit en utilisant les propriétés des applications orthogonales (l'isométrie vectorielle), que nous définissons dans le paragraphe suivant. □

**Définition 30.** L'ensemble des matrices orthogonales de déterminant égal à 1 est appelé **groupe spécial orthogonal** et est noté  $SO(n)$ . Autrement dit :

$$SO(n) := \{M \in O(n) / \det(M) = 1\}.$$

### 2.5.2 Endomorphismes orthogonaux

On se pose la même question que dans le paragraphe précédent, mais cette fois dans le cadre d'un espace euclidien quelconque (que nous n'avons pas identifié à  $\mathbb{R}^n$  muni du produit scalaire usuel).

**Théorème 13.** Soient  $E$  un espace euclidien et  $f \in \text{End}(E)$ . Les propriétés suivantes sont équivalentes :

1.  $f$  conserve le produit scalaire, i.e.

$$\forall x, y \in E, \quad \langle f(x), f(y) \rangle = \langle x, y \rangle. \tag{12}$$

2.  $f$  conserve la norme, i.e.

$$\forall x \in E, \quad \|f(x)\| = \|x\|.$$

3. si  $\{e_1, \dots, e_n\}$  est une base orthonormée de  $E$  alors la matrice  $M$  de  $f$  dans cette base est orthogonale.

Si l'une de ces conditions est satisfaite, on dit que l'endomorphisme  $f$  est **orthogonal** (ou, rappelons-nous, une isométrie vectorielle). De plus si  $f$  est orthogonal alors  $\det(M) = \pm 1$  et  $f$  est une bijection.

**Remarque 16.** Notons à l'aide de ce théorème que si une application envoie une base orthonormée particulière sur une base orthonormée, il envoie toute base orthonormée sur une base orthonormée (puisqu'il satisfait (12)).

**Démonstration.**  $1 \Rightarrow 2$  : ceci est immédiat en prenant  $y = x$ .

$2 \Rightarrow 1$  : ceci s'obtient de la relation de calcul de  $\langle f(x), f(y) \rangle$  :

$$\langle f(x), f(y) \rangle = \frac{\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2}{2}$$

$1 \Leftrightarrow 3$  : si  $X$  et  $Y$  sont les vecteurs colonnes des coordonnées de  $x$  et  $y$  de  $\mathbb{R}^n$  dans la base  $(e_1, \dots, e_n)$  alors  $f(x)$  (respectivement  $f(y)$ ) a comme coordonnées  $MX$  (resp.  $MY$ ) dans cette base. Comme la base  $(e_1, \dots, e_n)$  est orthonormée, il suit que :

$$\langle f(x), f(y) \rangle = {}^t(MX)MY = {}^tX({}^tMM)Y.$$

$$\langle x, y \rangle = {}^tXY,$$

donc (12) est équivalent à  ${}^tMM = I_n$ , c'est-à-dire  $M \in O(n)$ .

Si  ${}^tMM = I_n$  il résulte que  $1 = \det I_n = \det({}^tMM) = \det({}^tM) \det(M) = \det(M)^2$  donc  $\det(M) = \pm 1$ . En particulier  $\det(M) \neq 0$  donc  $f$  est bijective (ce qu'on savait puisque  ${}^tMM = I_n$ ).  $\square$

On vient d'utiliser le lemme élémentaire suivant ;

**Lemme 4.** Soient  $E$  un espace euclidien et  $f, g \in \text{End}(E)$ . On a  $f = g$  si, et seulement si,

$$\langle f(x), y \rangle = \langle g(x), y \rangle, \quad \forall x, y \in E.$$

Soient  $A, B \in M_n(\mathbb{R})$ . On a  $A = B$  si, et seulement si,

$$\langle AU, V \rangle = \langle BU, V \rangle, \quad \forall U, V \in \mathbb{R}^n.$$

**Démonstration.** Par linéarité, il suffit de montrer la première équivalence lorsque  $g = 0$ . Si maintenant

$$\langle f(x), y \rangle = 0, \quad \forall x, y \in E,$$

on en déduit, en prenant  $y = f(x)$ , que

$$\|f(x)\|^2 = 0, \quad \forall x \in E,$$

et donc bien  $f = 0$ . Les autres implications sont triviales ou leurs preuves sont semblables.  $\square$

## 3 Formes bilinéaires et quadratiques

### 3.1 Motivation

Lorsqu'on optimise une fonctionnelle  $F : E \rightarrow \mathbb{R}$  (c'est-à-dire qu'on en recherche selon les cas un maximum ou un minimum), il s'agit d'habitude d'applications plus complexes que des applications linéaires ou affines (au demeurant celles-ci n'ont pas de minimum en général ; par exemple  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = ax + b$  a  $-\infty$  comme infimum).

L'exemple le plus simple de fonctions après les applications affines est alors donnée par les polynômes de degré au plus 2 (mais de plusieurs variables). De plus, celui-ci se rencontre souvent dans les applications, voir par exemple le paragraphe 2.3.2. On considère donc  $E = \mathbb{R}^n$  et la fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  : pour  $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ ,

$$f(x_1, \dots, x_n) = \alpha + \sum_{i=1}^n \alpha_i x_i + \sum_{i,j=1}^n \beta_{ij} x_i x_j.$$

Une question naturelle est de déterminer s'il existe un minimum et d'expliquer comment le trouver. On sait lorsqu'il n'y a qu'une variable que si  $X_0$  est un minimum de  $f$  alors la dérivée de  $f$  est nulle au point  $X_0$ . Cela reste vrai à plusieurs variables : les dérivées partielles de  $f$  obtenues en gelant  $(n-1)$  coordonnées et en dérivant par rapport à la variable restante sont nulles.

En changeant la variable  $X$  par  $Y = X - X_0$ , le développement de Taylor à l'ordre 2 (à plusieurs variables) montre que :

$$\begin{aligned} f(X) &= f(X_0) + f'(X_0)(X - X_0) + Q(X - X_0, X - X_0) \\ &= f(X_0) + Q(Y, Y), \end{aligned}$$

où  $Q$  est une "forme quadratique", c'est-à-dire un polynôme homogène de degré 2 :

$$Q(Y, Y) = \sum_{i,j=1}^n \beta_{ij} y_i y_j.$$

La question de la minimalité est donc résolue par l'analyse de la forme quadratique  $Q$ . En particulier, des questions importantes associées à cette analyse sont les suivantes :

- la positivité de  $Q$  : (si elle a lieu alors  $X_0$  est un minimum dans l'exemple ci-dessus),
- l'existence d'une base où cette forme s'écrit plus simplement.

Dans une autre ligne d'idées, l'étude des formes bilinéaires permettra également de mieux comprendre à quoi ressemble un produit scalaire général.

### 3.2 Formes bilinéaires

Rappelons la définition d'une forme bilinéaire (voir page 9).

**Définition 31.** Soit  $E$  un espace vectoriel sur  $\mathbb{K} = \mathbb{R}$ . Une **forme bilinéaire** sur  $E$  est une application  $f : E \times E \rightarrow \mathbb{K}$ , linéaire par rapport à chacune de ses variables.

**Exemple.** Le produit scalaire dans un espace euclidien est une forme bilinéaire. Mais c'est loin d'être le seul exemple, on peut ainsi considérer dans  $\mathbb{R}^2$  la forme bilinéaire  $f((x_1, x_2), (y_1, y_2)) = -x_1y_1 - 2x_1y_2 + x_2y_2$  (mais pas  $g((x_1, x_2), (y_1, y_2)) = -x_1^2 - 2y_1y_2 + x_1y_2!$ )

Soient maintenant un espace vectoriel réel  $E$  et  $\{e_1, \dots, e_n\}$  une base de  $E$ . Alors pour  $u = \sum_{i=1}^n U^i e_i$  et  $v = \sum_{i=1}^n V^i e_i$  on aura

$$f(u, v) = f\left(\sum_{i=1}^n U^i e_i, \sum_{i=1}^n V^i e_i\right) = \sum_{i,j=1}^n U^i V^j f(e_i, e_j). \quad (13)$$

**Définition 32.** On appelle **matrice de la forme bilinéaire**  $f$  dans la base  $\{e_1, \dots, e_n\}$ , la matrice carrée  $A$  de taille  $n$ , d'entrées  $A_{ij} = f(e_i, e_j)$  (c'est-à-dire la matrice ayant  $f(e_i, e_j)$  en ligne  $i$  et colonne  $j$ ).

**Exemples.** Le produit scalaire exprimé dans une base orthonormée a comme matrice la matrice identité. La matrice dans la base canonique de  $\mathbb{R}^2$  de la forme bilinéaire donnée par  $f((x_1, x_2), (y_1, y_2)) = -x_1y_1 - 2x_1y_2 + x_2y_2$  est la suivante

$$A = \begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix}.$$

**Proposition 19.** Pour  $u = \sum_{i=1}^n U^i e_i$  et  $v = \sum_{i=1}^n V^i e_i$ , on note par  $U$  le vecteur colonne  $(U^i)_{i=1}^n$  et  $V$  le vecteur colonne  $(V^i)_{i=1}^n$ . On a alors :

$$f(u, v) = {}^t U A V = \langle U, A V \rangle_{\mathbb{R}^n} = \langle {}^t A U, V \rangle_{\mathbb{R}^n}. \quad (14)$$

**Remarque 17.** Ci-dessus, la matrice  ${}^t U$  est une matrice ligne, la matrice  $V$  est une matrice colonne, la matrice  $A$  est une matrice carrée. Donc le produit matriciel  ${}^t U A V = {}^t U (A V)$  est le produit d'une matrice ligne et d'une matrice colonne (dans cet ordre), donc c'est une matrice scalaire (réduite à une entrée). Cela explique pourquoi on peut avoir l'égalité avec  $f(u, v)$  qui est un élément de  $\mathbb{K}$ . En outre, le produit scalaire  $\langle U, A V \rangle_{\mathbb{R}^n}$  est bien défini, puisqu'il s'opère entre deux vecteurs de  $\mathbb{R}^n$  (écrits sous forme colonne).

**Démonstration.** C'est un simple calcul :

$$f(u, v) = \sum_{i,j=1}^n U^i V^j f(e_i, e_j) = \sum_{i=1}^n U^i \sum_{j=1}^n V^j A_{ij} = \sum_{i=1}^n U^i (A V)_i = {}^t U A V,$$

et les deux dernières égalités ont déjà été vues précédemment. □

On rappelle également la définition de la symétrie d'une forme bilinéaire.

**Définition 33.** La forme bilinéaire  $f : E \times E \rightarrow \mathbb{R}$  est dite **symétrique** lorsqu'elle satisfait

$$\forall u, v \in E, f(u, v) = f(v, u).$$

**Exemples.** La forme bilinéaire sur  $\mathbb{R}^2$  donnée par  $g((x_1, x_2), (y_1, y_2)) = -x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2$  est symétrique. La forme bilinéaire  $f((x_1, x_2), (y_1, y_2)) = -x_1y_1 - 2x_1y_2 + x_2y_2$  ne l'est pas.

**Définition 34.** Une matrice carrée  $A \in M_n(\mathbb{K})$  est dite **symétrique** lorsque  ${}^tA = A$ , c'est-à-dire lorsque  $A_{ij} = A_{ji}$ , pour tous  $i, j = 1, \dots, n$ .

**Exemples.** La matrice  $\begin{pmatrix} 1 & 6 \\ 6 & 0 \end{pmatrix}$  est symétrique, pas la matrice  $\begin{pmatrix} 1 & 6 \\ 7 & 0 \end{pmatrix}$ .

On a alors l'équivalence suivante.

**Proposition 20.** Soit  $f$  une forme bilinéaire sur un espace euclidien  $E$  et soit  $A$  sa matrice dans une base  $(e_i)_{i=1}^n$  de  $E$ . Alors  $f$  est symétrique si et seulement si  $A$  l'est.

**Démonstration.** De  $f(e_i, e_j) = A_{ij}$ , on déduit que  $A$  est symétrique si, et seulement si,  $f(e_i, e_j) = f(e_j, e_i)$  pour tout  $i, j$ . Evidemment, si  $f$  est symétrique alors  $f(e_i, e_j) = f(e_j, e_i)$  pour tout  $i, j$ . Réciproquement, si  $f(e_i, e_j) = f(e_j, e_i)$  pour tout  $i, j$ , on calcule

$$f(u, v) = \sum_{ij} U^i U^j f(e_i, e_j) = \sum_{ij} U^i U^j f(e_j, e_i) = f(v, u),$$

où on a exprimé  $u, v$  dans la base  $(e_i)_{1 \leq i \leq n}$ . □

**Proposition 21.** Soit  $f$  une forme bilinéaire symétrique sur  $E$  un espace réel, on a pour  $u = \sum_{i=1}^n U^i e_i$  et  $v = \sum_{i=1}^n V^i e_i$  :

$$f(u, v) = \langle AU, V \rangle_{\mathbb{R}^n} = \langle U, AV \rangle_{\mathbb{R}^n}.$$

**Démonstration.** Cela vient de (14). □

### 3.3 Formes quadratiques

**Définition 35.** Soit  $f : E \times E \rightarrow \mathbb{K}$  une forme bilinéaire sur un espace vectoriel  $E$ . On appelle **forme quadratique** associée à  $f$  l'application  $s : E \rightarrow \mathbb{K}$  donnée par

$$s(u) = f(u, u). \tag{15}$$

On appelle forme quadratique (tout court), une application  $s : E \rightarrow \mathbb{K}$  telle qu'il existe  $f$  une forme bilinéaire pour laquelle la formule (15) est valable.

**Proposition 22.** Soit  $(e_i)_{i=1}^n$  une base de  $E$ , et soit  $A$  la matrice de la forme bilinéaire  $f$  dans cette base. Pour  $u = \sum_{i=1}^n U^i e_i$ , on note par  $U$  le vecteur colonne  $(U^i)_{i=1}^n$ ; on a alors :

$$s(u) = {}^t U A U = \langle U, A U \rangle_{\mathbb{R}^n}.$$

En particulier  $s$  est un polynôme homogène de degré 2 en les  $U^i$  :

$$s(u) = \sum_{i,j=1}^n U^i U^j A_{ij},$$

d'où le nom de “quadratique”.

**Remarque 18.** Des formes bilinéaires différentes peuvent donner la même forme quadratique! Par exemple, la forme bilinéaire  $g((x_1, x_2), (y_1, y_2)) = x_1 y_1 - 2x_1 y_2 + 2x_2 y_1 + x_2 y_2$  et la forme bilinéaire  $h((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2$  donnent la même forme quadratique  $s(x_1, x_2) = x_1^2 + x_2^2$ .

**Proposition 23.** Soit  $s$  une forme quadratique. Parmi toutes les formes bilinéaires dont la forme quadratique associée est  $s$ , il en existe une unique qui soit symétrique. Celle-ci est donnée par

$$f(u, v) = \frac{s(u+v) - s(u) - s(v)}{2}. \quad (16)$$

**Démonstration.** Si  $s$  est associée à une forme bilinéaire symétrique  $f$ , alors on a

$$s(u+v) = f(u+v, u+v) = f(u, u) + f(u, v) + f(v, u) + f(v, v) = s(u) + 2f(u, v) + s(v).$$

Donc  $f$  est nécessairement donnée par (16), il y a donc au plus une forme bilinéaire symétrique qui fonctionne.

Réciproquement, soit  $s$  une forme quadratique, vérifions que  $f$  donnée par (16) est bien bilinéaire symétrique et que  $s$  est sa forme quadratique associée. Comme  $s$  est une forme quadratique, par définition, elle est associée à une certaine forme bilinéaire  $g$  (non nécessairement symétrique). Le même calcul que précédemment appliqué à  $g$  montre que :

$$\frac{g(u, v) + g(v, u)}{2} = \frac{s(u+v) - s(u) - s(v)}{2} = f(u, v).$$

Il est facile de voir sur cette formule que  $f$  est bien bilinéaire symétrique et que  $s$  est sa forme quadratique associée.  $\square$

**Définition 36.** On appellera **matrice d'une forme quadratique**  $Q$ , la matrice de la forme bilinéaire symétrique qui lui est associée.

**Proposition 24.** *Au niveau matriciel, si une forme quadratique  $q$  donnée dans une base  $(e_i)$  par*

$$q\left(\sum_i X_i e_i\right) = \sum_{\substack{i,k=1,\dots,n \\ i \leq k}} q_{ik} X_i X_k,$$

*alors la matrice  $A$  de la forme bilinéaire symétrique associée est  $A_{ik} = q_{ii}$  si  $i = k$  et  $A_{ik} = q_{ik}/2$  sinon.*

**Démonstration.** Il suffit d'observer que la matrice  $A$  en question est symétrique, et qu'elle donne bien la forme linéaire  $q$  : le calcul est direct (le " $q_{ik}/2$ " vient de ce que nous avons regroupé le terme en  $X_i X_k$  avec le terme en  $X_k X_i$  dans l'expression ci-dessus).  $\square$

**Remarque.** On appelle également *forme quadratique* une expression de la forme

$$Q(X_1, \dots, X_n) = \sum_{i,k=1,\dots,n; i \leq k} q_{ik} X_i X_k.$$

Elle est donc associée à une *forme quadratique  $q$  définie sur un espace vectoriel  $E$*  grâce à la Proposition 24.

**Exemple.** Si  $q(X, Y) = X^2 + XY - Y^2$ . La matrice  $A$  doit satisfaire  $q(X, Y) = (X, Y)A^t(X, Y)$  donc  $A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & -1 \end{pmatrix}$ .

### 3.4 Méthode de Gauss

Commençons ce paragraphe par une définition concernant les formes quadratiques et bilinéaires **réelles**.

**Définition 37.** *Soit un espace vectoriel **réel**  $E$ . Une forme quadratique  $s : E \rightarrow \mathbb{R}$ , est dite **positive** si pour tout  $u \in E$ ,  $s(u) \geq 0$ . Une forme bilinéaire  $f : E \times E \rightarrow \mathbb{R}$ , est dite **positive** si la forme quadratique associée est positive, c'est-à-dire si  $f(u, u) \geq 0$  pour tout  $u \in E$ .*

La question que nous soulevons dans ce paragraphe est la suivante. Étant donnée une forme quadratique, peut-on l'exprimer sous une forme plus simple, qui en particulier permette de déterminer si elle est positive ou non ?

**Rappel.** Un petit rappel avant de commencer : la **forme canonique** d'un polynôme de degré 2 à une variable est la suivante :

$$aX^2 + bX + c = a\left(X + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right).$$

**Premiers exemples.** Pour introduire la méthode de Gauss, commençons par quelques exemples de formes quadratiques à deux variables  $X$  et  $Y$  :

- $Q_1(X, Y) = X^2 + 3Y^2$  est clairement positive.
- $Q_2(X, Y) = 3X^2 - Y^2$  n'est clairement pas positive (prendre  $(X, Y) = (0, 1)$ ).
- $Q_3(X, Y) = (X + Y)^2 + 3(X - Y)^2$  est positive.
- $Q_4(X, Y) = 3(X + Y)^2 - (X - Y)^2$  n'est pas positive (prendre  $(X, Y) = (-1, 1)$ , ou, de manière plus générale, un couple  $(X, Y)$  qui annule  $X + Y$  sans annuler  $X - Y$ ).

Remarquons maintenant qu'en développant on obtient

$$Q_3(X, Y) = 4X^2 - 4XY + 4Y^2 \text{ et } Q_4(X, Y) = 2X^2 + 8XY + 2Y^2. \quad (17)$$

Sous ces formes-là, il est beaucoup moins facile de voir que  $Q_3$  est positive et pas  $Q_4$ !

D'où la question naturelle : trouver une méthode qui permette de remonter de la forme (17) à une forme telle que décrite plus haut. Un algorithme général permettant de faire ce travail est appelé *méthode (ou réduction) de Gauss*.

Décrivons-là sur les exemples du dessus : partons de la formule (17). On fait les opérations suivantes.

- On regroupe les termes en  $X$ . S'il n'y a pas de terme en  $X^2$ , on travaillera plutôt avec  $Y$  en premier. (S'il n'y a pas non plus de terme en  $Y^2$ , alors c'est plus compliqué ; voir plus bas).

Dans le cas de (17), c'est déjà fait, on écrit juste

$$Q_3(X, Y) = [4X^2 - 4XY] + 4Y^2 \text{ et } Q_4(X, Y) = [2X^2 + 8XY] + 2Y^2.$$

- Ensuite, on met *sous forme canonique* en  $X$  (comme pour les polynômes de degré 2 en une variable!), on obtient :

$$Q_3(X, Y) = [4(X - Y/2)^2 - Y^2] + 4Y^2 \text{ et } Q_4(X, Y) = [2(X + 2Y)^2 - 8Y^2] + 2Y^2.$$

Nous avons donc trouvé des formes réduites à une somme de carrés :

$$Q_3(X, Y) = 4(X - Y/2)^2 + 3Y^2 \text{ et } Q_4(X, Y) = 2(X + 2Y)^2 - 6Y^2.$$

Mais nous voyons qu'il **n'y a pas unicité** de ce type d'écriture !

Cependant, sous les formes que nous venons d'obtenir, la positivité ou non de la forme quadratique se reconnaît facilement : soit tous les coefficients des "carrés" sont positifs ou nuls et la réponse est positive, soit ce n'est pas le cas, et on montre que la réponse est

négative en prenant des vecteurs qui annulent les termes positifs. (On utilise pour cela le fait que les formes linéaires intervenant au sein des carrés sont indépendantes). Au passage, sur nos exemples, le résultat que nous trouvons est cohérent avec les formes dont nous sommes partis...

**Principe général.** De manière plus générale, l'algorithme est le suivant pour une forme quadratique générale à  $n$  variables  $X_1, \dots, X_n$  :

1. Regrouper tous les termes en  $X_1$ . S'il n'y a pas de terme en  $X_1^2$ , passer à  $X_2$  et ainsi de suite. S'il n'y a aucun terme en  $X_k^2$ , passer à l'étape 3, sinon à l'étape 2.
2. Mettre sous forme canonique en la variable sélectionnée. Aller à l'étape 4.
3. S'il n'y a pas de terme en  $X_k^2$ , on isole les termes en  $X_1$  et  $X_2$  (s'il y en a, sinon on fait la même chose avec des variables pour lesquelles il y a effectivement des termes!). Nous obtenons une partie de la forme quadratique (on garde les autres termes pour plus tard) de la forme :



$$\tilde{Q} := \alpha X_1 X_2 + X_1 L_1 + X_2 L_2,$$

où  $L_1$  et  $L_2$  sont des expressions linéaires des variables restantes  $X_3, X_4$ , etc. ( $L_1$  et  $L_2$  sont éventuellement nulles, par exemple s'il n'y a plus d'autres variables). On écrit alors

$$\tilde{Q} := \alpha \left[ \left( X_1 + \frac{L_2}{\alpha} \right) \left( X_2 + \frac{L_1}{\alpha} \right) - \frac{L_1 L_2}{\alpha^2} \right].$$

On utilise enfin l'identité remarquable  $ab = \frac{1}{4}[(a+b)^2 - (a-b)^2]$  pour obtenir

$$\tilde{Q} := \frac{\alpha}{4} \left[ \left( X_1 + X_2 + \frac{L_1 + L_2}{\alpha} \right)^2 - \left( X_1 - X_2 + \frac{L_1 - L_2}{\alpha} \right)^2 \right] - \frac{L_1 L_2}{\alpha}.$$

4. On itère avec les termes "restants" (ceux que nous n'avons pas mis sous forme de carré ou de somme de carrés) : ceux-ci ont strictement moins de variables!

Cela permet d'obtenir le théorème suivant.

**Théorème 14.** Soit  $Q$  une forme quadratique sur  $\mathbb{K}^n$ . Alors il existe un entier  $r \leq n$ ,  $r$  formes linéaires  $l_1, \dots, l_r$  sur  $\mathbb{K}^n$  **linéairement indépendantes** et  $r$  coefficients  $\alpha_1, \dots, \alpha_r$  de  $\mathbb{K}$  tous non nuls, tels que pour tout  $x$  de  $\mathbb{K}^n$  on ait

$$Q(x) = \sum_{i=1}^r \alpha_i l_i(x)^2. \tag{18}$$

Une telle écriture d'une forme quadratique est appelée **forme réduite**.

**Remarque 19.** Il faut bien retenir que les formes linéaires sont linéairement indépendantes. En particulier, il y en a moins que  $n$  ( $r \leq n$ ). Cela fait partie de la définition d'une forme réduite.

On peut alors répondre à notre question initiale, qui rappelons-le ne concerne que le cas de formes quadratiques réelles. Nous nous placerons dans ce cas jusqu'à la fin du Chapitre 3.

**Proposition 25.** *Une forme quadratique  $Q$  sur  $\mathbb{R}^n$  est positive si et seulement si tous les  $\alpha_i$  apparaissant dans la forme réduite (18) sont positifs.*

**Démonstration.** Ceci est une conséquence du fait que les  $(l_i)$  sont linéairement indépendants et du lemme suivant.

**Lemme 5.** *Soit  $E$  un espace de dimension finie, et soit  $(f_i)_{i=1}^k$  une famille libre de  $E^*$ . Il existe une base  $(e_j)_{j=1 \dots n}$  de  $E$  telle que pour tout  $i = 1 \dots k$  et  $j = 1 \dots n$  on ait  $f_i(e_j) = \delta_{ij}$ .*

**Preuve.** Soit  $n$  la dimension de  $E$  (et donc de  $E^*$ ). On complète les  $(f_i)_{i=1}^k$  en une base de  $E^*$ , mettons  $(f_i)_{i=1}^n$ . Alors l'application  $\Phi : E \rightarrow \mathbb{R}^n$  donnée par  $x \mapsto (f_1(x), \dots, f_n(x))$  est un isomorphisme. En effet, la linéarité est claire, et il suffit donc de montrer  $\Phi$  est injectif (pour des raisons de dimension). Maintenant, pour tout vecteur non nul  $v$ , il existe une forme linéaire ne s'annulant pas en ce point (par exemple  $\langle v, \cdot \rangle$ ). Donc si  $v \neq 0$ , il ne peut pas appartenir à  $\text{Ker}(\Phi)$ , puisque les  $(f_i)_{i=1}^n$  forment une base de  $E^*$ . Par suite, l'image réciproque de la base canonique de  $\mathbb{R}^n$  par  $\Phi$  répond à la question, ce qui achève la preuve du lemme.  $\square$

Revenons à la preuve de la Proposition 25. Si les  $\alpha_i$  sont tous positifs ou nuls, alors clairement  $Q$  est positive. Réciproquement, si les  $\alpha_i$  ne sont pas tous positifs ou nuls, soit  $k$  un indice tel que  $\alpha_k < 0$ . On introduit la base  $(e_j)_{j=1 \dots n}$  comme dans le Lemme 5. On vérifie alors que  $Q(e_k) < 0$ .  $\square$

**Remarque 20.** *Une bonne nouvelle : on peut voir en suivant l'algorithme précédent que si à un moment ou un autre on atteint l'étape 3 décrite précédemment ("S'il n'y a pas de carré"), cela veut dire que dans la réduction de Gauss que nous obtiendrons à la fin, il y aura des coefficients  $\alpha_i$  négatifs. On peut donc s'arrêter là si la seule question qui se posait était la positivité.*

**Remarque 21.** *On peut donner une forme un peu plus compacte à (18). En effet, réordonnons les  $\alpha_i l_i(x)^2$  pour mettre ceux qui correspondent à  $\alpha_i > 0$  en premier (mettons de 1 à  $p$ ) puis ceux qui correspondent à  $\alpha_i < 0$  (de  $p+1$  à  $r$ ). Alors on pose  $\tilde{l}_i = \sqrt{|\alpha_i|} l_i$  pour  $1 \leq i \leq r$  et on obtient la forme*

$$Q(x) = \sum_{i=1}^p \tilde{l}_i(x)^2 - \sum_{i=p+1}^r \tilde{l}_i(x)^2. \quad (19)$$

### 3.5 Signature d'une forme quadratique

**Théorème 15** (Signature). *Soit  $E$  un espace euclidien de dimension  $n$ . Soit  $q : E \rightarrow \mathbb{R}$  une forme quadratique. Alors il existe une base  $(e_1, \dots, e_n)$  de  $E$ , pas nécessairement*

orthonormée, et des entiers  $p$  et  $r$  avec  $p \leq r \leq n$ , tels que pour  $u = \sum_{i=1}^n x_i e_i$ , la forme  $q$  s'écrit :

$$q(u) = \sum_{i=1}^p x_i^2 - \sum_{i=p+1}^r x_i^2.$$

De plus les entiers  $p$  et  $r$  sont indépendants de la base  $(e_i)_{i=1, \dots, n}$  qui met  $q$  sous cette forme. Le couple d'entiers  $(p, r - p)$  sera appelé la **signature** de la forme quadratique  $q$  et sera noté  $\text{sign}(q) = (p, r - p)$ . L'entier  $r$  est appelé **rang** de la forme quadratique.

**Remarque 22.** L'écriture précédente sous-entend que la première somme (respectivement la seconde) est nulle (car l'ensemble d'indices est vide) si  $p = 0$  (resp. si  $p = r$ ).

**Démonstration.** L'existence d'une base  $(e_i)$  vient de la forme réduite (19) : on introduit les  $\tilde{l}_i$  comme dans (19), puis des  $(e_i)$  qui leur sont associés par le Lemme 5. Les  $r$  premières coordonnées de  $x$  dans la base  $(e_i)$  sont alors précisément les  $\tilde{l}_i(x)$ .

Ensuite, pour l'unicité du couple  $(p, r)$  : soient  $(e_i)_{i=1 \dots n}$  et  $(e'_i)_{i=1 \dots n}$  deux bases et  $u = \sum_{i=1}^n x_i e_i = \sum_{j=1}^n y_j e'_j$  ; on a



$$q(u) = \sum_{i=1}^p x_i^2 - \sum_{i=p+1}^r x_i^2 = \sum_{j=1}^{p'} y_j^2 - \sum_{j=p'+1}^{r'} y_j^2. \quad (20)$$

On introduit les notations :

$$\begin{aligned} F &= \text{Vect} \{e_1, \dots, e_p\}, & F' &= \text{Vect} \{e'_1, \dots, e'_{p'}\}, \\ G &= \text{Vect} \{e_{p+1}, \dots, e_r\}, & G' &= \text{Vect} \{e'_{p'+1}, \dots, e'_{r'}\}, \\ H &= \text{Vect} \{e_{r+1}, \dots, e_n\}, & H' &= \text{Vect} \{e'_{r'+1}, \dots, e'_n\}. \end{aligned}$$

Montrons que seules les paires  $(F, F')$ ,  $(G, G')$  et  $(H, H')$  peuvent avoir une intersection non triviale i.e.,  $F \cap G' = \{0\} = F \cap H' = G \cap F' = \dots$ , etc. Si  $u \in F \cap G$  et  $u \neq 0$ , alors d'après (20),  $u \in F$  implique  $q(u) > 0$  et  $u \in G'$  implique  $q(u) < 0$ . Ceci est une contradiction, donc  $F \cap G' = \{0\}$ .

Par conséquent, on a par exemple que les espaces  $F$ ,  $G'$  et  $H'$  sont en somme directe :  $F \oplus G' \oplus H' \subset E$ . Il suit que  $\dim(F) + \dim(G') + \dim(H') \leq \dim(E)$  c'est-à-dire  $p + r' - p' + n - r' \leq n$  et donc,  $p \leq p'$ . Mais en changeant les rôles des espaces nous obtenons aussi  $p' \leq p$ , donc  $p = p'$ . On applique ensuite le même raisonnement pour montrer  $r = r'$ .  $\square$

**Exemples.** La signature de la forme quadratique à trois variables  $X^2 + (Y + Z)^2 - (Y - Z)^2$  est  $(2, 1)$ . Celle de  $X^2 - (Y - Z)^2$  est  $(1, 1)$ .

On peut caractériser la positivité d'une forme quadratique en fonction de sa signature. Plus généralement que la Définition 37, on a les définitions suivantes.

**Définition 38.** Une forme quadratique  $q(u) : E \rightarrow \mathbb{R}$  est dite *positive* si  $q(x) \geq 0, \forall x \in E$ . Elle est dite **définie positive** si  $q(x) > 0, \forall x \neq 0$ . De même pour une forme **négative** ou **définie négative**. Une matrice symétrique  $A$  est dite *positive* ou *définie positive* (ce que l'on notera  $A \geq 0$  et  $A > 0$  respectivement) si la forme quadratique  $q(x) = \langle Ax, x \rangle$  l'est.

**Exemples.** Dans  $\mathbb{R}^2$ , la forme quadratique  $Q_1(X, Y) = X^2 + (X + Y)^2$  est définie positive. La forme quadratique  $Q_2(X, Y) = (X + Y)^2$  est positive, mais pas définie positive.

Introduisons également la définition de la non dégénérescence.

**Définition 39.** Une forme bilinéaire  $f : E \times E \rightarrow \mathbb{R}$  est dite **non dégénérée** lorsque pour tout  $u \in E \setminus \{0\}$ , il existe  $v \in E$  tel que  $f(u, v) \neq 0$ .

**Corollaire 3.** Soit  $q(u) : E \rightarrow \mathbb{R}$  une forme quadratique,  $n = \dim(E)$  et  $\text{sign}(q) = (n_+, n_-)$ . Alors :

- $q$  est définie positive si et seulement si  $n_+ = n$  (et donc  $n_- = 0$ );  $q$  est positive si et seulement si  $n_- = 0$  (et donc  $n_+ \leq n$ ),
- la forme bilinéaire symétrique  $f$  associée à  $q$  est non dégénérée si et seulement si  $n_+ + n_- = n$ .

**Démonstration.** Introduisons une base  $(e_1, \dots, e_n)$  comme dans le Théorème 15. Si  $n_+ = n$ , il est clair que  $q$  est définie positive. Si  $n_+ < n$ , alors pour  $n_+ < j \leq n$ , on voit que  $q(e_j) \leq 0$ , et donc  $q$  n'est pas définie positive. De même dans le cas positif : il s'agit qu'il n'y ait pas de vecteur de base tel que  $q(e_j) < 0$ .

Prouvons maintenant la deuxième assertion. Notons que la forme symétrique  $f$  associée à  $q$  a la forme suivante : pour  $u = \sum_{i=1}^p x_i e_i$  et  $v = \sum_{i=1}^r y_i e_i$  :

$$f(u, v) = \sum_{i=1}^p x_i y_i - \sum_{i=p+1}^r x_i y_i.$$

En effet, cette forme est clairement symétrique et satisfait  $f(u, u) = q(u)$ . Notons que sous cette forme et par définition de la signature, on a  $r = n_- + n_+$ . Il est clair que si  $r < n$ , alors on a  $f(e_n, \cdot) = 0$  et la forme est clairement dégénérée. Mais si  $r = n$ , alors quel que soit  $u \neq 0$ , il existe  $i \in \{1, \dots, n\}$  tel que la  $i$ -ème coordonnée de  $v$  dans la base  $(e_1, \dots, e_n)$  soit non nulle. Alors on voit que  $f(u, e_i) \neq 0$ .  $\square$

**Remarque 23.** On voit facilement à partir du Corollaire 3 qu'il y a équivalence pour une forme bilinéaire symétrique réelle **positive** entre "définie positive" et "non dégénérée positive". Mais on peut le montrer plus facilement en utilisant pour le sens non trivial ("non dégénérée positive"  $\Rightarrow$  "définie positive") :

$$\forall u, v \in E, \forall t \in \mathbb{R}, f(u + tv, u + tv) = f(u, u) + 2tf(u, v) + t^2 f(v, v) \geq 0,$$

donc le discriminant de ce polynôme en  $t$  est négatif ou nul :

$$f(u, v)^2 \leq f(u, u)f(v, v).$$

(C'est Cauchy-Schwarz pour une forme bilinéaire dont on ne sait pas encore que c'est un produit scalaire.) Pour  $u \in E \setminus \{0\}$ , on introduit  $v$  venant de la non dégénérescence de  $f$ , et on applique !

**Remarque 24.** Pour une forme bilinéaire non nécessairement positive, l'équivalence entre "non dégénérée" et "définie" est fautive en général.

## 4 Théorème de diagonalisation

Dans ce chapitre, nous utiliserons de nouveau les deux corps  $\mathbb{R}$  et  $\mathbb{C}$ .

### 4.1 Motivation

Un des grands enjeux de l'algèbre linéaire est de "réduire" les matrices d'endomorphismes, par exemple de les diagonaliser (c'est-à-dire de trouver une base où la matrice de l'endomorphisme est diagonale) lorsque cela est possible. Cela permet de simplifier considérablement certains calculs, comme par exemple les puissances  $n$ -ièmes de ces matrices.

La même question peut se poser pour une forme bilinéaire : existe-t-il des bases où la matrice de la forme bilinéaire est plus simple à écrire, par exemple diagonale ? On voit en particulier que si la matrice  $A$  de la forme bilinéaire  $f$  est diagonale dans la base  $(e_i)$ , notons  $A = \text{diag}(\alpha_i)$ , on peut écrire facilement la forme bilinéaire  $f(u, v) = \sum_i \alpha_i U^i V^i$ , où  $U$  et  $V$  sont les vecteurs colonnes des coordonnées de  $u$  et  $v$  dans la base  $(e_i)$ .

Nous allons décrire dans la suite un théorème qui montre que certaines matrices sont diagonalisables (à la fois en tant que matrice d'endomorphisme et en tant que matrice de forme quadratique).

### 4.2 Changement de base pour les matrices de forme bilinéaire

Soit  $f$  une forme bilinéaire sur un espace réel de dimension finie  $E$ , et soit  $(e_i)_{i=1}^n$  une base de  $E$ . On note  $A$  la matrice d'une forme bilinéaire  $f$  dans cette base. La question dans ce paragraphe est la suivante : comment s'écrit la matrice de  $f$  dans une autre base  $(\tilde{e}_i)_{i=1}^n$  ?

**Proposition 26.** *Soit  $P_{\tilde{e} \leftarrow e}$  la matrice de passage de  $(e_i)_{i=1}^n$  vers  $(\tilde{e}_i)_{i=1}^n$ . Alors la matrice de  $f$  dans la base  $(\tilde{e}_i)_{i=1}^n$  est donnée par*

$$\tilde{A} = {}^t P_{\tilde{e} \leftarrow e} A P_{\tilde{e} \leftarrow e}.$$

**Démonstration.** Soient  $u$  et  $v$  dans  $E$  ; notons  $U$  et  $V$  les vecteurs colonnes des coordonnées de  $u$  et  $v$  dans la base  $(e_i)_{i=1}^n$  et  $\tilde{U}$  et  $\tilde{V}$  les vecteurs colonnes des coordonnées de  $u$  et  $v$  dans la base  $(\tilde{e}_i)_{i=1}^n$ . On sait alors que

$$\tilde{U} = P_{\tilde{e} \leftarrow e}^{-1} U \text{ et } \tilde{V} = P_{\tilde{e} \leftarrow e}^{-1} V$$

Mais en utilisant la Proposition 19, on voit que que

$$f(u, v) = {}^t U A V = {}^t (P_{\tilde{e} \leftarrow e} \tilde{U}) A P_{\tilde{e} \leftarrow e} \tilde{V},$$

d'où le résultat. □

### 4.3 Vecteurs et valeurs propres

**Définition 40.** Soit  $M$  une matrice  $n \times n$  à entrées dans  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ . On dit que  $\lambda$  est une **valeur propre** de  $M$  lorsqu'il existe  $v \neq \mathbf{0}$  de  $\mathbb{R}^n$  (qu'on appelle **vecteur propre associé**) tel que

$$Mv = \lambda v.$$

**Remarque 25.** On parle de la même façon de **valeurs et vecteurs propres de l'endomorphisme**  $f$  représenté par  $M$  : l'équation devient

$$f(v) = \lambda v.$$

**Exemples.**

- $M = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ . Alors  $\lambda = 3$  est une valeur propre et  $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  un vecteur propre associé car

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

- $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ . Alors  $\lambda = i$  est une valeur propre et  $v = \begin{pmatrix} 1 \\ i \end{pmatrix}$  un vecteur propre associé car

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

**Remarque 26.** Si  $v$  est vecteur propre, tout  $\alpha v$  avec  $\alpha \in \mathbb{K}$  **non nul**, l'est également (pour la même valeur propre). Si  $v$  et  $w$  sont deux vecteurs propres non opposés associés à une même valeur propre, alors  $v + w$  l'est également.

**Définition 41.** Pour  $\lambda$  une valeur propre de la matrice  $M$ , l'ensemble des vecteurs propres associés à cette valeur propre, réuni avec  $\{0\}$ , forme un s.e.v. de  $\mathbb{R}^n$ . Ce sous espace  $V_\lambda$  est appelé **sous-espace propre** associé à  $\lambda$ .

**Remarque 27.** On voit que

$$V_\lambda = \{v / Mv = \lambda v\} = \text{Ker}(M - \lambda I).$$

Une conséquence importante du fait que  $\mathbb{C}$  soit algébriquement clos, c'est-à-dire que les polynômes de degré  $n$  aient toujours  $n$  racines comptées avec multiplicité (Théorème de d'Alembert-Gauss) est la suivante.

**Lemme 6.** Soit  $A$  une matrice carrée sur  $\mathbb{K}$ . Toute racine du polynôme caractéristique  $P(\lambda) = \det(A - \lambda I)$  est valeur propre de  $A$ . Par conséquent,  $A$  a toujours au moins une valeur propre complexe.

**Preuve.** Ce lemme est classique : on sait que  $\det(A - \lambda I) = 0 \Leftrightarrow \text{Ker}(A - \lambda I) \neq \{0\}$ , ce qui donne le résultat. Rappelons au passage que le polynôme caractéristique d'une matrice carrée de taille  $n$  est de degré  $n$  (ce qui vient des propriétés de multilinéarité du déterminant).  $\square$

On remarque que même si  $M$  a toutes ses entrées réelles il n'est pas nécessaire que les valeurs propres ou vecteurs propres complexes soient réels. Cependant ceci est vrai pour les matrices symétriques (c'est-à-dire, rappelons-nous, qui satisfont  ${}^tM = M$ ).

**Lemme 7.** Soit  $A$  une matrice symétrique réelle de taille  $n \times n$ .

1. Soit  $\lambda \in \mathbb{C}$  une valeur propre de  $A$ . Alors  $\lambda \in \mathbb{R}$  et il existe un vecteur propre réel  $v \in \mathbb{R}^n$  associé à  $\lambda$ .
2. Des vecteurs propres  $v_i$  et  $v_j$  de  $A$  qui correspondent à des valeurs propres différentes  $\lambda_i$  et  $\lambda_j$  sont orthogonaux deux à deux.

**Démonstration.** Notons tout d'abord que le fait que  $A = (A_{ij})_{i,j=1,\dots,n}$  soit symétrique réelle permet d'écrire :

$$\forall u, v \in \mathbb{C}^n, \langle Au, v \rangle_{\mathbb{C}^n} = \langle u, \overline{{}^tAv} \rangle_{\mathbb{C}^n} = \langle u, Av \rangle_{\mathbb{C}^n}. \quad (21)$$

Pour le premier point, nous introduisons  $w$  un vecteur propre associé à une valeur propre  $\lambda$  (a priori les deux sont complexes). Alors on a :

$$\langle Aw, w \rangle_{\mathbb{C}^n} = \lambda \langle w, w \rangle_{\mathbb{C}^n} = \lambda \|w\|^2,$$

et par ailleurs, grâce à (21), on a

$$\langle Aw, w \rangle_{\mathbb{C}^n} = \langle w, Aw \rangle_{\mathbb{C}^n} = \overline{\lambda} \|w\|^2.$$

Il suit que  $\lambda$  est réel. Donc il existe un vecteur propre réel, car la matrice réelle  $A - \lambda I$  n'est pas inversible.

Pour le deuxième point, on écrit

$$\lambda_i \langle v_i, v_j \rangle_{\mathbb{C}^n} = \langle Av_i, v_j \rangle_{\mathbb{C}^n} = \langle v_i, Av_j \rangle_{\mathbb{C}^n} = \lambda_j \langle v_i, v_j \rangle_{\mathbb{C}^n},$$

puisque  $\lambda_j$  est réel. Donc  $\langle v_i, v_j \rangle_{\mathbb{C}^n} = 0$ .  $\square$

## 4.4 Théorème de diagonalisation

Le théorème suivant est l'un des points les plus importants de ce cours.

**Théorème 16** (Diagonalisation des matrices symétriques réelles). Une matrice symétrique réelle  $A \in M_n(\mathbb{R})$  est diagonalisable en base orthonormée, i.e. il existe  $P \in O(n)$  tel que

$$\Delta = P^{-1}AP \text{ soit une matrice diagonale.} \quad (22)$$

De plus, les colonnes de  $P$  sont **des** vecteurs propres de  $A$  qui forment une base orthonormée de  $\mathbb{R}^n$  (en particulier elles sont deux à deux orthogonales). Les éléments diagonaux de  $\Delta$  sont **les** valeurs propres de  $A$ .

**Remarque 28.** Lorsque l'on peut obtenir la relation (22) avec  $P$  inversible seulement (mais pas nécessairement orthogonale), on dit que  $A$  est diagonalisable (tout court). On retiendra que **diagonaliser** une matrice, c'est déterminer les matrices  $A$  et  $P$  (et éventuellement  $P^{-1}$ ).

Le théorème précédent a pour corollaire le théorème suivant.

**Théorème 17** (Diagonalisation des formes bilinéaires). Soit  $f$  une forme bilinéaire symétrique sur l'espace euclidien  $E$  de dimension  $n$ . Alors il existe une base orthonormée  $(e_i)_{i=1}^n$  de  $E$  telle que la matrice de  $f$  par rapport à cette base soit diagonale, c'est-à-dire

$$f(e_i, e_j) = 0 \text{ si } i \neq j$$

L'action de  $f$  se calcule alors par

$$f(u, v) = \sum_{i=1}^n \alpha_i \langle u, e_i \rangle \langle v, e_i \rangle.$$

**Démonstration du Théorème 16.** Soit  $A \in M_n(\mathbb{R})$  une matrice symétrique réelle (notons  $f$  l'endomorphisme associé qui a pour matrice  $A$  dans la base canonique). Notons  $\{\lambda_1, \dots, \lambda_k\}$  l'ensemble de ses valeurs propres, c'est-à-dire des racines complexes de son polynôme caractéristique (comptées avec multiplicité, il y en aurait  $n$ ). D'après le Lemme 7, les  $\lambda_i$  sont réels. Notons  $V_i = \text{Ker}(A - \lambda_i I)$  l'espace propre dans  $\mathbb{R}^n$  associé à  $\lambda_i$ . D'après le Lemme 7, on voit que pour  $i \neq j$ ,  $V_i \perp V_j$ .

Montrons que

$$E = V_1 + \dots + V_k. \quad (23)$$

Si ce n'est pas le cas, notons  $F = V_1 + \dots + V_k$ , on a donc  $F^\perp$  (son orthogonal dans  $\mathbb{R}^n$ ) non réduit à  $\{0\}$ . Alors en utilisant (21) on voit que pour  $x \in F^\perp$ , on a pour tout  $v_i \in V_i$ ,

$$\langle v_i, Ax \rangle = \langle Av_i, x \rangle = \lambda_i \langle v_i, x \rangle = 0.$$

Il suit que pour  $x \in F^\perp$ ,  $Ax \in F^\perp$ . Donc la restriction  $f|_{F^\perp}$  de  $f$  à  $F^\perp$  peut être vue comme un endomorphisme de  $F^\perp$ . Donc il admet une valeur propre complexe. Celle-ci est bien entendu aussi une valeur propre de  $f$ , elle est donc réelle. Donc  $f|_{F^\perp}$  admet un vecteur propre réel dans  $F^\perp$ ! Ce qui contredit le fait que nous avons considéré toutes les valeurs et vecteurs propres précédemment.

Une fois (23) démontré, on choisit une base orthonormée dans chaque  $V_i$ . Comme  $V_i \perp V_j$  pour  $i \neq j$ , leur réunion fait une base orthonormée de  $E$ . Dans cette base composée de vecteurs propres, la matrice de l'endomorphisme  $f$  est donc diagonale avec les éléments diagonaux les valeurs propres de  $f$ .

Enfin on sait que la matrice de passage de la base canonique de  $\mathbb{R}^n$  vers une autre base orthonormée est une matrice orthogonale (voir la Proposition 17).  $\square$

**Démonstration du Théorème 17.** On construit la matrice  $A$  de  $f$  dans une base orthonormée de  $E$ . Cette matrice est diagonalisable en base orthonormée d'après ce qui précède : il existe  $P \in O(n)$  telle que (22) soit satisfait. Mais alors la matrice de  $f$  dans la nouvelle base est

$${}^tPAP = P^{-1}AP = \Delta,$$

c'est-à-dire que dans cette base, la matrice de la forme bilinéaire  $f$  est diagonale (voir la Proposition 26).  $\square$

**Remarque 29.** La diagonalisation en base orthonormée d'une matrice symétrique réelle n'est pas unique, dans le sens où  $P$  et  $\Delta$  dans l'écriture (22) ne le sont pas. Dans la preuve du Théorème 16, on a vu que l'on avait le choix de l'ordre des valeurs propres et sous-espaces propres, mais aussi le choix d'une base orthonormée dans chaque sous-espace propre.

## 4.5 Critères de positivité des formes quadratiques

Un corollaire du Théorème 17 est le suivant.

**Corollaire 4.** Soit  $A \in M_n(\mathbb{R})$  une matrice symétrique réelle. Soit  $(n_+, n_-)$  la signature de la forme quadratique  $q(x) = \langle Ax, x \rangle$  associée. Alors  $n_+$  (respectivement  $n_-$ ) est le nombre de valeurs propres strictement positives (resp. négatives) de  $A$ .

Par conséquent,  $A$  est positive (respectivement définie positive) si et seulement si les valeurs propres de  $A$  sont positives (resp. strictement positives).

**Démonstration.** Nous pouvons utiliser le Théorème 17 pour donner une nouvelle démonstration de la partie existence du Théorème 15 : il suffit de prendre la base  $(e_i)$  qui diagonalise  $A$ , puis de "normaliser" les  $e_i$  (les remplacer par  $e_i/\sqrt{|\alpha_i|}$  si  $\alpha_i \neq 0$  et ne pas transformer les autres) pour ramener les  $\alpha_i$  à  $\pm 1$ . Au passage, on voit bien que  $n_+$  (respectivement  $n_-$ ) est le nombre de valeurs propres strictement positives (resp. négatives) de  $A$ .  $\square$

Un autre critère classique de positivité des formes quadratiques est le suivant.

**Théorème 18 (Sylvester).** Soit  $A$  une matrice symétrique réelle. Alors  $A$  est définie positive si et seulement si tous les "mineurs principaux"  $\det \left( (A_{ij})_{i,j=1}^l \right)$  sont strictement positifs pour  $l = 1, \dots, n$ .



**Démonstration.** On démontre le théorème par récurrence.

Pour  $n = 1$ ,  $A$  est un nombre réel  $A = (A_{11})$ , et  $\langle Ax, x \rangle = A_{11}x^2$  est positif si et seulement si  $A_{11} > 0$ ; par ailleurs  $A_{11}$  est le (seul) mineur principal.

Supposons maintenant le résultat vrai pour des matrices de taille  $n - 1$  et prouvons-le pour des matrices de taille  $n$ . Soit  $A$  une telle matrice. Notons par  $(\lambda_i, v_i)$  des couples valeurs propres/vecteurs propres de la matrice  $A$ , les  $v_i$  formant une base orthonormée de  $\mathbb{R}^n$ .

Nous observons que si l'on note par  $A_{n-1} \in M_{n-1}(\mathbb{R})$  la matrice carrée "nord-ouest" des  $n - 1$  premières lignes et colonnes de  $A$ , alors :

$$\left\langle A \begin{pmatrix} x_1 \\ \dots \\ x_{n-1} \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ \dots \\ x_{n-1} \\ 0 \end{pmatrix} \right\rangle_{\mathbb{R}^n} = \left\langle A_{n-1} \begin{pmatrix} x_1 \\ \dots \\ x_{n-1} \end{pmatrix}, \begin{pmatrix} x_1 \\ \dots \\ x_{n-1} \end{pmatrix} \right\rangle_{\mathbb{R}^n}.$$

*Sens direct.* Nous savons que  $\langle Ax, x \rangle > 0$  pour tout  $x \in \mathbb{R}^n$ ,  $x \neq 0$ , donc aussi en particulier pour tous les  $x$  de la forme

$$x = \begin{pmatrix} x_1 \\ \dots \\ x_{n-1} \\ 0 \end{pmatrix}.$$

Par conséquent,  $A_{n-1}$  est une matrice définie positive donc tous les mineurs de  $A_{n-1}$  sont strictement positifs, donc les  $n - 1$  premiers mineurs de  $A$  sont strictement positifs. Il ne reste donc plus qu'à montrer que  $\det(A) > 0$ . Mais le fait que  $A$  soit définie positive implique que  $\lambda_i > 0$  pour tous les  $i$  et comme  $\det(A) = \prod_{i=1}^n \lambda_i$ , il suit que  $\det(A) > 0$ .

*Sens réciproque.* Si tous les mineurs sont strictement positifs on applique l'hypothèse pour déduire que  $A_{n-1}$  est une matrice définie positive. Comme  $\det(A) > 0$  on a l'alternative suivante : soit toutes les valeurs propres sont strictement positives (et donc  $A$  est définie positive), soit au moins deux valeurs propres  $\lambda_i$  et  $\lambda_j$  sont strictement négatives. Dans ce cas, il existe au moins une combinaison linéaire  $\alpha v_i + \beta v_j$  avec  $\alpha^2 + \beta^2 \neq 0$  telle que  $\alpha v_i + \beta v_j$  ait zéro sur la dernière coordonnée. Puisque nous avons démontré que  $A_{n-1}$  était définie positive il suit que  $\langle A(\alpha v_i + \beta v_j), \alpha v_i + \beta v_j \rangle > 0$ . Mais d'un autre coté

$$\langle A(\alpha v_i + \beta v_j), \alpha v_i + \beta v_j \rangle = \alpha^2 \lambda_i + \beta^2 \lambda_j < 0,$$

d'où une contradiction. □

## 4.6 Deux applications de la diagonalisation

Donnons deux applications immédiates de la diagonalisation. Soit  $A \in M_n(\mathbb{R})$  une matrice diagonalisable.

**1. Puissance d'une matrice, récurrence linéaire.** Si l'on a la relation (22), alors il suit immédiatement que

$$A^k = P \Delta^k P^{-1}.$$

Or  $\Delta^k$  est très simple à calculer :

$$\text{si } \Delta = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}, \text{ alors } \Delta^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix}.$$

*Application : récurrence linéaire.* La diagonalisation est très utilisée pour calculer le résultat d'une suite récurrente vectorielle : si la suite  $(x_k)_{k \in \mathbb{N}}$  d'éléments de  $\mathbb{R}^n$  est donnée par la relation de récurrence

$$x_{k+1} = Ax_k,$$

alors on montre facilement que

$$x_k = A^k x_0.$$

*Cas non homogène.* On peut aussi considérer les suites satisfaisant des récurrences du type :

$$x_{k+1} = Ax_k + b, \tag{24}$$

où  $b$  est un vecteur fixe de  $\mathbb{R}^n$ . La méthode suivante fonctionne si 1 n'est pas valeur propre de  $A$  (au moins dans ce cas devrions-nous dire). Il s'agit de déterminer  $v \in \mathbb{R}^n$  un "point fixe" de la relation de récurrence, i.e. un point satisfaisant :

$$v = Av + b.$$

Si 1 n'est pas valeur propre de  $A$ , cela est toujours possible car dans ce cas  $A - I$  est injectif donc inversible, et on prend  $v = -(A - I)^{-1}(b)$ . Et dans ce cas, on observe qu'une suite  $(x_n)$  satisfait (24) si et seulement si la suite de terme général  $y_n = x_n - v$  satisfait

$$y_{k+1} = Ay_k,$$

ce que l'on sait calculer par la puissance de matrice !

**2. Exponentielle de matrice.** On se propose ici de donner un sens à l'exponentielle d'une matrice carrée, qu'on notera  $\exp(A)$ , dans le cas où  $A$  est symétrique réelle. Comme on sait que pour  $x$  réel, on a :

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

on a envie de poser

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Grâce au calcul précédent, on peut voir que pour les matrices symétriques réelles  $A$  diagonalisées comme dans (22), on a

$$\sum_{k=0}^{\infty} \frac{A^k}{k!} = P \left( \sum_{k=0}^{\infty} \frac{\Delta^k}{k!} \right) P^{-1}.$$

Cela permet de voir qu'on a bien convergence (élément par élément de la matrice) vers

$$\sum_{k=0}^{\infty} \frac{A^k}{k!} = P \begin{pmatrix} \exp(\lambda_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \exp(\lambda_n) \end{pmatrix} P^{-1}.$$

**Remarque 30.** *En fait, on peut définir l'exponentielle de matrice de toute matrice carrée sur  $\mathbb{R}$  ou  $\mathbb{C}$  (non nécessairement diagonalisable), mais cela est un peu plus délicat dans le cas général (et c'est beaucoup plus difficile à calculer). Et on peut étendre le procédé à d'autres fonctions que la fonction exponentielle, en particulier aux fonctions entières (c'est-à-dire développables en série entières avec rayon de convergence infini).*

L'exponentielle de matrice a d'importantes conséquences en théorie des équations différentielles vectorielles (souvent appelées *systèmes différentiels*). On montre que  $A$  étant une matrice fixée, l'unique solution  $y$  (dépendant du temps et à valeurs dans  $\mathbb{R}^n$ ), de l'équation différentielle vectorielle :

$$y' = Ay,$$

avec condition initiale

$$y(0) = y_0,$$

est donnée par

$$y(t) := \exp(tA)y_0.$$

Ce n'est pas sans rappeler le cas où  $y$  est à valeurs réelles !

## 5 Espace vectoriel normé

### 5.1 Evm et sa structure “topologique”

On rappelle que  $(\mathbb{R}, +, \cdot, \leq)$  est l’unique “corps” totalement ordonné qui satisfait à l’une des trois (et donc aux trois) propriétés suivantes

1. Tout ensemble non vide majoré possède une borne supérieure, ou de manière équivalente : si  $(x_n)$  est une suite réelle croissante et majorée (i.e.  $x_n \leq x_{n+1}$ ,  $\exists M, x_n \leq M$  pour tout  $n \in \mathbb{N}$ ) alors il existe  $\ell \in \mathbb{R}$  tel que  $x_n \rightarrow \ell$  lorsque  $n \rightarrow \infty$ .
2. Tout ensemble non vide fermé borné est compact, ou de manière équivalente : si  $(x_n)$  est une suite réelle telle que  $x_n \in [-M, M]$ ,  $M \geq 0$ , alors on peut en extraire une sous-suite convergente (i.e.  $\exists(x_{n_k}), \exists \ell \in [-M, M]$  tels que  $x_{n_k} \rightarrow \ell$  lorsque  $k \rightarrow \infty$ ).
3. Toute suite de Cauchy est convergente, ou de manière équivalente : si  $(x_n)$  est une suite réelle telle que pour tout  $\varepsilon > 0$  il existe  $N = N(\varepsilon)$  de sorte que  $|x_n - x_m| \leq \varepsilon$  pour tout  $m, n \geq N$ , alors il existe  $\ell \in \mathbb{R}$  tel que  $x_n \rightarrow \ell$  lorsque  $n \rightarrow \infty$ .

**Remarque 31.**  $(\mathbb{Q}, +, \cdot, \leq)$  est un “corps” totalement ordonné, mais il ne satisfait à aucune de ces trois propriétés. On dit que  $\mathbb{R}$  est complet, localement compact et satisfait à l’axiome de la borne supérieure.

On rappelle que dans  $\mathbb{R}^d$ , on définit la norme

$$\forall x = (x_1, \dots, x_d) \in \mathbb{R}^d, \quad \|x\|_\infty := \max\{|x_i|, 1 \leq i \leq d\}.$$

Dans un espace vectoriel normé  $(E, N)$ , on dit qu’une suite  $(x_n)$  converge vers  $\bar{x} \in E$  si  $N(x_n - \bar{x}) \rightarrow 0$  lorsque  $n \rightarrow \infty$ .

**Lemme 8.** On munit  $\mathbb{R}^d$  de la norme  $\|\cdot\|_\infty$ . Les boules fermées

$$B(0, R) := \{x \in \mathbb{R}^d, \|x\|_\infty \leq R\}$$

sont compactes.

**Démonstration du Lemme 8.** On considère une suite  $(x_n)$  de  $B(0, R)$  et on note  $x_n := (x_n^1, \dots, x_n^d)$ . Comme  $(x_n)$  est une suite réelle telle que  $|x_n^1| \leq R$  pour tout  $n \in \mathbb{N}$ , on peut en extraire une sous-suite  $(x_{n_k}^1)$  convergente. A nouveau, la suite réelle  $(x_{n_k}^2)$  satisfait  $|x_{n_k}^2| \leq R$  pour tout  $k \in \mathbb{N}$  et on peut en extraire une sous-suite  $(x_{n_k}^2)$  convergente. On répète  $d$  fois le procédé en remarquant que  $\{n_k^{i+1}, k \in \mathbb{N}\} \subset \{n_k^i, k \in \mathbb{N}\}$ . De la sorte, on exhibe  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^d)$  tels que  $x_{n_k}^i \rightarrow \bar{x}^i$  lorsque  $k \rightarrow \infty$  pour tout  $i \in \{1, \dots, d\}$ , ce qui signifie exactement  $x_{n_k}^d \rightarrow \bar{x}$  dans  $\mathbb{R}^d$  au sens de la norme  $\|\cdot\|_\infty$  lorsque  $k \rightarrow \infty$ .  $\square$

**Théorème 19.** Dans  $\mathbb{R}^d$  toutes les normes sont équivalentes.

**Démonstration du Théorème 19.** Il suffit de montrer que toutes les normes de  $\mathbb{R}^d$  sont équivalentes à la norme  $\|\cdot\|_\infty$ .

*Etape 1.* Soit  $N$  une deuxième norme de  $\mathbb{R}^d$ . Avec les notations habituelles, on a

$$\begin{aligned} N(x) &\leq N(x_1, 0, \dots, 0) + \dots + N(0, \dots, 0, x_d) \\ &= |x_1|N(e_1) + \dots + |x_d|N(e_d) \\ &\leq \{N(e_1) + \dots + N(e_d)\}\|x\|_\infty, \end{aligned}$$

ce qui donne une première inégalité.

*Etape 2.* Supposons par l'absurde que l'on n'a pas

$$\exists C \geq 0, \quad \forall x \in \mathbb{R}^d, \quad \|x\|_\infty \leq CN(x).$$

Cela signifie que

$$\forall n \geq 1, \quad \exists x_n \in \mathbb{R}^d, \quad \|x_n\|_\infty > nN(x_n) \geq 0.$$

On définit  $y_n := x_n/\|x_n\|_\infty$  qui satisfait donc

$$nN(y_n) \leq \|y_n\|_\infty = 1, \quad \forall n \geq 1.$$

Par le lemme 8, on en déduit qu'il existe une sous-suite  $(y_{n_k})$  et un vecteur  $\bar{y} \in \mathbb{R}^d$  tels que  $y_{n_k} \rightarrow \bar{y}$  dans  $\mathbb{R}^d$  au sens de la norme  $\|\cdot\|_\infty$ , et donc également au sens de la norme  $N$  d'après l'étape 1.

On en déduit d'une part

$$\left| \|\bar{y}\|_\infty - 1 \right| = \left| \|\bar{y}\|_\infty - \|y_{n_k}\|_\infty \right| \leq \|\bar{y} - y_{n_k}\|_\infty \rightarrow 0,$$

et donc  $\|\bar{y}\|_\infty = 1, \bar{y} \neq 0$ .

On en déduit d'autre part

$$N(\bar{y}) \leq N(\bar{y} - y_{n_k}) + N(y_{n_k}) \leq C\|\bar{y} - y_{n_k}\|_\infty + 1/n_k \rightarrow 0,$$

et donc  $N(\bar{y}) = 0, \bar{y} = 0$ . □

**Corollaire 5.** Dans  $\mathbb{R}^d$ , les suites de Cauchy et les suites convergentes sont les mêmes quelle que soit la norme considérée. Dorénavant et en général, on ne précisera plus la norme considérée dans  $\mathbb{R}^d$ .

**Corollaire 6.** Tout espace vectoriel normé de dimension finie est complet : toute les suites de Cauchy sont convergentes.

**Démonstration du Corollaire 6.** Soit  $(E, N)$  un evn de dimension finie. Une base  $(e_1, \dots, e_d)$  étant choisie, celui-ci est isomorphe à  $(\mathbb{R}^d, \nu)$  avec

$$\nu(x_1, \dots, x_d) := N\left(\sum_{i=1}^d x_i e_i\right).$$

Il suffit donc de montrer que  $(\mathbb{R}^d, \|\cdot\|_\infty)$  est complet. Considérons une suite de Cauchy  $(x_n)$  de  $\mathbb{R}^d$ . Chaque composante  $(x_n^i)$  est alors une suite de Cauchy réelle et est donc convergente. On en déduit que la suite  $(x_n)$  est convergente dans  $\mathbb{R}^d$ . □

**Lemme 9.** Soit  $E$  un evn. Il y a équivalence entre

1.  $E$  est complet (les suites de Cauchy sont convergentes).
2. Les séries normalement convergentes sont convergentes.

**Démonstration du Lemme 9.** On ne démontre que l'implication (1)  $\Rightarrow$  (2) qui nous sera utile dans une prochaine section. On suppose donc que  $E$  est complet. Soit  $((x_n))$  une série de  $E$  normalement convergente au sens où

$$\sum_{k=1}^{\infty} \|x_k\| < \infty.$$

On rappelle que cela implique

$$\sum_{k=N}^{\infty} \|x_k\| \rightarrow 0 \text{ lorsque } N \rightarrow \infty.$$

On considère maintenant la somme partielle

$$S_n := \sum_{k=1}^n x_k.$$

On observe alors que pour  $m \geq n$ , on a

$$\|S_m - S_n\| = \left\| \sum_{k=n+1}^m x_k \right\| \leq \sum_{k=n+1}^{\infty} \|x_k\| \rightarrow 0 \text{ lorsque } n \rightarrow \infty,$$

ce qui signifie que  $(S_n)$  est une suite de Cauchy. On en déduit que la suite  $(S_n)$  est convergente dans  $E$ . On note

$$S := \lim_{n \rightarrow \infty} \sum_{k=1}^n x_k$$

cette limite. □

## 5.2 Norme d'opérateurs (ou norme subordonnée)

**Proposition 27.** Soient  $E, F$  deux evn et  $T : E \rightarrow F$  une application linéaire. Il y a équivalence entre

1.  $T$  est bornée :  $\sup_{x \in B_E} \|Tx\|_F < \infty$ ,  $B_E := \{x \in E; \|x\|_E \leq 1\}$  ;
2.  $T$  est continue :  $Tx_n \rightarrow Tx$  dans  $F$  si  $x_n \rightarrow x$  dans  $E$  ;
3.  $T$  est Lipschitzienne :  $\exists L < \infty$  telle que  $\|Tx - Ty\|_F \leq L\|x - y\|_E, \forall x, y \in E$ .

**Démonstration de la Proposition 27.** On ne démontre que l'implication (1)  $\Rightarrow$  (3). Les autres implications, plus simples, sont laissées en exercices. On suppose (1) et donc qu'il existe  $L < \infty$  telle que  $\|Tz\|_F \leq L$  pour tout  $z \in B_E$ . Soient  $x, y \in E$ . Si  $x = y$ , on a évidemment (3). Sinon,  $\|x - y\|_E \neq 0$  et on calcule

$$\begin{aligned} \|Tx - Ty\|_F &= \|T(x - y)\|_F \\ &= \|x - y\|_E \left\| T\left(\frac{x - y}{\|x - y\|_E}\right) \right\|_F \leq \|x - y\|_E L, \end{aligned}$$

puisque  $(x - y)/\|x - y\|_E \in B_E$ .  $\square$

**Proposition 28.** Dans  $\mathbb{R}^d$ , toutes les applications linéaires sont bornées, et donc également continues, Lipschitziennes.

**Démonstration de la Proposition 28.** Etant donnée une base  $(e_1, \dots, e_d)$ , on pose  $L = \|Te_1\|_F + \dots + \|Te_d\|_F < \infty$ . On a alors

$$\|Tx\|_F \leq |x_1| \|Te_1\|_F + \dots + |x_d| \|Te_d\|_F \leq L,$$

pour tout  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ ,  $\|x\|_\infty \leq 1$ , ce qui signifie que  $T$  est bornée. On conclut grâce à la Proposition 27.  $\square$

**Proposition 29.** Pour  $T \in \mathcal{L}(E, F)$ ,  $E = \mathbb{R}^d$ , on définit

$$\| \|T\| \| := \sup_{B_E} \|Tx\|_F = \sup_{x \in E \setminus \{0\}} \frac{\|Tx\|_F}{\|x\|_E}. \quad (25)$$

Alors  $\| \cdot \|$  est une norme sur  $\mathcal{L}(E, F)$ . De plus, si  $F = E$ , on a

$$\| \|ST\| \| \leq \| \|S\| \| \| \|T\| \|, \quad \forall S, T \in \mathcal{L}(E). \quad (26)$$

**Démonstration de la Proposition 29.** On commence par montrer (25). On a d'une part

$$\sup_{x \in E \setminus \{0\}} \frac{\|Tx\|_F}{\|x\|_E} \geq \sup_{x \in B_E \setminus \{0\}} \frac{\|Tx\|_F}{\|x\|_E} \geq \sup_{x \in B_E \setminus \{0\}} \|Tx\|_F = \| \|T\| \|.$$

On a d'autre part

$$\frac{\|Tx\|_F}{\|x\|_E} = \left\| T\left(\frac{x}{\|x\|_E}\right) \right\|_F \leq \| \|T\| \|, \quad \forall x \in E \setminus \{0\}.$$

On laisse la preuve du fait que  $\| \cdot \|$  est une norme en exercice. Enfin, pour  $S, T \in \mathcal{L}(E)$ , on calcule

$$\begin{aligned} \| \|ST\| \| &= \sup_{x \in B_E} \|S(Tx)\|_E \\ &\leq \sup_{x \in B_E} \| \|S\| \| \|Tx\|_E = \| \|S\| \| \| \|T\| \|, \end{aligned}$$

ce qui démontre (26).  $\square$

A cause de la propriété (26), on dit qu'une telle norme est une norme d'algèbre. Par récurrence, on a également

$$\| \|T^k\| \| \leq \| \|T\| \|^k, \quad \forall T \in \mathcal{L}(E), \quad \forall k \in \mathbb{N}.$$

### 5.3 La fonction exponentielle de matrices

Pour  $A \in M_d(\mathbb{R})$  (et on pourrait faire la même chose pour  $u \in \mathcal{L}(\mathbb{R}^d)$ ), on définit

$$e^A = \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

**Lemme 10.**  $\exp(A)$  est bien définie dans  $M_d(\mathbb{R})$ .

**Démonstration du Lemme 10.** On munit  $M_d(\mathbb{R})$  d'une norme d'opérateurs, notée  $\|\cdot\|$ , associée à une norme de  $\mathbb{R}^d$ . En utilisant la propriété (26), on observe que

$$\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\| \leq \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = \exp(\|A\|) < \infty,$$

de sorte que la série  $((A^k/k!))$  est normalement convergente. L'espace  $M_d(\mathbb{R})$  étant de dimension finie ( $d^2$ ), le Lemme 9 implique que la série  $((A^k/k!))$  est convergente. Cela définit  $\exp(A)$ .  $\square$

**Lemme 11.** Si  $A$  et  $B$  commutent, i.e.  $AB = BA$ , alors  $\exp(A + B) = \exp(A)\exp(B)$ .

**Démonstration du Lemme 11.** On calcule (formellement)

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{1}{n!} (A+B)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^n \frac{n!}{j!(n-j)!} A^j B^{n-j} \\ &= \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \mathbf{1}_{n-j \geq 0} \frac{1}{j!(n-j)!} A^j B^{n-j} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{j!k!} A^j B^k = e^A e^B. \end{aligned}$$

$\square$

**Corollaire 7.** Soit  $A \in M_d(\mathbb{R})$ . L'application  $t \mapsto \exp(tA)$  est un homéomorphisme de groupe de  $\mathbb{R}$  dans le groupes  $GL(n)$  des matrices inversibles. En particulier, on a

1.  $e^{0A} = I$ ;
2.  $e^{-A}e^A = e^Ae^{-A} = I$ , soit donc  $(e^A)^{-1} = e^{-A}$ ;
3.  $e^{(s+t)A} = e^{sA}e^{tA}$  pour tout  $s, t \in \mathbb{R}$ .

**Lemme 12.** Pour tout  $A \in M_d(\mathbb{R})$  et  $t \in \mathbb{R}$ , on a

$$\frac{d}{dt} e^{tA} := \lim_{s \rightarrow 0} \frac{e^{(s+t)A} - e^{tA}}{s} = A e^{tA}.$$

**Démonstration du Lemme 12.** Pour  $t = 0$ , on écrit

$$\begin{aligned} e^{sA} - \mathbf{I} - sA &= \sum_{n=2}^{\infty} \frac{1}{n!} s^n A^n \\ &= s^2 A^2 \sum_{n=0}^{\infty} \frac{s^n A^n}{(n+2)!}. \end{aligned}$$

Pour  $s \in [-1, 1]$ , on en déduit

$$\|e^{sA} - \mathbf{I} - sA\| \leq s^2 \|A^2\| \sum_{n=0}^{\infty} \frac{\|A^n\|}{(n+2)!} \leq s^2 \|A\|^2 e^{\|A\|}$$

Cela implique

$$\left\| \frac{e^{sA} - \mathbf{I}}{s} - A \right\| = O(s) \rightarrow 0$$

lorsque  $s \rightarrow 0$ . On en déduit le résultat pour  $t \neq 0$  en utilisant la propriété de groupe 3) du Corollaire 7.  $\square$

**Lemme 13.** Soient  $A \in M_d(\mathbb{R})$ ,  $y_0 \in \mathbb{R}^d$  et  $y(t) := e^{tA} y_0$ . Alors

$$y'(t) = Ay(t), \quad \forall t \in \mathbb{R}, \quad y(0) = y_0.$$

**Démonstration du Lemme 12.** On calcule

$$\begin{aligned} y'(t) &:= \lim_{s \rightarrow 0} \frac{y(t+s) - y(t)}{s} \\ &= \lim_{s \rightarrow 0} \frac{e^{(s+t)A} - e^{tA}}{s} y_0 \\ &= Ae^{tA} y_0 = Ay(t). \end{aligned}$$

$\square$

## 6 Applications

### 6.1 Coniques



On considère les courbes du plan définies par une équation polynomiale du deuxième degré, c'est-à-dire pour  $a, b, c, d, e, f \in \mathbb{R}$  l'ensemble des points

$$\mathcal{C} = \{(X, Y) \in \mathbb{R}^2 / aX^2 + bXY + cY^2 + dX + eY + f = 0\}. \quad (27)$$

Les courbes qui satisfont cette équation sont appelées “coniques”.

Avec les notations

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}, \quad B = \begin{pmatrix} d \\ e \end{pmatrix}, \quad u = \begin{pmatrix} X \\ Y \end{pmatrix},$$

l'équation de (27) admet l'écriture matricielle

$$\mathcal{C} = \{u \in \mathbb{R}^2 / \langle Au, u \rangle + \langle B, u \rangle + f = 0\}.$$

Nous distinguerons plusieurs cas.

**Premier cas.** Si  $\text{rang}(A) = 0$ . Dans ce cas, on a donc  $A = 0$ ; l'équation sera  $\langle B, u \rangle + f = 0$  et donc la courbe est une droite.

**Deuxième cas.** Si  $\text{rang}(A) = 2$ . On peut alors résoudre de manière unique l'équation  $2Aw + B = 0$  d'inconnue  $w$  (car  $A$  est inversible). Ensuite, par le changement d'inconnue  $u = w + v$  on obtient l'équation suivante, en posant  $\alpha = -\langle Aw, w \rangle + \langle B, w \rangle + f$  :

$$\mathcal{C} = \{u = w + v \in \mathbb{R}^2 / \langle Av, v \rangle = \alpha\}.$$

On écrit alors la décomposition

$$A = {}^tU \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} U,$$

(avec  $\lambda$  et  $\mu$  non nuls pour des raisons de rang). Cherchons ce que devient l'équation dans la base des vecteurs colonnes de  $U$ . On note  $V = Uv$  et avec la notation  $V = (X, Y)$  l'équation devient

$$\lambda X^2 + \mu Y^2 = \alpha.$$

La courbe sera donc

- une ellipse si  $\lambda$  et  $\mu$  sont tous les deux positifs ou tous les deux négatifs (strictement pour des raisons de rang), cette ellipse étant éventuellement vide si  $\alpha$  a le mauvais signe.

- une hyperbole sinon.

*Exemple.*  $(X/2)^2 + Y^2 = 2$  donne une ellipse,  $Y^2 - X^2 = 1$  une hyperbole.

**Troisième cas.** Si  $\text{rang}(A) = 1$ . Dans ce cas, on peut écrire

$$A = {}^tU \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} U,$$

(avec  $\lambda \neq 0$ ). Après changement de repère (comme précédemment) on obtient l'équation  $\lambda X^2 + \alpha X + \beta Y + \gamma = 0$ . Alors nous avons deux sous-cas différents.

- Si  $\beta = 0$ , alors

$$\mathcal{C} = \{u = (X, Y) \in \mathbb{R}^2 / \lambda X^2 + \alpha X + \gamma = 0\}.$$

Selon le nombre de solutions de l'équation en  $X$ , on obtient l'ensemble vide, une ou deux droites parallèles à l'axe des ordonnées dans les coordonnées  $(X, Y)$ , d'équation  $X = x_1$  et  $X = x_2$ .

- Si  $\beta \neq 0$ , on obtient la parabole

$$\mathcal{C} = \left\{ u = (X, Y) \in \mathbb{R}^2 / Y = -\frac{\lambda X^2 + \alpha X + \gamma}{\beta} \right\}.$$

## 6.2 Interprétation géométrique des groupes $O(2)$ et $O(3)$

Dans cette partie, nous décrivons géométriquement les isométries vectorielles dans le plan et l'espace tridimensionnel euclidiens.

### 6.2.1 Le groupe $O(2)$

Soit  $A \in O(2)$ ,  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Les vecteurs  $\begin{pmatrix} a \\ c \end{pmatrix}$  et  $\begin{pmatrix} b \\ d \end{pmatrix}$  forment donc une base orthonormée de  $\mathbb{R}^2$ . On en déduit que :

- $a^2 + c^2 = 1$  donc il existe  $\theta$  avec  $a = \cos \theta$ ,  $c = \sin \theta$ ,
- $b^2 + d^2 = 1$  donc il existe  $\phi$  avec  $b = \cos \phi$ ,  $d = \sin \phi$ ,
- $ab + cd = 0$  donc  $\cos \theta \cos \phi + \sin \theta \sin \phi = 0$  donc  $\cos(\phi - \theta) = 0$ ; en particulier  $\sin(\phi - \theta) = \pm 1$ .

Donc  $d = \sin(\phi - \theta + \theta) = \sin(\phi - \theta) \cos \theta + \cos(\phi - \theta) \sin \theta = \sin(\phi - \theta) \cos \theta$ . Par conséquent, nous sommes dans l'une des deux situations suivantes.

**Première situation.** Si  $\sin(\phi - \theta) = 1$ , alors  $d = \cos \theta$  et  $b = -\sin \theta$ . Donc

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Dans ce cas  $\det A = 1$  donc  $A \in SO(2)$ .

*Interprétation géométrique.* Soit le vecteur  $\begin{pmatrix} x \\ y \end{pmatrix}$  dont l'écriture en coordonnées polaires est  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix}$ . Alors on a :

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} = \begin{pmatrix} \rho \cos(\alpha + \theta) \\ \rho \sin(\alpha + \theta) \end{pmatrix}$$

donc  $A$  est une rotation d'angle  $\theta$  dans le sens trigonométrique.

**Deuxième situation.** Si  $\sin(\phi - \theta) = -1$ , alors  $d = -\cos \theta$  et  $b = \sin \theta$ . Donc

$$A = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Dans ce cas  $\det A = -1$  donc  $A \notin SO(2)$ .

*Interprétation géométrique.* Soit le vecteur  $\begin{pmatrix} x \\ y \end{pmatrix}$  dont l'écriture en coordonnées polaires est  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix}$ . Alors

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} = \begin{pmatrix} \rho \cos(\theta - \alpha) \\ \rho \sin(\theta - \alpha) \end{pmatrix}$$

donc  $A$  est une symétrie (sur le cercle de rayon  $\rho$ ) par rapport à la droite qui fait un angle  $\theta/2$  avec  $Ox$  dans le sens trigonométrique.

### 6.2.2 Le groupe $O(3)$

**Théorème 20.** Soit  $A \in O(3)$  et  $f$  un endomorphisme de  $\mathbb{R}^3$  dont  $A$  est la matrice dans la base canonique notée  $(e_1, e_2, e_3)$ . Alors il existe une base orthonormée  $(v_1, v_2, v_3)$  telle que la matrice  $M$  de  $f$  dans cette nouvelle base soit

$$M = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & \det(A) \end{pmatrix}.$$



**Démonstration.** Remarquons d'abord que si  $\lambda$  est valeur propre réelle de  $A$  alors  $\lambda = \pm 1$ ; en effet  $Av = \lambda v$  mais  $\|Av\| = \|v\|$  donc  $|\lambda| = 1$ .

Prouvons ensuite que  $\det(A)$  est valeur propre de  $A$ . En effet, les valeurs propres de  $A$  sont des racines du polynôme caractéristique  $\det(A - \lambda I)$  qui est à coefficients réels et de degré 3. Il a donc soit une, soit trois racines réelles. Examinons ces deux cas.

- S'il a une racine réelle  $\lambda$  et deux complexes  $\lambda_2$  et  $\lambda_3$ . Il suit que  $\lambda_3 = \overline{\lambda_2}$ . Alors  $\det(A) = \lambda\lambda_2\lambda_3 = \lambda|\lambda_2|^2$  donc  $\det(A)$  et  $\lambda$  ont le même signe, et comme les deux sont de module 1, elles sont égales.
- S'il a trois valeurs propres réelles  $\lambda_1, \lambda_2$  et  $\lambda_3$ , elles sont nécessairement toutes  $\pm 1$ . Comme  $\det(A) = \lambda_1\lambda_2\lambda_3$  au moins une aura le signe du  $\det(A)$  et on obtient donc la conclusion.

Soit donc maintenant  $v$  un vecteur propre (de norme 1) correspondant à la valeur propre  $\lambda = \det(A)$ . On complète  $v$  en une base orthonormée de  $\mathbb{R}^3$  notée  $(v_1, v_2, v_3)$  avec  $v = v_3$ . Les entrées de la matrice  $M$  de  $A$  dans cette base sont alors  $M_{ij} = \langle f(v_j), v_i \rangle_{\mathbb{R}^3}$ . Par exemple  $M_{i3} = \langle f(v_3), v_i \rangle_{\mathbb{R}^3} = \langle \lambda v_3, v_i \rangle_{\mathbb{R}^3} = \lambda \delta_{3i}$ . Donc

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ 0 & 0 & \det(A) \end{pmatrix}.$$

Mais, comme  $f$  est orthogonale,  $M$  l'est aussi. En particulier les lignes 1 et 2 de  $M$  sont des vecteurs orthogonaux à la ligne 3, donc  $M_{13} = M_{23} = 0$ .

Notons maintenant

$$M_2 = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

ce qui nous permet d'écrire

$$M = \begin{pmatrix} M_2 & 0 \\ 0 & 0 & \det(A) \end{pmatrix}.$$

Par conséquent

$$I_3 = {}^t M = \begin{pmatrix} {}^t M_2 M_2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

donc  $M_2 \in O(2)$ . De plus  $\det(A) = \det(M) = \det(M_2) \det(A)$  donc  $\det(M_2) = 1$ . □

Par conséquent, la matrice  $A \in O(3)$  est dans l'une des situations suivantes :

- si  $\det(A) = 1$  alors  $A \in SO(3)$  donc c'est une rotation dans le plan  $\{v\}^\perp$  orthogonal au vecteur propre de la valeur propre 1,
- si  $\det(A) = -1$  alors  $A \notin SO(3)$  et c'est une rotation dans le plan  $\{v\}^\perp$  orthogonal au vecteur propre de la valeur propre  $-1$ , suivie d'une symétrie par rapport au même plan.

En pratique, pour déterminer la transformation il faut déterminer  $v$ , ensuite considérer un vecteur  $u$  orthogonal à  $v$  et calculer l'angle  $\theta$  entre  $u$  et  $f(u)$ . La transformation sera soit une rotation, soit une "rotation + symétrie" selon le signe du  $\det(A)$ . Remarquons enfin que l'on peut calculer l'angle  $\theta$  en remarquant que  $Tr(A) = Tr(M) = \det(A) + 2 \cos(\theta)$ .

### 6.3 Une deuxième preuve du théorème de diagonalisation

L'objectif de cette section est de montrer sous forme d'exercices le théorème de diagonalisation sans avoir recours au théorème de d'Alembert (accepté en cours) affirmant qu'un polynôme de degré  $d$  possède exactement  $d$  racines.

On note  $\|\cdot\|$  la norme euclidienne de  $\mathbb{R}^d$ ,  $d \geq 1$ , et  $(\cdot, \cdot)$  le produit scalaire associé. On considère une matrice  $A$  réelle symétrique  $d \times d$ , et on définit la forme quadratique

$$s(u) := (Au, u), \quad \forall u \in \mathbb{R}^d.$$

On note  $VP(A)$  l'ensemble des valeurs propres **réelles** de  $A$  et on note

$$m := \inf_{u \in \mathbb{R}^d, \|u\|=1} s(u), \quad M := \sup_{u \in \mathbb{R}^d, \|u\|=1} s(u).$$

1. Quelques questions préliminaires.

(a) Soit  $B$  une matrice réelle  $d \times d$ . Montrer qu'il existe une constante  $C \geq 0$  telle que  $\|Bu\| \leq C\|u\|$  pour tout  $u \in \mathbb{R}^d$ .

(b) Montrer que

$$\|u\| = \max_{v \in \mathbb{R}^d, \|v\|=1} (u, v), \quad \forall u \in \mathbb{R}^d.$$

(c) Montrer que si  $\lambda \in VP(A)$  alors  $m \leq \lambda \leq M$ , soit donc  $VP(A) \subset [m, M]$ .

(d) Montrer que si  $s = 0$  alors  $A = 0$ . En déduire que dans ce cas  $VP(A) = \{0\}$ .

(e) Montrer que si  $\{\lambda_1, \dots, \lambda_\ell\} \subset VP(A)$ , avec  $\ell \geq 2$  et les  $\lambda_i$  distincts, et si on définit  $E_i := \text{Ker}(A - \lambda_i I)$ , alors

$$E_\ell \perp (E_1 + \dots + E_{\ell-1}).$$

En déduire que le cardinal de  $VP(A)$  est inférieur ou égal à  $d$ .

2. On définit la forme bilinéaire

$$a(u, v) := (Mu - Au, v).$$

(a) Montrer que  $a$  est une forme positive et symétrique. **Démontrer** que

$$a(u, v) \leq a(u, u)^{1/2} a(v, v)^{1/2}, \quad \forall (u, v) \in \mathbb{R}^d \times \mathbb{R}^d.$$

(b) En déduire que

$$\|Mu - Au\| \leq (M - m)^{1/2} (Mu - Au, u)^{1/2}, \quad \forall u \in \mathbb{R}^d.$$

(c) Montrer qu'il existe une suite  $(u_n)$  de  $\mathbb{R}^d$  telle que  $\|u_n\| = 1$  et  $(Au_n, u_n) \rightarrow M$  lorsque  $n \rightarrow \infty$ . En déduire que

$$\|Mu_n - Au_n\| \rightarrow 0 \quad \text{lorsque } n \rightarrow \infty.$$

- (d) Montrer que si  $MI - A$  était inversible alors la suite  $(u_n)$  vérifierait  $u_n \rightarrow 0$  lorsque  $n \rightarrow \infty$ . En déduire que  $M \in VP(A)$ .
3. Montrer que l'on a l'alternative
- (i) ou bien  $VP(A) = \{M\}$  ;
  - (ii) ou bien  $F := (\text{Ker}(A - MI))^\perp \neq \{0\}$  et
 
$$\sup_{u \in F, \|u\|=1} s(u) < M.$$
4. Conclure que  $VP(A) = \{\lambda_1, \dots, \lambda_k\}$  avec  $1 \leq k \leq d$ ,  $\lambda_1 = M$ ,  $\lambda_k = m$ , et qu'il existe une base orthonormée de  $\mathbb{R}^d$  constituée de vecteurs propres de  $A$ .

## 6.4 Matrice carrée symétrique positive et rayon spectral

L'objectif de cette section est de fournir sous forme d'exercices une caractérisation des matrices carrées symétriques positives et de faire le lien entre norme d'opérateurs et plus grande valeur propre.

Soit  $A$  une matrice carrée réelle de  $M_n(\mathbb{R})$  et  $S := {}^t A A$ .

1. Montrer que  $S$  est symétrique et positive.  
On calcule  ${}^t S = {}^t({}^t A A) = {}^t A A = S$  et  $\langle Su, u \rangle = \langle Su {}^t A A, u \rangle = \langle Au, Au \rangle \geq 0$ .
2. Montrer que  $\text{Ker} S = \text{Ker} A$  puis que  $\text{rg} S = \text{rg} A$ .  
Evidemment  $\text{Ker} A \subset \text{Ker} S$ . Réciproquement, si  $Su = 0$  alors  $0 = \langle Su, u \rangle = \|Au\|^2$  et donc  $Au = 0$ . On en déduit  $\text{rg} S = \text{rg} A$  par le théorème du rang.
3. Montrer que  $S$  est définie positive si, et seulement si  $A$  est inversible.  
 $S$  est définie positive si, et seulement si,  $\text{Ker} S = \{0\}$ , ce qui est équivalent à  $\text{Ker} A = \{0\}$  et donc à  $A$  est inversible.
4. On note  $\|\cdot\|$  la norme d'opérateur sur  $M_n(\mathbb{R})$  associée à la norme euclidienne. Montrer que

$$\|A\| = \sqrt{\|S\|}.$$

Réciproquement, on considère  $S$  une matrice réelle symétrique et positive.

5. Montrer qu'il existe une matrice réelle symétrique et positive  $A$  telle que  $S = A^2$ .  
(Ind. On pensera à utiliser un théorème de diagonalisation que l'on énoncera proprement).
6. Montrer que si  $B$  est une matrice réelle symétrique et positive telle que  $S = B^2$  alors  $\text{Ker}(B - \lambda) \subset \text{Ker}(S - \lambda^2)$  pour toute valeur propre  $\lambda$  de  $B$ . En déduire qu'il y a unicité de la matrice réelle symétrique et positive  $R$  telle que  $S = R^2$ .
7. Avec les notations de la question 4, montrer que

$$\|S\| = \rho(S) := \sup\{\lambda, \lambda \text{ valeur propre de } S\}.$$

(Ind. On pensera à utiliser le même théorème de diagonalisation qu'à la question 5).

## Références

- [1] Joseph Grifone. *Algèbre linéaire*. Éditions Cépaduès, Toulouse, 1990.
- [2] Bernard Guerrien. *Algèbre linéaire pour économistes*. Economica, 1997.
- [3] David Lay. *Algèbre linéaire. Théorie, exercices et applications*. De Boeck, 2004.
- [4] François Liret. *Mathématiques pour le DEUG. Algèbre et Géométrie ; 2ème année*. Dunod, 1999.

# Index

- Angle non orienté, 19
- Anneau, 58
- Annulateur, 9
- Application antilinéaire, 13
- Application bilinéaire, 12
- Application linéaire, 3
- Application semilinéaire, 13
- Application sesquilinéaire, 13
  
- Base, 5
- Base duale, 9
- Base orthogonale, 20
- Base orthonormée, 20
  
- Changement de base, 29, 43
- Combinaison linéaire, 4
- Conditions d'optimalité, 52
- Conique, 52
- Coordonnées, 5
- Corps, 1, 59
  
- Dépendance linéaire, 4
- Dérivée partielle, 50
- Dimension, 6
- Distance, 10
- Droite de régression linéaire, 26
- Dual, 8
  
- Élément inversible, 57
- Élément neutre, 57
- Endomorphisme, 3
- Endomorphisme orthogonal, 31
- Équivalence de normes, 11
- Espace euclidien, 16
- Espace vectoriel, 1
- Espace vectoriel normé, 10
- Exponentielle de matrice, 49
  
- Famille génératrice, 4
- Famille libre, 5
- Forme bilinéaire, 33
- Forme définie positive, 41
- Forme linéaire, 8
- Forme non dégénérée, 13, 41
- Forme positive, 13, 14, 37, 46
- Forme quadratique, 35
- Forme réduite, 39
- Forme symétrique, 13, 34
- Formule de Taylor-Young, 51
  
- Gradient, 51
- Groupe, 57
- Groupe orthogonal, 31, 54
- Groupe spécial orthogonal, 31
  
- Identité du parallélogramme, 18
- Image, 3
- Inégalité de Cauchy-Schwarz, 14
- Inverse, 57
- Isométrie vectorielle, 23
- Isomorphisme, 28
- Isomorphisme d'espaces euclidiens, 23
  
- Loi associative, 57
- Loi commutative, 57
- Loi de composition externe, 57
- Loi de composition interne, 56
  
- Matrice d'une application linéaire, 6
- Matrice d'un endomorphisme, 29
- Matrice d'une forme bilinéaire, 34
- Matrice d'une forme quadratique, 36
- Matrice de passage, 7, 29
- Matrice hessienne, 51
- Matrice orthogonale, 31
- Matrice symétrique, 35, 44
- Matrice unitaire, 31
- Mineurs principaux, 47
- Minimum, maximum, 50
- Minimum, maximum local, 51
  
- Norme, 10

Norme euclidienne, 15  
Norme hermitienne, 15  
Norme induite par le produit scalaire, 15  
Noyau, 3

Orthogonal d'une partie, 24  
Orthogonalité, 16

Problème des moindres carrés, 26  
Procédé d'orthonormalisation de Schmidt, 21,  
25

Produit matriciel, 7  
Produit scalaire, 13, 14  
Produit scalaire canonique, 16, 17  
Produit scalaire hermitien, 14  
Projection orthogonale, 24  
Puissance d'une matrice, 48

Récurrence linéaire, 48  
Réduction de Gauss, 37

Signature, 40  
Somme directe  $\oplus$ , 3  
Sous-espace engendré, 4  
Sous-espace propre, 44  
Sous-espace vectoriel, 2  
Sous-espaces supplémentaires, 3  
Symbole de Kronecker  $\delta_{ij}$ , 9  
Symétrie hermitienne, 14  
Système différentiel, 49

Théorème de diagonalisation, 45  
Théorème de la base incomplète, 6  
Théorème de Pythagore, 19  
Théorème de Sylvester, 47  
Théorème du rang, 6  
Transposée, 16

Valeur propre, 43  
Vecteur propre, 43