

# **Algèbre, géométrie, transformation, calcul numérique**

Benjamin Charlier

1<sup>er</sup> mars 2025

# Table des matières

<b>1</b>	<b>Géométrie vectorielle</b>	<b>5</b>
1.1	Espace vectoriels, sous espaces vectoriels et bases . . . . .	5
1.1.1	Définition et exemples . . . . .	5
1.1.2	Bases et dimension . . . . .	6
1.2	Espaces affines . . . . .	7
1.2.1	Points et vecteurs . . . . .	7
1.2.2	Repère . . . . .	8
1.2.3	Droite et plan affines . . . . .	9
1.3	Calcul Vectoriel . . . . .	9
1.3.1	Produit scalaire et espace euclidien . . . . .	9
1.3.2	Orientation . . . . .	10
1.3.3	Produit mixte et produit vectoriel . . . . .	11
1.3.4	Produit vectoriel dans l'espace . . . . .	12
1.3.5	Produit mixte dans l'espace . . . . .	14
1.4	Droites, plans, cercles et sphères . . . . .	14
1.4.1	Droites dans le plan . . . . .	15
1.4.2	Cercles dans le plan . . . . .	15
1.4.3	Plans dans l'espace . . . . .	15
1.4.4	Droites dans l'espace . . . . .	16
1.4.5	Sphères dans l'espace . . . . .	16
<b>2</b>	<b>Transformation linéaires</b>	<b>17</b>
2.1	Définitions et premières propriétés . . . . .	17
2.1.1	Définitions . . . . .	17
2.1.2	Noyau et image . . . . .	18
2.2	Matrices et applications linéaires . . . . .	19
2.3	Inversion de matrices . . . . .	22
2.3.1	Déterminant . . . . .	22
2.3.2	Calcul de l'inverse en pratique . . . . .	23
2.3.3	Les matrices orthogonales . . . . .	24
2.3.4	Transformer un champs de vecteur normal . . . . .	26
2.4	Changement de base . . . . .	26
2.4.1	Pour un vecteur . . . . .	26
2.4.2	Pour les applications linéaires . . . . .	28
2.4.3	Matrices semblables . . . . .	30
<b>3</b>	<b>Réduction des endomorphismes</b>	<b>31</b>
3.1	Valeurs propres et espaces propres . . . . .	31
3.1.1	Valeurs propres et polynôme caractéristique . . . . .	31

3.1.2	Théorème de Cayley-Hamilton . . . . .	33
3.2	Diagonalisation . . . . .	34
3.2.1	Condition suffisante . . . . .	34
3.2.2	Condition nécessaire et suffisante . . . . .	35
3.3	Triangularisation . . . . .	36
3.4	Projecteurs . . . . .	37
<b>4</b>	<b>ACP</b> . . . . .	<b>41</b>
4.1	Matrices symétriques . . . . .	41
4.2	Décomposition en valeurs singulières (SVD) . . . . .	41
4.2.1	Valeurs singulières et vecteurs singuliers . . . . .	41
4.2.2	La SVD réduite . . . . .	42
4.3	Meilleure approximation . . . . .	43
4.4	Pseudo-inverse . . . . .	44
4.5	Analyse en composante principale (ACP ou PCA) . . . . .	45
4.5.1	Interprétation et régression . . . . .	47



# Chapitre 1

## Géométrie vectorielle

### 1.1 Espace vectoriels, sous espaces vectoriels et bases

#### 1.1.1 Définition et exemples

Qu'est ce qu'il y a de commun entre

1. Les  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ ,
2. L'ensemble des suites (l'ensemble  $\mathbb{R}^{\mathbb{N}}$ ),
3. Les fonctions de  $\mathbb{R} \rightarrow \mathbb{R}$  (l'ensemble  $\mathbb{R}^{\mathbb{R}}$ ) ?

Le fait que l'on puisse ajouter les éléments dans chacun de ces espaces, mais aussi les multiplier par un scalaire. En bref, la **structure d'espace vectoriel**.

**Définition 1.1.1.** Un  $\mathbb{R}$  **espace vectoriel** est un triplet  $(E, +, \cdot)$  formé :

- d'un ensemble  $E$  dont tous les éléments sont appelés vecteurs ;
- d'une loi d'addition qui est une application

$$\begin{aligned} + : E \times E &\rightarrow E \\ (u, v) &\mapsto u + v \end{aligned}$$

- d'une loi de multiplication par un scalaire

$$\begin{aligned} \cdot : \mathbb{R} \times E &\rightarrow E \\ (\lambda, v) &\mapsto \lambda v \end{aligned}$$

tels que

1.  $(E, +)$  est associatif, commutatif, existence d'un élément neutre et d'un opposé (groupe abélien)
2.  $\forall (\lambda, \mu) \in \mathbb{R}^2$  et  $\forall (u, v) \in E^2$  on a

$$\left\{ \begin{array}{l} (\lambda\mu)u = \lambda(\mu u) \\ (\lambda + \mu)u = \lambda u + \mu u \\ \lambda(u + v) = \lambda u + \lambda v \\ 1u = u \end{array} \right.$$

**Exemple 1.1.1.** 1. Soient  $a_0, a_1, \dots, a_n \in \mathbb{R}$ . On considère  $F$  l'ensemble des fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$  telles que  $f(a_0) = f(a_1) = \dots = f(a_n) = 0$ .  
2.  $G$ , l'ensemble des solutions de l'équation différentielle  $y' = y$ .

3.  $H$ , l'ensemble des matrices  $3 \times 3$  triangulaires supérieures.
4.  $E = \{P \in \mathbb{R}_3[X] \mid P(X^2) = X^2P(X)\}$ .

Pour vérifier que les exemples précédemment cités conviennent, il faut vérifier les 8 axiomes, ce qui n'est pas forcément difficile mais un peu long. Pour montrer qu'un espace est un espace vectoriel, on préfère souvent montrer que c'est un sous-espace vectoriel.

**Définition 1.1.2.** Soit  $(E, +, \cdot)$  un espace vectoriel et  $F \subset E$ .  $F$  est appelé **sous-espace vectoriel (sev)** de  $E$  si  $\forall (u, v) \in F^2$  et  $(\lambda, \mu) \in \mathbb{R}^2$  alors  $\lambda u + \mu v \in F$ . On note  $(F, +, \cdot)$  cet espace vectoriel.

**Définition 1.1.3.** — Étant donné des vecteurs  $u_1, \dots, u_p$  de  $E$ , on appelle combinaison linéaire à coefficients réels de  $u_1, \dots, u_p$  tout vecteur de la forme

$$\lambda_1 u_1 + \dots + \lambda_p u_p \quad \text{avec } \lambda_1, \dots, \lambda_p \in \mathbb{R}.$$

- L'espace des combinaisons linéaires finies de  $u_1, \dots, u_p$  sera noté  $\text{Vect}(u_1, \dots, u_p)$  et appelé espace engendré par  $u_1, \dots, u_p$ .
- Plus généralement, si  $A$  est une partie non vide, finie ou infinie de  $E$ ,  $\text{Vect}(A)$  est l'ensemble des combinaisons linéaires d'éléments de  $A$  et est appelé espace engendré par  $A$ .

## 1.1.2 Bases et dimension

**Définition 1.1.4.** Une famille  $(u_1, \dots, u_p)$  de vecteurs de  $E$  telle que  $\text{Vect}(u_1, \dots, u_p) = E$  est appelée **famille génératrice** de  $E$  (i.e. si tout  $v \in E$  peut s'exprimer comme combinaison linéaire de  $u_1, \dots, u_p$ ).

**Définition 1.1.5.** On dit qu'une famille  $(u_1, \dots, u_p)$  de vecteurs de  $E$  est une **famille libre** de  $E$  si tout vecteur de  $\text{Vect}(u_1, \dots, u_p)$  s'exprime de manière unique comme combinaison linéaire de  $u_1, \dots, u_p$ . Cela revient à dire que

$$(\forall (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p : \lambda_1 u_1 + \dots + \lambda_p u_p = 0) \Leftrightarrow \lambda_1 = \dots = \lambda_p = 0.$$

Une famille de vecteurs qui n'est pas libre est dite **liée**. Dans ce cas, au moins l'un de ses éléments est combinaison linéaire des autres (ce qui revient à dire que  $\det(u_1, \dots, u_p) = 0$ ). Nous sommes maintenant en mesure de définir la notion de base dans un espace vectoriel.

**Définition 1.1.6.** On dit qu'une famille  $\mathcal{B} = (u_1, \dots, u_p)$  est une **base d'un espace vectoriel**  $E$  si c'est une famille libre et génératrice de  $E$ . C'est à dire si tout vecteur de  $E$  s'écrit de manière unique comme linéaire des vecteurs de la base  $\mathcal{B}$ , i.e.,

$$\forall v \in E, \exists! (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p \text{ t.q. } v = \lambda_1 u_1 + \dots + \lambda_p u_p.$$

**Théorème 1.1.1.** Soit  $E$  un espace vectoriel de dimension finie. Toute famille libre  $C = (u_1, \dots, u_p)$  de vecteurs de  $E$  peut être complétée en une base  $\mathcal{B} = (u_1, \dots, u_p, u_{p+1}, \dots, u_n)$  de  $E$ .

**Définition - Proposition 1.1.1.** Dans un espace vectoriel de dimension finie, toutes les bases ont le même nombre d'éléments. Ce nombre est appelé **dimension de l'espace vectoriel**  $E$  et est noté  $\dim(E)$ .

**Proposition 1.1.1.** Dans un espace vectoriel de dimension  $n$  :

1. Toute famille libre  $(u_1, \dots, u_n)$  de  $n$  vecteurs de  $E$  est une base (famille libre maximale).
2. Toute famille génératrice  $(u_1, \dots, u_n)$  de  $n$  vecteurs de  $E$  est une base (famille génératrice minimale).

**Proposition 1.1.2.**

1. Un sous-espace  $F$  d'un espace vectoriel  $E$  de dimension finie est de dimension finie.
2. On a alors  $\dim(F) \leq \dim(E)$ .
3. Si  $\dim(F) = \dim(E)$ , on a  $F = E$ .

On peut construire une infinité de bases d'un même espace vectoriel  $E$ . Suivant le problème auquel nous sommes confrontés, certaines bases se prêteront mieux à sa résolution. Il est donc non seulement important de savoir gérer les changements de bases mais aussi de trouver des bases "pertinentes" suivant le problème considéré.

**Définition 1.1.7.** Soient  $F_1$  et  $F_2$  deux sous-espaces vectoriels d'un espace vectoriel  $E$ . On dit que  $F_1$  et  $F_2$  sont en **somme directe** si, pour tout élément  $u$  de la somme  $F_1 + F_2 = \{v \in E \mid v = u_1 + u_2, u_1 \in F_1, u_2 \in F_2\}$ , il existe un unique couple  $(u_1, u_2)$  de  $F_1 \times F_2$  tel que  $u = u_1 + u_2$ .

**Exemple 1.1.2. Décomposition d'un espace vectoriel en somme directe** Démontrer que l'ensemble  $P$  des fonctions paires et l'ensemble des fonctions impaires  $I$  sont deux espaces supplémentaires de  $\mathbb{R}^{\mathbb{R}}$ . *Note : on pourra faire un lien utile avec les séries de Fourier...*

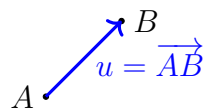
## 1.2 Espaces affines

### 1.2.1 Points et vecteurs

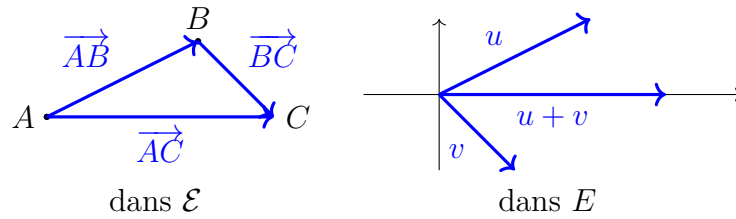
On modélise le plan et l'espace par des espaces affines de dimension 2 et 3 respectivement. Pour l'espace à 3 dimensions on considère l'ensemble  $\mathbb{R}^3$  des triplets de réels :

1. Ensemble  $\mathcal{E}$  des points : On peut décrire un point  $M$  de l'espace (*i.e.* une **position**) par un triplet  $(x, y, z) \in \mathbb{R}^3$  de réels.
2. Ensemble  $E$  de vecteurs : on peut aussi voir  $E = \mathbb{R}^3$  comme un espace vectoriel. Un vecteur  $u \in E$  modélise alors un "**déplacement**" entre 2 points  $A$  et  $B$  de  $\mathcal{E}$ . Si  $A = (a_1, a_2, a_3)$  et  $B = (b_1, b_2, b_3)$  on définit le vecteur  $u = \overrightarrow{AB}$  comme étant le triplet de réels  $u = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$ . On a alors :

(a) Pour tout  $u \in E$  et  $(A, B) \in \mathcal{E} : u + A = \overrightarrow{AB} + A = B \in \mathcal{E}$



(b) Relation de Chasles :  $\forall A, B, C \in \mathcal{E}$  on a  $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ .



En résumé, les éléments de  $\mathcal{E}$  sont considérés comme des points (notés en majuscules) sur lesquels opèrent des vecteurs (notés en minuscules) de  $E$ .

## 1.2.2 Repère

**Définition 1.2.1.** Un **repère cartésien** de l'espace  $\mathcal{E}$  est la donnée d'un point  $\Omega$  (l'origine du repère) et de trois vecteurs  $(e_1, e_2, e_3)$  formant une base (*i.e.* une famille libre génératrice) de l'espace  $E$ .

**Définition 1.2.2.** Si  $M$  est un point de  $\mathcal{E}$  alors le vecteur  $\overrightarrow{\Omega M}$  se décompose de façon unique sur les vecteurs de base :

$$\overrightarrow{\Omega M} = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

Le triplet  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\mathcal{R}}$  (ou plus simplement  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$  si il n'y a pas d'ambiguïtés) contient les **coordonnées cartésiennes** du point  $M$  dans le repère  $\mathcal{R}$ .

**Remarque 1.** Une fois une origine  $\Omega$  de repère choisie, on peut identifier un point  $M$  de  $\mathcal{E}$  avec le vecteur  $\overrightarrow{\Omega M} \in E$ . Alors, le choix d'un système de coordonnées permet de ramener tous les calculs à des calculs dans  $\mathbb{R}^3$ . Lorsque l'on ne précise pas la base, c'est que l'on se place dans l'espace vectoriel  $E = \mathbb{R}^3$  muni de la base canonique composée des vecteurs  $i = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $j = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  et  $k = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ .

**Remarque 2. Changement de repère :** Soit deux repères cartésiens  $\mathcal{R} = (\Omega, e_1, e_2, e_3)$  et  $\mathcal{R}' = (\Omega', e'_1, e'_2, e'_3)$  et un point  $M$ . On note :

- $\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{R}}$  les coordonnées de  $M$  dans  $\mathcal{R}$  et  $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\mathcal{R}'}$  les coordonnées de  $M$  dans  $\mathcal{R}'$ .
- $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}_{\mathcal{R}}$  les coordonnées de  $\Omega'$  dans  $\mathcal{R}$ .
- $\begin{pmatrix} a_{1,i} \\ a_{2,i} \\ a_{3,i} \end{pmatrix}_{\mathcal{R}}$  les coordonnées des  $e'_i$  (pour  $i = 1, 2, 3$ ) dans  $\mathcal{R}$ .

Alors les coordonnées du point  $M$  dans le repère  $\mathcal{R}$  s'expriment en fonction des coordonnées de  $M$  dans le repère  $\mathcal{R}'$  :

$$\begin{cases} x = \alpha + a_{1,1}x' + a_{1,2}y' + a_{1,3}z' \\ y = \beta + a_{2,1}x' + a_{2,2}y' + a_{2,3}z' \\ z = \gamma + a_{3,1}x' + a_{3,2}y' + a_{3,3}z' \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{R}} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}_{\mathcal{R}} + \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}_{\mathcal{R} \leftarrow \mathcal{R}'} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\mathcal{R}'}$$

**Remarque 3.** A propos des notations : un élément de  $\mathbb{R}^n$  sera noté par convention  $(x_1, \dots, x_n)$  (c'est la donnée de  $n$  réels). Lorsque l'on note  $(x_1 \ \dots \ x_n)$  (resp.  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ) on manipule en fait une matrice ligne (resp. colonne) dont les entrées sont les coordonnées d'un élément  $(\tilde{x}_1, \dots, \tilde{x}_n)$  dans une base donnée (ainsi lorsque les virgules disparaissent il y eut choix de base...). En général, les matrices colonnes représentent des vecteurs et les matrices lignes des formes linéaires (des fonctions linéaires à valeurs réelles).



### 1.2.3 Droite et plan affines

**Définition 1.2.3.** Soit  $\mathcal{E}$  un espace affine dirigé par un espace vectoriel  $E$  :

1. Dans  $E$  : la **droite vectorielle** engendrée par un vecteur  $u \in E$  non-nul est l'ensemble des vecteurs colinéaires à  $u$  :

$$\text{Vect}(u) = \{v \in E, \exists \lambda \in \mathbb{R}, v = \lambda u\} = \mathbb{R}u.$$

Si deux vecteurs  $(u, v) \in E^2$  sont non-colinéaires, le **plan vectoriel** engendré par  $(u, v)$  est l'ensemble des combinaisons linéaires de  $u$  et  $v$  :

$$\text{Vect}(u, v) = \{w \in E, \exists \lambda, \mu \in \mathbb{R}, w = \lambda u + \mu v\}.$$

2. Dans  $\mathcal{E}$  : la **droite affine**  $\mathcal{D}$  passant par un point  $A$  et dirigée par un vecteur non-nul  $u \in E$  est l'ensemble des points

$$\mathcal{D} = \{D \in \mathcal{E} \mid D = A + \lambda u, \lambda \in \mathbb{R}\} \subset \mathcal{E}.$$

On note  $\mathcal{D} = A + \text{Vect}(u)$  et on dit que  $\text{Vect}(u)$  est la direction de  $\mathcal{D}$ . De même, le **plan affine**  $\mathcal{P}$  dirigé par  $u, v$  est

$$\mathcal{P} = \{A + (\lambda u + \mu v), \lambda, \mu \in \mathbb{R}\} = A + \text{Vect}(u, v) \subset \mathcal{E}.$$

## 1.3 Calcul Vectoriel

### 1.3.1 Produit scalaire et espace euclidien

**Définition 1.3.1.** Soit  $u = (u_1, u_2, u_3)$  et  $v = (v_1, v_2, v_3)$  deux vecteurs de  $\mathbb{R}^3$ . Le **produit scalaire** de  $u$  et  $v$  est

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

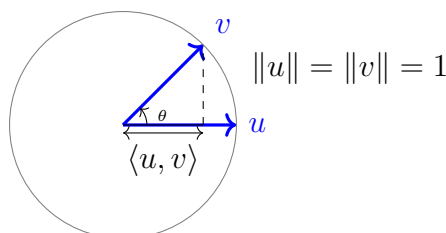
et on définit la **norme** d'un vecteur  $u = (u_1, u_2, u_3)$  par

$$\|u\|^2 = u_1^2 + u_2^2 + u_3^2.$$

**Remarque 4.** 1. Les définitions s'étendent sans problème à  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$

2. Le produit scalaire et la norme (au carré) sont donc des relations polynômiales homogènes (de degré 2) des coordonnées de  $u$  et  $v$ .
3. On peut définir des produits scalaires "à poids" (forme bilinéaire, symétrique, définie positive)

**Interprétation géométrique :** Soit  $u$  et  $v$  deux vecteurs de  $\mathbb{R}^3$  dont la norme est égale à 1 :



Le produit scalaire de deux vecteurs unitaires est donc le cosinus de l'angle  $\theta$  entre les vecteurs. De plus, on a la formule suivante pour tout  $u, v \in \mathbb{R}^3$  :

$$\langle u, v \rangle = \|u\| \|v\| \cos(\theta).$$

En découle l'**inégalité de Cauchy-Swarz**

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

avec cas d'égalité si  $u$  et  $v$  sont colinéaires. Rigoureusement, on démontre cette majoration à partir des propriétés algébrique du produit scalaire :

**Proposition 1.3.1.** Le produit scalaire satisfait aux règles de calculs suivantes :

1. **Bilinéarité** : soient  $u, v, w \in \mathbb{R}^3$  et  $\lambda, \mu \in \mathbb{R}$  :
  - (a)  $\langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle$ ,
  - (b)  $\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle$ .
2. **Symétrie** : soient  $u, v \in \mathbb{R}^3$   $\langle u, v \rangle = \langle v, u \rangle$ .
3. **Positivité** : soient  $u \in \mathbb{R}^3$   $\langle u, u \rangle \geq 0$  et  $\langle u, u \rangle = 0 \Leftrightarrow u = 0$ .

**Définition 1.3.2.** 1. On dit que deux vecteurs sont **orthogonaux** lorsque leur produit scalaire est nul.

2. On dit qu'un vecteur est **unitaire** lorsque sa norme vaut 1.
3. On dit qu'une base est **orthonormale** lorsque les trois vecteurs de la base sont orthogonaux deux à deux et unitaires. On utilise l'acronyme "B.O.N." pour base orthonormale.
4. On dit qu'un repère est **orthonormé** lorsque sa base est orthonormale.

**Proposition 1.3.2 coordonnées d'un vecteur dans une B.O.N..** Dans une base orthonormale  $(e_1, e_2, e_3)$  de  $\mathbb{R}^3$ , un vecteur  $u \in \mathbb{R}^3$  se décompose sous la forme

$$u = \langle e_1, u \rangle e_1 + \langle e_2, u \rangle e_2 + \langle e_3, u \rangle e_3.$$

**Proposition 1.3.3 Calcul du produit scalaire dans une B.O.N..** Si  $(e_1, e_2, e_3)$  est une base orthonormale quelconque de  $\mathbb{R}^3$  et si  $u = x_1 e_1 + x_2 e_2 + x_3 e_3$  et  $v = y_1 e_1 + y_2 e_2 + y_3 e_3$  sont deux vecteurs de  $\mathbb{R}^3$ , alors le produit scalaire de  $u$  et  $v$  s'écrit :

$$\langle u, v \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

**Définition 1.3.3.** On définit la **distance euclidienne** entre deux points  $A$  et  $B$  de l'espace  $\mathcal{E}$  par :

$$d(A, B) = \|\overrightarrow{AB}\|.$$

Si  $\mathcal{R}$  est un repère orthonormé on a

$$d(A, B) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

où  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$  sont les coordonnées dans  $\mathcal{R}$  de  $A$  et de  $B$  respectivement.

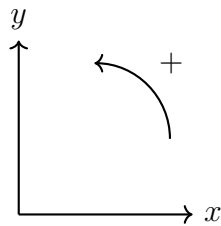
## 1.3.2 Orientation

Un repère orthonormé  $\mathcal{R} = (\Omega, e_1, e_2, e_3)$  étant choisi, on rappelle que, pour exprimer les coordonnées d'un point dans un autre repère  $\mathcal{R}' = (\Omega', e'_1, e'_2, e'_3)$  il est nécessaire d'utiliser la

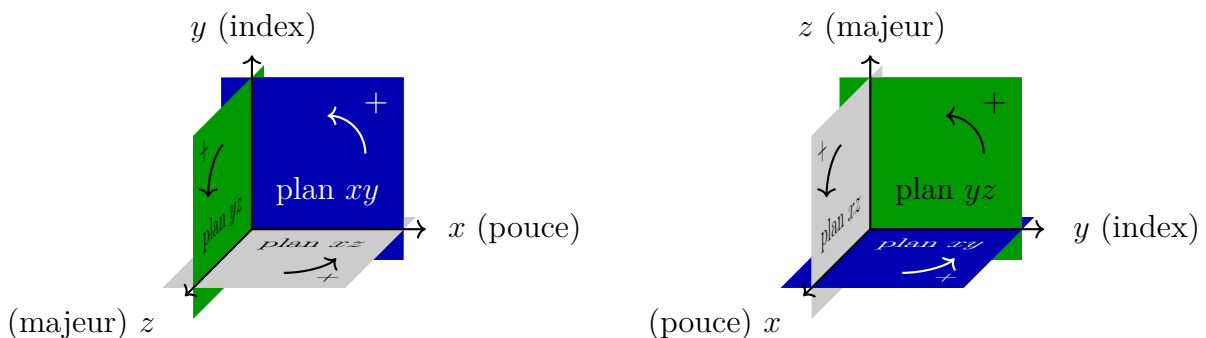
matrice de passage  $B$  de la base  $(e_1, e_2, e_3)$  à la base  $(e'_1, e'_2, e'_3)$  (voir la Remarque 2). Si la nouvelle base est orthonomée, on peut montrer que  $\det(B) \in \{1, -1\}$  et a donc seulement deux valeurs possibles qui correspondent à deux orientations de l'espace.

**Remarque 5.** D'un point de vue pratique, ceci s'illustre par le fait qu'un contorsionniste aussi doué qu'il soit ne pourra pas superposer ses deux mains, droite et gauche.

Il faut alors choisir une convention pour "orienter l'espace" et on se contente ici de décrire l'une des règles classiques dite "règle des 3 doigts de la main droite" (il en existe d'autres comme la règle du "bonhomme d'ampère" ou la règle du "tire bouchon" qui suivent toutes la même convention) : on dit qu'un repère est **direct** si le pouce, l'index et le majeur de la main droite peuvent être placés (direction et sens) suivant les vecteurs  $(e_1, e_2, e_3)$ .



Repère direct du plan



Repères directs de l'espace

Une permutation circulaire de trois vecteurs ne modifie pas son orientation et les deux repères ci-dessus sont directs. Si on change l'orientation d'un des vecteurs, on change l'orientation du repère (orientation **indirecte**). Dans la suite de ce cours nous considérerons toujours des repères directs.

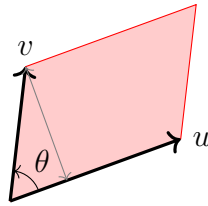
### 1.3.3 Produit mixte et produit vectoriel

**Définition 1.3.4.** Soit  $u = (u_1, u_2)$  et  $v = (v_1, v_2)$  deux vecteurs de  $\mathbb{R}^2$ . On appelle **déterminant** ou **produit mixte** des deux vecteurs, le nombre réel

$$\det(u, v) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = u_1 v_2 - u_2 v_1.$$

**Remarque 6.** Même remarque que pour le produit scalaire. On a toujours affaire à des polynômes homogènes de degré 2.

**Interprétation géométrique :** Le produit mixte  $\det(u, v)$  de deux vecteurs  $u, v$  de  $\mathbb{R}^2$  représente l'aire orientée du parallélogramme s'appuyant sur les deux vecteurs :



De plus, si  $\theta$  désigne l'angle orienté entre  $u$  et  $v$  on a

$$\det(u, v) = \|u\| \|v\| \sin(\theta).$$

En découle la propriété suivante :

$$|\det(u, v)| \leq \|u\| \|v\|$$

avec égalité si et seulement si les vecteurs sont orthogonaux. Ici aussi, la démonstration rigoureuse de ces relations se fait à partir des propriétés algébriques du produit mixte. En voici quelques unes :

**Proposition 1.3.4.** Le produit mixte satisfait aux règles de calcul suivantes :

1. **Bilinéarité** : soient  $u, v, w \in \mathbb{R}^2$  et  $\lambda, \mu \in \mathbb{R}$  :
  - (a)  $\det(\lambda u + \mu v, w) = \lambda \det(u, w) + \mu \det(v, w)$ ,
  - (b)  $\det(u, \lambda v + \mu w) = \lambda \det(u, v) + \mu \det(u, w)$ .
2. **Antisymétrie** : soient  $u, v \in \mathbb{R}^2$  on a  $\det(u, v) = -\det(v, u)$ .

**Remarque 7.** 1. Attention donc à l'ordre d'écriture des facteurs dans le produit mixte !  
 2. L'Antisymétrie implique que pour tout  $u \in \mathbb{R}^2$  on a  $\det(u, u) = 0$  (le parallélogramme est plat!). Voir la proposition suivante :

**Proposition 1.3.5 Condition de colinéarité.** Deux vecteurs  $u$  et  $v$  du plan sont liés si et seulement si  $\det(u, v) = 0$ .

**Remarque 8.** Soit  $u = (u_1, u_2) \in \mathbb{R}^2$  un vecteur non normé. Alors le vecteur  $u^\perp = (-u_2, u_1)$  est également un vecteur normé tel que  $\langle u, u^\perp \rangle = 0$ . On en déduit que  $(u, u^\perp)$  forme une base orthonormée de  $\mathbb{R}^2$ .

Si  $(e_1, e_2)$  est une base orthonormée de  $\mathbb{R}^2$  on peut montrer qu'on a deux possibilités :  $e_2 = e_1^\perp$  ou  $e_2 = -e_1^\perp$ . Dans le premier cas, on dira que la base  $(e_1, e_2)$  est directe.

**Proposition 1.3.6 (Calcul du produit mixte dans une B.O.N. directe).** Si  $(e_1, e_2)$  est une base orthonormale directe de  $\mathbb{R}^2$  et si  $u = x_1 e_1 + x_2 e_2$  et  $v = y_1 e_1 + y_2 e_2$  sont deux vecteurs du plan, alors le produit mixte est

$$\det(u, v) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

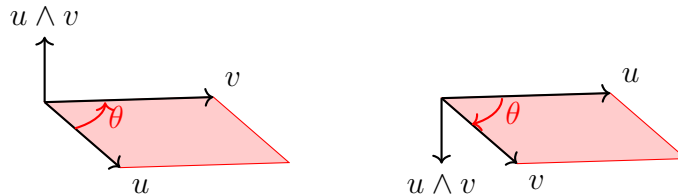
### 1.3.4 Produit vectoriel dans l'espace

**Définition 1.3.5.** On appelle **produit vectoriel** de deux vecteurs  $u = (u_1, u_2, u_3)$  et  $v =$

$(v_1, v_2, v_3)$  de  $\mathbb{R}^3$  le vecteur :

$$u \wedge v = \left( \begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \right).$$

**Interprétation géométrique :** Le produit vectoriel  $u \wedge v$  est orthogonal au plan engendré par  $u$  et  $v$  et sa norme est égale à l'aire du parallélogramme engendré par  $u$  et  $v$ .



On a  $\|u \wedge v\| = \|u\| \|v\| |\sin(\theta)|$  et l'**identité de Lagrange**

$$\|u \wedge v\|^2 + |\langle u, v \rangle|^2 = \|u\|^2 \|v\|^2.$$

On a encore des propriétés algébriques similaires car le calcul du produit vectoriel met en jeu des polynômes homogènes de degré 2 en les coordonnées.

**Proposition 1.3.7.** Le produit vectoriel satisfait aux règles de calculs suivantes :

1. **Bilinéarité :** soient  $u, v, w \in \mathbb{R}^3$  et  $\lambda, \mu \in \mathbb{R}$  :
  - (a)  $(\lambda u + \mu v) \wedge w = \lambda(u \wedge w) + \mu(v \wedge w)$ ,
  - (b)  $u \wedge (\lambda v + \mu w) = \lambda(u \wedge v) + \mu(u \wedge w)$ .
2. **Antisymétrie :**  $u \wedge v = -v \wedge u$ .

**Proposition 1.3.8.** On a les propriétés suivantes :

1. Le produit vectoriel est nul si et seulement si les vecteurs sont colinéaires.
2. Le produit vectoriel est un vecteur orthogonal aux deux vecteurs. C'est à dire que pour tout  $u, v \in \mathbb{R}^3$   $\langle u, u \wedge v \rangle = \langle v, u \wedge v \rangle = 0$ .

**Définition 1.3.6.** Soit  $\mathcal{B} = (e_1, e_2, e_3)$  une base orthonormale de  $\mathbb{R}^3$ . On dit que  $\mathcal{B}$  est **directe** si  $e_1 \wedge e_2 = e_3$ .

**Remarque 9.** Dans une base orthonormale directe on a les égalités suivantes :

- $(e_1 \wedge e_2) \wedge e_3 = (e_3 \wedge e_1) \wedge e_2 = (e_1 \wedge e_3) \wedge e_2 = \dots = 0$ ,
- $(e_1 \wedge e_2) \wedge e_1 = e_2, (e_1 \wedge e_2) \wedge e_2 = -e_1, \dots$

**Proposition 1.3.9 Calcul du produit vectoriel dans une B.O.N. directe.** Si  $(e_1, e_2, e_3)$  est une base orthonormale directe de  $\mathbb{R}^3$ , et si  $u = x_1 e_1 + x_2 e_2 + x_3 e_3$  et  $v = y_1 e_1 + y_2 e_2 + y_3 e_3$  sont deux vecteurs de  $\mathbb{R}^3$  alors

$$u \wedge v = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} e_3.$$

**Remarque 10.** Le produit vectoriel n'est pas associatif : en général, on a  $(u \wedge v) \wedge w \neq u \wedge (v \wedge w)$ . Plus précisément, on a la **formule du double produit vectoriel** suivante :

$$u \wedge (v \wedge w) = \langle u, w \rangle v - \langle u, v \rangle w$$

et

$$(u \wedge v) \wedge w = \langle u, w \rangle v - \langle v, w \rangle u.$$

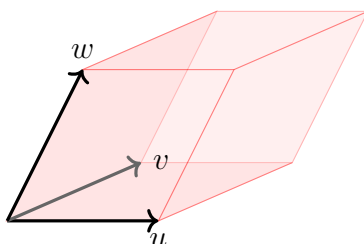
En d'autres termes,  $u \wedge (v \wedge w)$  appartient au plan engendré par  $v$  et  $w$  et  $(u \wedge v) \wedge w$  au plan engendré par  $u$  et  $v$ .

### 1.3.5 Produit mixte dans l'espace

**Définition 1.3.7.** On appelle **produit mixte** (ou **déterminant**) de trois vecteurs  $u, v, w \in \mathbb{R}^3$ , le nombre réel :

$$\det(u, v, w) = \langle (u \wedge v), w \rangle.$$

**Interprétation géométrique :** Le déterminant représente le volume orientée du parallélépipède construit sur  $u, v$  et  $w$ .



Le déterminant possède des propriétés algébriques simples (ici, le calcul met en jeu des polynômes homogènes de degré 3).

**Proposition 1.3.10.** Le produit mixte satisfait aux règles de calcul suivantes :

1. **Trilinéarité :** le produit mixte est linéaire par rapport à chacun des trois vecteurs.
2. **Antisymétrie :** en permutant 2 vecteurs, on change le produit mixte en son opposé.

**Proposition 1.3.11 Condition de coplanarité.** Trois vecteurs de  $\mathbb{R}^3$  sont coplanaires si et seulement si leur produit mixte est nul.

**Proposition 1.3.12 Calcul du produit mixte dans une B.O.N. directe - Règle de Sarrus.** Si  $(e_1, e_2, e_3)$  est une base orthonormale directe, et si  $u = x_1e_1 + x_2e_2 + x_3e_3$ ,  $v = y_1e_1 + y_2e_2 + y_3e_3$  et  $w = z_1e_1 + z_2e_2 + z_3e_3$  alors

$$\begin{aligned} \det(u, v, w) &= \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = z_1 \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - z_2 \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \\ &= x_1y_2z_3 + y_1z_2x_3 + z_1x_2y_3 - x_3y_2z_1 - y_3z_2x_1 - z_3x_2y_1. \end{aligned}$$

## 1.4 Droites, plans, cercles et sphères

La **distance euclidienne** entre un point  $A$  et un ensemble  $\mathcal{C}$  est définie par :

$$\text{dist}(A, \mathcal{C}) = \inf_{P \in \mathcal{C}} \|\overrightarrow{AP}\|$$

En générale, aucun point ou plusieurs points  $P$  de  $\mathcal{C}$  peuvent réaliser cette distance. Mais, si un point  $P$  réalise cette distance alors  $P$  est un projeté orthogonale de  $A$  sur  $\mathcal{C}$ .

### 1.4.1 Droites dans le plan

Plusieurs manières équivalentes de définir une droite passant par  $A = (\alpha, \beta)$  :

1. Dirigée par un vecteur  $u \in \mathbb{R}^2$  (car c'est un sous espace affine de dimension 1)

$$\mathcal{D} = \left\{ (x, y) \in \mathbb{R}^2 \mid \det(u, (x - \alpha, y - \beta)) = 0 \right\}.$$

2. Normale à un vecteur  $n \in \mathbb{R}^2$  (car c'est un sous espace affine de codimension 1)

$$\mathcal{D} = \left\{ (x, y) \in \mathbb{R}^2 \mid \langle n, (x - \alpha, y - \beta) \rangle = 0 \right\}.$$

**Proposition 1.4.1.** La distance entre un point  $M$  et une droite  $\mathcal{D}$  passant par les points  $A$  et  $B$

$$\text{dist}(M, \mathcal{D}) = \frac{|\det(\overrightarrow{AB}, \overrightarrow{AM})|}{\|\overrightarrow{AB}\|}.$$

### 1.4.2 Cercles dans le plan

**Définition 1.4.1.** Le cercle  $\mathcal{C}$  de centre  $\mathcal{O} = (x_0, y_0) \in \mathbb{R}^2$  et de rayon  $r > 0$  est

$$\begin{aligned} \mathcal{C} &= \left\{ M \in \mathbb{R}^2 \mid d(M, \mathcal{O}) = r \right\} \\ &= \left\{ (x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = r^2 \right\}. \end{aligned}$$

Autrement dit, un cercle est l'ensemble de niveau  $r$  pour la distance euclidienne à son centre.

### 1.4.3 Plans dans l'espace

Plusieurs manières équivalentes de définir un plan passant par  $A = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$  :

1. Dirigé par deux vecteurs non colinéaires  $u, v \in \mathbb{R}^3$  (car c'est un sous espace affine de dimension 2)

$$\mathcal{P} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \det \left( \begin{pmatrix} x - \alpha \\ y - \beta \\ z - \gamma \end{pmatrix}, u, v \right) = 0 \right\}.$$

2. Normal à un vecteur  $n \in \mathbb{R}^3$  (car c'est un sous espace affine de codimension 1)

$$\mathcal{P} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \left\langle n, \begin{pmatrix} x - \alpha \\ y - \beta \\ z - \gamma \end{pmatrix} \right\rangle = 0 \right\}.$$

**Proposition 1.4.2.** La distance entre un point  $M$  et un plan  $\mathcal{P}$  passant par les points non alignés  $A, B$  et  $C$

$$\text{dist}(M, \mathcal{P}) = \frac{|\det(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AM})|}{\|\overrightarrow{AB} \wedge \overrightarrow{AC}\|}.$$

### 1.4.4 Droites dans l'espace

Plusieurs manières équivalentes de définir une droite passant par  $A = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$  :

1. Dirigée par un vecteur  $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3$  (car c'est un sous espace affine de dimension 1)

$$\mathcal{D} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \begin{cases} x = u_1 t + \alpha \\ y = u_2 t + \beta \\ z = u_3 t + \gamma \end{cases}, t \in \mathbb{R} \right\}.$$

2. Intersection de deux plans passant par  $A = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$  et  $B = \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix}$  et normaux à  $n = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$  et  $n' = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix}$  non nuls et non colinéaires. On note  $d = \langle n, \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \rangle$  et  $d' = \langle n', \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix} \rangle$ . On a alors :

$$\mathcal{D} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases} = \left\{ \begin{cases} \langle n, \begin{pmatrix} x-\alpha \\ y-\beta \\ z-\gamma \end{pmatrix} \rangle = 0 \\ \langle n', \begin{pmatrix} x-\alpha' \\ y-\beta' \\ z-\gamma' \end{pmatrix} \rangle = 0 \end{cases} \right\}.$$

**Proposition 1.4.3.** La distance entre un point  $M$  et une droite  $\mathcal{D}$  passant par les points  $A$  et  $B$

$$\text{dist}(M, \mathcal{D}) = \frac{\|\overrightarrow{AB} \wedge \overrightarrow{AM}\|}{\|\overrightarrow{AB}\|}.$$

### 1.4.5 Sphères dans l'espace

**Définition 1.4.2.** La sphère  $\mathcal{S}$  de centre  $\mathcal{O} = (x_0, y_0, z_0) \in \mathbb{R}^3$  et de rayon  $r > 0$  est

$$\begin{aligned} \mathcal{S} &= \{M \in \mathbb{R}^3 \mid d(M, \mathcal{O}) = r\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2\}. \end{aligned}$$



# Chapitre 2

## Transformation linéaires

### 2.1 Définitions et premières propriétés

#### 2.1.1 Définitions

Les applications linéaires sont précisément les fonctions qui préservent la structure d'espace vectoriel.

**Définition 2.1.1.** Soient  $E$  et  $F$  deux espaces vectoriels sur  $\mathbb{R}$ . On appelle **application linéaire** de  $E$  dans  $F$  la donnée d'une application  $f : E \rightarrow F$  telle que :

$$\begin{aligned}\forall (u, v) \in E^2 : f(u + v) &= f(u) + f(v), \\ \forall u \in E, \forall \lambda \in \mathbb{R} : f(\lambda u) &= \lambda f(u).\end{aligned}$$

**Remarque 11.** L'image directe et l'image réciproque d'un espace vectoriel est un espace vectoriel. De plus,

- Le zéro (l'origine) de  $E$  est envoyé sur le zéro de  $F$  :  $f(0_E) = 0_F$ . En effet,  $f(u) = f(u + 0_E) = f(u) + f(0_E)$ .
- L'image d'une combinaison linéaire est la combinaison linéaire des images :  $f(\lambda_1 u_1 + \dots + \lambda_p u_p) = \lambda_1 f(u_1) + \dots + \lambda_p f(u_p)$ .

**Proposition 2.1.1.** La composée de deux applications linéaires est une application linéaire.

*Démonstration.* Soit  $f : E \rightarrow F$  et  $g : F \rightarrow G$  deux applications linéaires,  $u$  et  $v$  deux vecteurs de  $E$  et  $\lambda$  un réel. On a bien

$$\begin{aligned}(g \circ f)(u + v) &= g(f(u + v)) = g(f(u) + f(v)) = g(f(u)) + g(f(v)) = (g \circ f)(u) + (g \circ f)(v), \\ (g \circ f)(\lambda u) &= g(f(\lambda u)) = g(\lambda f(u)) = \lambda g(f(u)) = \lambda(g \circ f)(u).\end{aligned}\quad \square$$

**Définition 2.1.2.** Soit  $f : E \rightarrow F$  et  $g : E \rightarrow F$  deux applications linéaires. On définit  $f + g : E \rightarrow F$  la **somme de deux applications linéaires** en posant :

$$(f + g)(u) = f(u) + g(u).$$

On définit aussi  $\lambda f$  le **produit d'une application linéaire par un scalaire** en posant :

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

- Proposition 2.1.2.** 1. La somme de deux applications linéaires est linéaire.  
2. Le produit d'une application linéaire par un scalaire est une application linéaire.

*Démonstration.*

$$\begin{aligned}(f + g)(u + v) &= f(u + v) + g(u + v) = f(u) + f(v) + g(u) + g(v) = (f + g)(u) + (f + g)(v), \\ (f + g)(\lambda u) &= f(\lambda u) + g(\lambda u) = \lambda f(u) + \lambda g(u) = \lambda(f + g)(u).\end{aligned}\quad \square$$

**Corollaire 2.1.1.** Soient  $E$  et  $F$  deux espaces vectoriels. L'ensemble  $\mathcal{L}(E, F)$  des applications linéaires de  $E$  dans  $F$  a une structure d'espace vectoriel pour les opérations de somme et produit par un scalaire définies ci-dessus.

**Remarque 12.** On verra dans la suite qu'une application linéaire de  $E = \mathbb{R}^m$  dans  $F = \mathbb{R}^n$  admet une représentation matricielle (en fait, plusieurs... car cela dépend d'un choix de bases). L'ensemble des matrices comportant  $n$  lignes et  $m$  colonnes est noté  $\mathbb{R}^{n \times m}$ . Remarquons alors que l'ensemble des matrices de la forme

$$\delta_{ij} = \begin{pmatrix} 0 & \cdots & \overset{j}{\downarrow} 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & 0 & 0 & 0 & \vdots \\ \vdots & 0 & 1 & 0 & \vdots \\ \vdots & 0 & 0 & 0 & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \leftarrow i, \quad i = 1, \dots, n \text{ et } j = 1, \dots, m$$

forment alors une base de  $\mathbb{R}^{n \times m}$  dont la dimension est  $nm$ .

## 2.1.2 Noyau et image

**Définition 2.1.3.** Soient  $E$  et  $F$  deux espaces vectoriels et  $f : E \rightarrow F$  une application linéaire. On appelle **noyau de l'application linéaire**  $f$  et on note  $\text{Ker}(f)$  l'ensemble des vecteurs de  $E$  dont l'image est le vecteur nul de  $F$ , autrement dit  $\text{Ker}(f)$  est l'image réciproque du zéro de  $F$  :

$$\text{Ker}(f) = \{u \in E : f(u) = 0\}.$$

**Proposition 2.1.3.** Le noyau d'une application linéaire  $f : E \rightarrow F$  est un sous-espace vectoriel de  $E$ .

*Démonstration.* Soit  $u$  et  $v$  dans  $\text{Ker}(f)$  et  $(\lambda, \mu) \in \mathbb{R}^2$ . On a

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v) = 0 + 0 = 0.$$

□

**Remarque 13.** Dans le cas où  $E$  est de dimension finie et  $f$  est une forme linéaire (i.e.  $f$  est à valeur réelle,  $F = \mathbb{R}$ ), le noyau est un hyperplan de dimension  $\dim(E) - 1$ . On dit aussi que  $\text{Ker}(f)$  est de codimension 1.

On rappelle qu'une application  $f$  est injective si

$$x \neq x' \implies f(x) \neq f(x').$$

Dans les démonstrations, il est souvent plus simple de montrer la contraposée :

$$f(x) = f(x') \implies x = x'.$$

Grâce au noyau, on dispose maintenant d'un critère beaucoup plus simple pour montrer qu'une application est injective.

**Proposition 2.1.4 Critère d'injectivité.** Soit  $f : E \rightarrow F$  une application linéaire. On a :

$$f \text{ injective} \iff \text{Ker}(f) = \{0_E\}.$$

*Démonstration.* Supposons  $f$  injective. On sait que  $f(0) = 0$  (car  $f$  est une application linéaire) donc  $0 \in \text{Ker}(f)$ . Soit  $u \neq 0$ ,  $f$  étant injective  $f(u) \neq f(0) = 0$ , aucun autre élément de  $E$  ne peut avoir pour image 0, donc  $\text{Ker}(f) = \{0\}$ .

Réciproquement, soit  $u$  et  $v$  deux vecteurs de  $E$  ayant la même image ( $f(u) = f(v)$ ), on a alors :

$$f(u - v) = 0 \iff (u - v) \in \text{Ker}(f) \iff u = v. \quad \square$$

**Définition 2.1.4.** Soit  $E$  et  $F$  deux espaces vectoriels et  $f$  une application linéaire. On appelle image de  $f$  et on note  $\text{Im}(f)$  l'ensemble des vecteurs de  $F$  image d'au moins un vecteur de  $E$ , autrement dit :

$$\text{Im}(f) = \{v \in F : \exists u \in E \text{ tel que } f(u) = v\}$$

On définit alors le **rang de l'application linéaire**  $f$  comme étant égal à la dimension de  $\text{Im}(f)$ .

**Théorème 2.1.1 Théorème du rang.** Soit  $E$  un espace vectoriel de dimension finie,  $F$  un espace vectoriel et  $f : E \rightarrow F$  une application linéaire. Alors  $\text{Ker}(f)$  et  $\text{Im}(f)$  sont des espaces vectoriels de dimension finie et :

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

## 2.2 Matrices et applications linéaires

**Matrices d'applications linéaires.** Nous avons introduit les bases afin de pouvoir faire des calculs dans les espaces vectoriels de dimension finie. Maintenant que nous avons défini les applications linéaires et présenté quelques unes de leurs propriétés, nous allons introduire les matrices pour pouvoir faire des calculs sur les applications linéaires.

Soit  $E$  un espace vectoriel de dimension finie  $p \geq 1$ ,  $F$  un espace vectoriel de dimension finie  $n \geq 1$  et  $f : E \rightarrow F$  une application linéaire. On considère  $\mathcal{B}_E$  une base de  $E$  et  $\mathcal{B}_F$  une base de  $F$  :

$$\begin{cases} \mathcal{B}_E = (e_1, \dots, e_p), \\ \mathcal{B}_F = (\varepsilon_1, \dots, \varepsilon_n). \end{cases}$$

Connaître  $f$  revient à connaître l'image de tout vecteur  $w$  de  $E$  par  $f$ . Soit  $w$  quelconque fixé dans  $E$ ,  $w$  s'écrit  $w = a_1e_1 + \dots + a_p e_p$  dans  $\mathcal{B}_E$ . Comme  $f(w)$  est un vecteur de  $F$ , cherchons ses coordonnées dans  $\mathcal{B}_F$ . On rappelle que  $f$  est une application linéaire, et on a

$$f(w) = a_1f(e_1) + \dots + a_pf(e_p)$$

Les vecteurs  $f(e_i)$  ( $i = 1, \dots, p$ ) sont dans  $F$  et peuvent donc se décomposer dans la base  $\mathcal{B}_F$  :

$$\begin{aligned} f(e_1) &= b_{11}\varepsilon_1 + \dots + b_{n1}\varepsilon_n \\ f(e_2) &= b_{12}\varepsilon_1 + \dots + b_{n2}\varepsilon_n \\ &\vdots \quad \quad \quad \vdots \\ f(e_p) &= b_{1p}\varepsilon_1 + \dots + b_{np}\varepsilon_n \end{aligned}$$

Il vient donc

$$\begin{aligned} f(w) &= a_1f(e_1) + \dots + a_pf(e_p) \\ &= (a_1b_{11} + a_2b_{12} + \dots + a_pb_{1p})\varepsilon_1 + \dots + (a_1b_{n1} + a_2b_{n2} + \dots + a_pb_{np})\varepsilon_n \end{aligned}$$

Les coordonnées de  $f(w)$  dans  $\mathcal{B}_F$  peuvent se réécrire comme un produit matriciel

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$

Pour identifier une application linéaire  $f$  en connaissant la base de départ et la base d'arrivée, il suffit de connaître

$$\mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f) = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

Cette matrice est appelée **matrice de l'application linéaire  $f$  par rapport aux bases  $\mathcal{B}_E$  et  $\mathcal{B}_F$** .

**Remarque 14.** — Pour connaître une application linéaire  $f : E \rightarrow F$ , il faut se donner une base  $(e_1, \dots, e_p)$  de l'espace de départ, une base  $(\varepsilon_1, \dots, \varepsilon_n)$  de l'espace d'arrivée,  $f$  est alors entièrement déterminée par la matrice  $\mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f)$

- Réciproquement, toute matrice de dimension  $\dim(F) \times \dim(E)$  définit une application linéaire de  $E$  dans  $F$ .
- Suivant les bases que l'on considère, une application linéaire peut avoir plusieurs représentations matricielles. C'est pourquoi il faut toujours préciser les bases dans lesquelles les calculs sont menés.

**Exemple 2.2.1.** Si  $\text{can} = (e_1, e_2, e_3)$  est la base canonique de  $\mathbb{R}^3$  et  $\mathcal{B}_F = (\varepsilon_1, \varepsilon_2)$  celle de  $\mathbb{R}^2$

1. La matrice de l'application  $\text{Id} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  par rapport à la base canonique est :

$$\mathcal{M}_{\text{can}, \text{can}}(\text{Id}) = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. L'application linéaire  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  définie par :

$$f(e_1) = 3\varepsilon_1 - 4\varepsilon_2, \quad f(e_2) = 2\varepsilon_1 + 3\varepsilon_2, \quad f(e_3) = \varepsilon_1 + \varepsilon_2.$$

a pour matrice

$$\mathcal{M}_{\text{can}, \mathcal{B}_F}(f) = \begin{pmatrix} 3 & 2 & 1 \\ -4 & 3 & 1 \end{pmatrix}.$$

Si maintenant  $\mathcal{B}_E = (e_2, e_1, e_3)$  et  $\mathcal{B}'_F = (\varepsilon_2, \varepsilon_1)$ , on a

$$\mathcal{M}_{\mathcal{B}_E, \mathcal{B}'_F}(f) = \begin{pmatrix} 3 & -4 & 1 \\ 2 & 3 & 1 \end{pmatrix}.$$

**Matrice de la somme d'applications.** La matrice de la somme est la somme des matrices.

**Proposition 2.2.1 Somme de matrices.** Soit  $E$  et  $F$  deux espaces vectoriels. On considère  $\mathcal{B}_E$  et  $\mathcal{B}_F$  deux bases de  $E$  et  $F$ . Soit  $f_1$  et  $f_2$  deux applications linéaires de  $E$  dans  $F$ . Alors

1. La matrice de  $f_1 + f_2$  par rapport aux bases  $\mathcal{B}_E$  et  $\mathcal{B}_F$  est :

$$\mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f_1 + f_2) = \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f_1) + \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f_2)$$

2. La matrice de  $\lambda f$  par rapport aux bases  $\mathcal{B}_E$  et  $\mathcal{B}_F$  est :

$$\mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(\lambda f) = \lambda \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f)$$

**Matrice de la composée d'applications linéaires.** La matrice de la composée de deux applications linéaires est le produit des matrices.

On considère trois espaces vectoriels de dimensions finies :

1.  $E$  de dimension  $m$  avec une base  $\mathcal{B}_E = (e_1, \dots, e_m)$ ,
2.  $F$  de dimension  $n$  avec une base  $\mathcal{B}_F = (\varepsilon_1, \dots, \varepsilon_n)$ ,
3.  $G$  de dimension  $p$  avec une base  $\mathcal{B}_G = (\eta_1, \dots, \eta_p)$ .

**Proposition 2.2.2 Matrice de la composée.** Soit  $f : E \rightarrow F$  et  $g : F \rightarrow G$  deux applications linéaires. La matrice de l'application linéaire  $h = g \circ f$  par rapport aux bases  $\mathcal{B}_E$  et  $\mathcal{B}_G$  est donnée par :

$$\mathcal{M}_{\mathcal{B}_E, \mathcal{B}_G}(h) = \mathcal{M}_{\mathcal{B}_F, \mathcal{B}_G}(g) \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f).$$

On a le diagramme commutatif suivant :

$$\begin{array}{ccc} (E, \mathcal{B}_E) & \xrightarrow{f, \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f)} & (F, \mathcal{B}_F) \\ & \searrow^{h=g \circ f, \mathcal{M}_h = \mathcal{M}_{\mathcal{B}_F, \mathcal{B}_G}(g) \mathcal{M}_{\mathcal{B}_E, \mathcal{B}_F}(f)} & \downarrow^{g, \mathcal{M}_{\mathcal{B}_F, \mathcal{B}_G}(g)} \\ & & (G, \mathcal{B}_G) \end{array}$$

**Exemple 2.2.2.** Soit  $(e_1, e_2, e_3)$  et  $(\varepsilon_1, \varepsilon_2)$  les bases canoniques de  $\mathbb{R}^3$  et  $\mathbb{R}^2$ . On considère deux applications linéaires  $f$  et  $g$  dont les matrices par rapport aux bases canoniques sont

$$\mathcal{M}_g = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 1 & 3 \end{pmatrix}, \quad \mathcal{M}_f = \begin{pmatrix} 1 & 2 & 5 \\ 4 & 3 & 1 \end{pmatrix}$$

1. Les applications linéaires en jeu sont donc  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  et  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ .
2. L'application composée est  $h = g \circ f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ .
3. Pour trouver la représentation matricielle de  $h$  dans les bases canoniques, il suffit de calculer  $h(e_i)$ ,  $i = 1, 2, 3$ . On a

$$\begin{aligned} h(e_1) &= g(f(e_1)) = g(\varepsilon_1 + 4\varepsilon_2) = g(\varepsilon_1) + 4g(\varepsilon_2) \\ &= 2e_1 + 3e_2 + e_3 + 4(e_1 + 4e_2 + 3e_3) = 6e_1 + 19e_2 + 13e_3 \end{aligned}$$

De même, il vient  $h(e_2) = 7e_1 + 18e_2 + 11e_3$  et  $h(e_3) = 11e_1 + 19e_2 + 8e_3$ . Finalement,

$$M_h = \begin{pmatrix} 6 & 7 & 11 \\ 19 & 18 & 19 \\ 13 & 11 & 8 \end{pmatrix}$$

Il suffit alors de vérifier que  $\mathcal{M}_h = \mathcal{M}_g \mathcal{M}_f$ .

## 2.3 Inversion de matrices

**Proposition 2.3.1.** Soit  $f$  une application linéaire inversible de  $E$  dans  $E$ . On note  $f^{-1}$  son application inverse, i.e.,  $f^{-1} : E \rightarrow E$  et  $f \circ f^{-1} = f^{-1} \circ f = \text{Id}_E$ , alors

$$M_{\mathcal{B}_E, \mathcal{B}_E}(f^{-1}) = M_{\mathcal{B}_E, \mathcal{B}_E}^{-1}(f).$$

**Remarque 15.** Si on a deux matrices carrées  $A, B$  de même taille et inversibles on a

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

### 2.3.1 Déterminant

On a déjà vu comment calculer le déterminant d'une matrice  $2 \times 2$  et  $3 \times 3$ . Dans le cas général, le calcul met en jeu un polynôme homogène de degré  $n$  en les coordonnées de la matrice. Il existe des formules générales récursives mais de complexités importantes, typiquement  $O(n^3)$  voir plus si on utilise une implémentation naïve. Pour les détails, on renvoie le lecteur à [https://fr.wikipedia.org/wiki/D%C3%A9terminant\\_\(math%C3%A9matiques\)](https://fr.wikipedia.org/wiki/D%C3%A9terminant_(math%C3%A9matiques)) ou Chapitre 1 de [1].

Étant donnée une matrice carrée  $A$  de taille  $n$ , on note :

- $a_{ij}$  est l'entrée  $(i, j)$  de la matrice  $A$
- $A_{\overline{ij}}$  est la matrice carrée de taille  $(n-1) \times (n-1)$  contenant les entrées de  $A$  auquel on a enlevé la  $i$ -ème ligne et la  $j$ -ème colonne. Cette sous matrice est appelée  $(i, j)$ -mineure de  $A$ .

**Proposition 2.3.2.** Soit donc  $A \in \mathbb{R}^{n \times n}$ , on a pour tout  $i_0 = 1, \dots, n$  fixé :

$$\det(A) = \sum_{j=1}^n (-1)^{i_0+j} a_{i_0j} \det(A_{\overline{i_0j}}).$$

De plus, on a pour tout  $j_0 = 1, \dots, n$  fixé :

$$\det(A) = \sum_{i=1}^n (-1)^{i+j_0} a_{ij_0} \det(A_{\overline{ij_0}}).$$

Les propriétés du déterminant sont (très) riches. En voici quelques unes parmi les plus utiles :

**Proposition 2.3.3.** Soit  $A, B$  deux matrices carrées de taille  $n \times n$  et  $\lambda \in \mathbb{R}$ . Le déterminant satisfait les propriétés suivantes :

- $\det(\text{Id}) = 1$  et  $\det(\lambda A) = \lambda^n \det(A)$ ,
- $\det(A^t) = \det(A)$ ,
- si  $A$  est inversible :  $\det(A^{-1}) = \det(A)^{-1}$ ,
- $\det(AB) = \det(A) \det(B)$
- si on permute deux colonnes (ou deux lignes) de  $A$  on change le signe du déterminant.

### 2.3.2 Calcul de l'inverse en pratique

Étant donnée une matrice carrée  $A \in \mathbb{R}^{n \times n}$  et un vecteur  $u \in \mathbb{R}^n$ , il existe un grand nombre de méthodes algorithmiques (explicites ou approchées)

- calculer l'inverse  $A^{-1}$  de  $A$
- résoudre le système linéaire  $Ax = u$  d'inconnue  $x \in \mathbb{R}^n$ .

Ces deux problèmes sont reliés mais loin d'être équivalents car dans le second cas on a pas besoin de connaître (ni de stocker)  $A^{-1}$ , on veut juste  $x$ ...

**Cas général** Pour calculer l'inverse d'une matrice, nous verrons en TD le **pivot de Gauss** qui est décrit par exemple ici : [https://en.wikipedia.org/wiki/Gaussian\\_elimination](https://en.wikipedia.org/wiki/Gaussian_elimination).

**Inverser des petites matrices** La formule de Laplace permet de donner une solution explicite pour inverser une matrice  $A$  :

$$A \text{com}(A)^t = \text{com}(A)^t A = \det(A) \text{Id}$$

où  $\text{com}(A)$  est la **comatrice** de  $A$  calculée à partir des entrées de  $A$  (relation polynomiale) et qui existe même si  $A$  n'est pas inversible. La transposée  $\text{com}(A)^t$  est appelé matrice complémentaire (**adjugate** en anglais). On voit que si  $\det(A) \neq 0$  on a

$$A^{-1} = \frac{1}{\det(A)} \text{com}(A)^t.$$

Nous nous contenterons de donner la formule de la comatrice pour les cas  $2 \times 2$  et  $3 \times 3$ .

*Dimension 2.* Soit  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ , la comatrice est  $\text{com}(A) = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$  et on a

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}.$$

*Dimension 3.* Pour le cas  $3 \times 3$ , on peut calculer la comatrice comme suit

$$\text{com} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & + \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ - \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & - \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & - \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}$$

où  $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ . On remarque alors que si l'on note  $x = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}$ ,  
 $y = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$ ,  $z = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$  les trois colonnes de  $A$ , on a

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} (y \wedge z)^t \\ (z \wedge x)^t \\ (x \wedge y)^t \end{pmatrix}$$

**Remarque 16.** Des formules similaires existent pour la dimension 4. Cela est bien utile pour les vecteurs en coordonnées homogènes. Voir Section 1.7 de [1].

### 2.3.3 Les matrices orthogonales

Nous nous intéressons dans cette partie aux applications qui conservent le produit scalaire. Pour simplifier, on ne considérera des espaces vectoriels munis du produit scalaire canonique (on parle d'**espace euclidiens**).

**Définition - Proposition 2.3.1.** Soient  $(E, \langle \cdot, \cdot \rangle)$  et  $(F, \langle \cdot, \cdot \rangle)$  deux espaces euclidiens. Soit  $f : E \rightarrow F$  une application linéaire. Les conditions suivantes sont équivalentes :

1.  $f$  conserve les produits scalaires, c'est à dire que pour tous  $u$  et  $v$  de  $E$ , on a  $\langle f(u), f(v) \rangle = \langle u, v \rangle$
2.  $f$  conserve les normes, c'est à dire que pour tout  $u$  de  $E$  on a  $\|f(u)\| = \|u\|$ .

Une application linéaire qui vérifie une de ces deux conditions est appelée **application orthogonale** ou **isométrie**.

*Démonstration.* Si la première condition est vérifiée, on a :

$$\|f(u)\|^2 = \langle f(u), f(u) \rangle = \langle u, u \rangle = \|u\|^2.$$

Si la seconde condition est vérifiée, alors

$$2 \langle f(u), f(v) \rangle = \|f(u) + f(v)\|^2 - \|f(u)\|^2 - \|f(v)\|^2 = \|u + v\|^2 - \|u\|^2 - \|v\|^2 = 2 \langle u, v \rangle.$$

□

**Proposition 2.3.4 Matrices des transformations orthogonales.** Soit  $E$  un espace vectoriel,  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthonormée de  $E$  et  $f : E \rightarrow E$  une application linéaire. On pose  $A = \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)$

1. Si  $f$  est une isométrie, la matrice  $A$  :
  - est inversible,
  - vérifie  $A^t A = A A^t = I$ , autrement dit  $A^{-1} = A^t$ ,
  - est telle que  $\det(A) = 1$  ou  $-1$ .
2. Si  $A$  vérifie  $A^t A = A A^t = I$ , autrement dit si  $A^{-1} = A^t$ , alors  $f$  est une transformation orthogonale.

Une matrice  $A$  qui vérifie une de ces deux conditions est appelée **matrice orthogonale**.

Inverser une telle matrice ne coûte donc rien en temps de calcul ! L'ensemble des matrices orthogonales forme un sous-groupe de matrices (la multiplication de deux matrices orthogonales reste une matrice orthogonale) appelé  $O(n)$ . Les matrices de  $O(n)$  qui ont un déterminant positif (égale à 1 donc...) est appelé le groupe spécial orthogonal et est noté  $SO(n)$ .



*Démonstration.* 1. Soit  $f$  une transformation orthogonale, alors

$$\text{Ker}(f) = \{u \in E : f(u) = 0\} = \{u \in E : \|f(u)\| = 0\} = \{u \in E : \|u\| = 0\} = \{0\},$$

donc  $f$  est injective et donc bijective.  $A$  est donc inversible. Soit  $u$  et  $v$  deux vecteurs de  $E$  de coordonnées  $U$  et  $V$  dans  $\mathcal{B}$ . Comme  $\langle f(u), f(v) \rangle = \langle u, v \rangle$  on a

$$(AU)^t AV = U^t A^t AV = U^t V$$

d'où  $A^t A = I$  et par conséquent  $(\det(A))^2 = 1$ .

2. De même, pour deux vecteurs  $u$  et  $v$  dans  $E$  on a :

$$\langle f(u), f(v) \rangle = (AU)^t AV = U^t A^t AV = U^t V = \langle u, v \rangle.$$

□

On peut reconnaître une matrice orthogonale grâce à la proposition suivante :

**Proposition 2.3.5.** Soit  $E$  un espace euclidien,  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthonormée de  $E$ ,  $f : E \rightarrow E$  une application linéaire et  $A = \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)$  la matrice de  $f$  par rapport à  $\mathcal{B}$ . Alors :

1.  $f$  est une transformation orthogonale si et seulement si  $(f(e_1), \dots, f(e_n))$  est une base orthonormée de  $E$ .
2.  $A$  est une matrice orthogonale si et seulement si :
  - le produit scalaire entre les vecteurs colonnes de  $A$  est nul,
  - les vecteurs colonnes de  $A$  sont de normes 1.

*Démonstration.* Il est évident que les points 1 et 2 sont équivalents. Montrons 1.

- Soit  $f$  une transformation orthogonale. Alors  $\forall i = 1, \dots, n, \|f(e_i)\| = \|e_i\| = 1$  et  $\forall i \neq j, \langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = 0$
- Soit  $u = \sum_{i=1}^n u_i e_i$  et  $v = \sum_{i=1}^n v_i e_i$  deux vecteurs de  $E$ . On a

$$\begin{aligned} \langle f(u), f(v) \rangle &= \left\langle f \left( \sum_{i=1}^n u_i e_i \right), f \left( \sum_{i=1}^n v_i e_i \right) \right\rangle = \sum_{1 \leq i, j \leq n} u_i v_j \langle f(e_i), f(e_j) \rangle \\ &= \sum_{1 \leq i, j \leq n} u_i v_j = \langle u, v \rangle. \end{aligned}$$

□

**Corollaire 2.3.1.** Soit  $E$  un espace euclidien,  $\mathcal{B} = (e_1, \dots, e_n)$  et  $\mathcal{B}' = (e'_1, \dots, e'_n)$  deux bases orthonormées de  $E$ . La matrice de passage  $P = \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id})$  est une matrice orthogonale.

**Exemple 2.3.1.** Les matrices de rotation de  $\mathbb{R}^3$ . Les matrices de rotation d'un angle  $\theta \in \mathbb{R}$  autour des axes sont :

$$R_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}, \quad R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

On parle de lacet (yaw), tangage (pitch) et roulis (roll). On peut les composer, étant donné  $\alpha, \beta, \gamma \in \mathbb{R}$  :

$$\begin{aligned} R &= R_z(\alpha) R_y(\beta) R_x(\gamma) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha \cos \beta & \cos \alpha \sin \beta \sin \gamma - \sin \alpha \cos \gamma & \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma \\ \sin \alpha \cos \beta & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma \\ -\sin \beta & \cos \beta \sin \gamma & \cos \beta \cos \gamma \end{pmatrix} \end{aligned}$$





on a

$$\mathcal{M}_{\text{can}, \mathcal{B}_2}(\text{Id}) = \mathcal{M}_{\mathcal{B}_1, \mathcal{B}_2}(\text{Id}) \mathcal{M}_{\text{can}, \mathcal{B}_1}(\text{Id}) \iff \mathcal{M}_{\mathcal{B}_1, \mathcal{B}_2}(\text{Id}) = \mathcal{M}_{\text{can}, \mathcal{B}_2}(\text{Id}) \mathcal{M}_{\mathcal{B}_1, \text{can}}(\text{Id}).$$

On a donc

$$\mathcal{M}_{\mathcal{B}_1, \mathcal{B}_2}(\text{Id}) = \begin{pmatrix} 3 & 0 & 2 \\ 2 & -3 & -3 \\ 2 & -1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -7 & -2 \\ 10 & -15 & -5 \\ -7 & 11 & 3 \end{pmatrix},$$

et

$$\mathcal{M}_{\mathcal{B}_2, \mathcal{B}_1}(\text{Id}) = \mathcal{M}_{\mathcal{B}_1, \mathcal{B}_2}(\text{Id})^{-1} = \begin{pmatrix} 2 & -0.2 & 1 \\ 1 & 0.2 & 1 \\ 1 & -1.2 & -1 \end{pmatrix}.$$

## 2.4.2 Pour les applications linéaires

**Exemple 2.4.2.** Dans  $\mathbb{R}^2$  muni de la base canonique  $(e_1, e_2)$ , posons  $u = (3, 1)$  et  $v = (5, 2)$ . Les vecteurs  $u$  et  $v$  forment une base  $\mathcal{B}$  de  $\mathbb{R}^2$ . Considérons alors l'application linéaire de  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  définie par :

$$f(u) = 2u, \quad f(v) = -v$$

Par rapport à la base  $\mathcal{B}$ ,  $f$  a pour matrice une matrice diagonale :

$$\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Déterminons maintenant la matrice de  $f$  par rapport à la base canonique. Il est facile de voir que  $e_1 = 2u - v$  et  $e_2 = 3v - 5u$ . On en déduit :

$$\begin{cases} f(e_1) = 2f(u) - f(v) = 4u + v = 17e_1 + 6e_2, \\ f(e_2) = 3f(v) - 5f(u) = -3v - 10u = -45e_1 - 16e_2. \end{cases}$$

La matrice de  $f$  par rapport à la base canonique est donc :

$$\mathcal{M}_{\text{can}, \text{can}}(f) = \begin{pmatrix} 17 & -45 \\ 6 & -16 \end{pmatrix}.$$

Il est évident que la base  $\mathcal{B}$  est mieux adaptée aux calculs sur l'application  $f$ . Montrons le sur deux exemples :

1. La matrice de  $f^5$  dans la base  $\mathcal{B}$  est facile à calculer :

$$\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f^5) = \begin{pmatrix} 2^5 & 0 \\ 0 & (-1)^5 \end{pmatrix}.$$

2. La matrice de  $f^{-1}$  dans la base  $\mathcal{B}$  est :

$$\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f^{-1}) = \begin{pmatrix} 0.5 & 0 \\ 0 & -1 \end{pmatrix}.$$

On considère deux espaces vectoriels de dimension finie  $E$  et  $F$ , deux bases  $\mathcal{B}$  et  $\mathcal{B}'$  de  $E$ , deux bases  $\mathcal{C}$  et  $\mathcal{C}'$  de  $F$  et une application linéaire  $f : E \rightarrow F$ .

$$\begin{array}{ccc} (E, \mathcal{B}) & \xrightarrow{f, \mathcal{M}_{\mathcal{B}, \mathcal{C}}(f)} & (F, \mathcal{C}) \\ \uparrow \text{Id}_E, \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E) & & \downarrow \text{Id}_F, \mathcal{M}_{\mathcal{C}', \mathcal{C}}(\text{Id}_F)^{-1} \\ (E, \mathcal{B}') & \xrightarrow{f, \mathcal{M}_{\mathcal{B}', \mathcal{C}'}(f)} & (F, \mathcal{C}') \end{array}$$

La formule de changement de base pour l'application  $f$  est

$$\mathcal{M}_{\mathcal{B}', \mathcal{C}'}(f) = \mathcal{M}_{\mathcal{C}', \mathcal{C}}(\text{Id}_F)^{-1} \mathcal{M}_{\mathcal{B}, \mathcal{C}}(f) \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E).$$

**Proposition 2.4.3 Cas de  $\mathbb{R}^n$ .** Soit  $f$  une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ . Soit  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de  $\mathbb{R}^n$ .

- $\det(\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)) = \det_{\mathcal{B}', \mathcal{B}'}(f) = \det(f)$  (le déterminant de la matrice associée à  $f$  ne dépend pas de la base choisie). Ce déterminant est appelé **déterminant de l'application linéaire  $f$** .
- $f$  est inversible si et seulement si  $\det(f) \neq 0$  (dans ce cas,  $f$  est bijective).

**Application :** Les changements de base pour les applications linéaires peuvent être utilisés pour calculer les puissances ou les inverses de matrice :

$$\begin{aligned} \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f) &= \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E) \mathcal{M}_{\mathcal{B}', \mathcal{B}'}(f) \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E)^{-1} \\ \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)^n &= \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E) \mathcal{M}_{\mathcal{B}', \mathcal{B}'}(f)^n \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E)^{-1} \end{aligned}$$

Cette formule reste encore valable pour tout  $n \in \mathbb{Z}$  quand  $f$  est inversible :

$$\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)^{-1} = \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E) \mathcal{M}_{\mathcal{B}', \mathcal{B}'}(f)^{-1} \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E)^{-1}$$

**Exemple 2.4.3.** Appliquons les résultats précédents pour retrouver le changement de base de l'exemple 2.4.2. On connaît  $\mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)$  et on cherche  $\mathcal{M}_{\text{can}, \text{can}}(f)$ . On déduit du diagramme

$$\begin{array}{ccc} (\mathbb{R}^3, \mathcal{B}) & \xrightarrow{f, \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)} & (\mathbb{R}^2, \mathcal{B}) \\ \uparrow \text{Id}, \mathcal{M}_{\text{can}, \mathcal{B}}(\text{Id}) & & \downarrow \text{Id}, \mathcal{M}_{\text{can}, \mathcal{B}}(\text{Id})^{-1} \\ (\mathbb{R}^2, \text{can}) & \xrightarrow{f, \mathcal{M}_{\text{can}, \text{can}}(f)} & (\mathbb{R}^2, \text{can}) \end{array}$$

$$\mathcal{M}_{\text{can}, \text{can}}(f) = \mathcal{M}_{\text{can}, \mathcal{B}}(\text{Id})^{-1} \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f) \mathcal{M}_{\text{can}, \mathcal{B}}(\text{Id}).$$

On a

$$\mathcal{M}_{\mathcal{B}, \text{can}}(\text{Id}) = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \implies \mathcal{M}_{\text{can}, \mathcal{B}}(\text{Id}) = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$$

D'où

$$\mathcal{M}_{\text{can}, \text{can}}(f) = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 17 & -45 \\ 6 & -16 \end{pmatrix}.$$

**Remarque 17.** Soit  $E$  un espace vectoriel muni de deux bases  $\mathcal{B}$  et  $\mathcal{B}'$ . Les matrices  $\mathcal{M}_{\mathcal{B},\mathcal{B}'}(\text{Id}_E)$  et  $\mathcal{M}_{\mathcal{B}',\mathcal{B}}(\text{Id}_E)$  sont appelées **matrices de passage**. Les lettres  $P$  et  $Q$  sont souvent utilisées pour noter ces matrices. Mais attention, étant donné un vecteur  $u$  dans  $E$ , la matrice  $\mathcal{M}_{\mathcal{B},\mathcal{B}'}(\text{Id})$  permet de calculer les coordonnées de  $u$  dans  $\mathcal{B}'$  sachant ses coordonnées dans  $\mathcal{B}$  :

$$u_{\mathcal{B}'} = \mathcal{M}_{\mathcal{B},\mathcal{B}'}(\text{Id})u_{\mathcal{B}}$$

L'usage est cependant d'appeler cette matrice  $\mathcal{M}_{\mathcal{B},\mathcal{B}'}(\text{Id})$  matrice de passage de la base  $\mathcal{B}'$  à  $\mathcal{B}$ ...

### 2.4.3 Matrices semblables

**Définition 2.4.1.** On dit que deux matrices carrées  $A_1$  et  $A_2$  sont **semblables** s'il existe une matrice carrée  $P$  inversible telle que

$$A_1 = P^{-1}A_2P.$$

De manière équivalente, s'il existe une application linéaire  $f : E \rightarrow E$  et des bases  $\mathcal{B}_1$  et  $\mathcal{B}_2$  de  $E$  telles que :

$$A_1 = \mathcal{M}_{\mathcal{B}_1,\mathcal{B}_1}(f), \quad A_2 = \mathcal{M}_{\mathcal{B}_2,\mathcal{B}_2}(f).$$

La recherche d'une base bien adaptée au problème qu'on étudie est un thème important de l'algèbre linéaire. Si les calculs sont effectués dans une base mal adaptée au problème, ils peuvent rapidement devenir complexes. Il est donc nécessaire de développer des algorithmes permettant de se placer dans des bases où les calculs sont "simples". Enfin, certaines quantités sont invariantes par changement de base.

**Définition 2.4.2.** La **trace** d'une matrice carrée de taille  $n \times n$  est la somme de ses entrées diagonales. Si  $A = [a_{ij}]_{1 \leq i,j \leq n}$  alors

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

**Proposition 2.4.4.** Soit  $A_1, A_2$  deux matrices carrées semblables. Alors

1.  $\det(A_1) = \det(A_2)$ ,
2.  $\text{tr}(A_1) = \text{tr}(A_2)$ .

On peut calculer la trace et le déterminant d'un endomorphisme à l'aide de n'importe quelle matrice de cet endomorphisme (*i.e.* ces quantités ne dépendent pas des bases choisies).

# Chapitre 3

## Réduction des endomorphismes

Nous avons vu dans le chapitre précédent que pour certaines applications linéaires, il est possible de trouver des bases dans lesquelles les représentations matricielles de ces applications soient simples. Par exemple, pour  $u = (3, 1)$  et  $v = (5, 2)$ , l'application linéaire de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$  définie par :

$$f(u) = 2u, \quad f(v) = -v,$$

a pour représentation matricielle dans la base canonique :

$$\mathcal{M}_{\text{can,can}}(f) = \begin{pmatrix} 17 & -45 \\ 6 & -16 \end{pmatrix}$$

alors que dans la base  $\mathcal{B} = (u, v)$ ,  $f$  est représentée par une matrice diagonale :

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

**Question :** Comment trouver des bases dans lesquelles les représentations matricielle des **endomorphismes** (applications linéaires de  $E$  dans  $E$ ) sont diagonales ?

### 3.1 Valeurs propres et espaces propres

#### 3.1.1 Valeurs propres et polynôme caractéristique

Soit  $E$  un espace vectoriel de dimension finie et  $f$  un endomorphisme de  $E$ . On cherche une base  $\mathcal{B}$  de  $E$  telle que la matrice  $\mathcal{M}_{\mathcal{B},\mathcal{B}}(f)$  soit diagonale. Si  $\mathcal{B} = (v_1, \dots, v_n)$ , cela signifie que  $f(v_i) = \lambda_i v_i$  ou encore  $(f - \lambda I)(v_i) = 0$ ,  $v_i$  est donc un vecteur non nul de  $\ker(f - \lambda I)$  et  $f - \lambda I$  n'est pas injective.

**Exemple 3.1.1.** Soit

$$A = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}, \quad u = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

On a  $Au = \begin{pmatrix} -5 \\ -1 \end{pmatrix} \neq \lambda u$  et  $Av = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2v$ . Existe-t-il d'autres vecteurs  $w$  tels que  $Aw$  soit colinéaire à  $w$  ?

**Définition 3.1.1.** 1. Un **vecteur propre** d'une matrice  $A$  est un vecteur non nul tel que :  
 $Av = \lambda v$  pour un scalaire  $\lambda$  ;  
 2. Un scalaire  $\lambda$  est appelé **valeur propre** d'une matrice  $A$  lorsqu'il existe un vecteur  $v$  non nul tel que  $Av = \lambda v$ .

**Proposition 3.1.1.** Soit  $f$  un endomorphisme de  $E$  et  $A$  (resp.  $A'$ ) la représentation matricielle de  $f$  dans une base  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ). Alors  $A$  et  $A'$  possèdent les mêmes valeurs propres et les mêmes vecteurs propres. Ces valeurs propres (resp. vecteurs propres) sont appelées valeurs propres (resp. vecteurs propres) de l'application linéaire  $f$ .

*Démonstration.* Soit  $v$  un vecteur propre de la matrice  $A$  associé à la valeur propre  $\lambda$ . On note  $v_{\mathcal{B}}$  (resp.  $v_{\mathcal{B}'}$ ) le vecteur des coordonnées de  $V$  dans  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ). Soit  $P$  la matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$  :  $P = \mathcal{M}_{\mathcal{B}', \mathcal{B}}(f)$ . Il suffit de montrer  $A'v_{\mathcal{B}'} = \lambda v_{\mathcal{B}'}$ .

$$\begin{array}{ccc} (E, \mathcal{B}) & \xrightarrow{f, A} & (E, \mathcal{B}) \\ \uparrow \text{Id}, P & & \downarrow \text{Id}, P^{-1} \\ (E, \mathcal{B}') & \xrightarrow{f, A'} & (E, \mathcal{B}') \end{array}$$

On a  $A' = P^{-1}AP$  et aussi  $v_{\mathcal{B}} = Pv_{\mathcal{B}'}$ . Donc

$$A'v_{\mathcal{B}'} = P^{-1}APP^{-1}v_{\mathcal{B}} = P^{-1}Av_{\mathcal{B}} = \lambda P^{-1}v_{\mathcal{B}} = \lambda v_{\mathcal{B}'}. \quad \square$$

Il est facile de savoir si un vecteur  $v$  est vecteur propre d'une matrice (il suffit de faire le calcul). Il est en revanche plus compliqué de savoir si un scalaire est une valeur propre. La proposition suivante nous donne un moyen de pallier à cette difficulté.

**Proposition 3.1.2.** Un scalaire  $\lambda$  est une valeur propre de  $A$  si et seulement si l'équation vectorielle  $(A - \lambda I)u = 0$  admet un vecteur non nul comme solution. Autrement dit, quand  $\text{Ker}(A - \lambda I)$  n'est pas réduit à  $\{0\}$ .

**Théorème 3.1.1.** Un scalaire  $\lambda$  est une valeur propre de  $A$  si et seulement si  $\det(A - \lambda I) = 0$ . Le polynôme en  $\lambda \mapsto P(\lambda) = \det(A - \lambda I)$  est appelé **polynôme caractéristique** de  $A$  (ou de l'application linéaire  $f$  associée à  $A$ ).

Ce théorème nous permet de déterminer les valeurs propres d'un endomorphisme  $f$ , il suffit pour cela de connaître sa représentation matricielle  $A$  par rapport à une base quelconque. Les valeurs propres de  $f$  sont alors les solutions en  $\lambda$  de l'équation  $\det(A - \lambda I) = 0$ .

**Exemple 3.1.2.** Soit  $f$  un endomorphisme de  $\mathbb{R}^2$  dont la représentation matricielle dans la base canonique est

$$A = \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix}.$$

Pour trouver les valeurs propres de  $A$  on résout :

$$\det \begin{pmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{pmatrix} = 0 \iff \lambda^2 + 4\lambda - 21 = 0 \iff \lambda = 3 \text{ ou } \lambda = -7.$$



**Définition 3.1.2.** Si un scalaire  $\lambda$  est une valeur propre de  $A$ , le noyau  $\text{Ker}(A - \lambda I)$  est appelé **sous-espace propre** de  $A$  associé à la valeur propre  $\lambda$ . Dit autrement, le sous-espace propre associé à la valeur propre  $\lambda$  est formé par les vecteurs  $v$  de  $E$  tels que  $f(v) = \lambda v$  ( $f$  étant l'application linéaire associée à la matrice  $A$ ).

**Exemple 3.1.3.** Déterminons les sous-espaces propres associés aux deux valeurs propres de l'exemple précédent. Le sous-espace propre  $E_{\lambda_1}$  associé à la valeur propre  $\lambda_1 = 3$  est par définition le sous espace  $\text{Ker}(f - 3\text{Id})$  :

$$E_{\lambda_1} = \text{Ker}(f - 3\text{Id}) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : A \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix} \right\}.$$

On résout le système suivant pour décrire déterminer  $E_{\lambda_1}$  :

$$\begin{cases} 2x + 3y = 3x \\ 3x - 6y = 3y \end{cases} \iff \begin{cases} -x + 3y = 0 \\ 3x - 9y = 0 \end{cases} \iff x = 3y.$$

Par suite

$$E_{\lambda_1} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x = 3y \right\} = \left\{ y \begin{pmatrix} 3 \\ 1 \end{pmatrix} : y \in \mathbb{R} \right\} = \text{Vect}(3e_1 + e_2).$$

Par un raisonnement analogue, on trouve :

$$E_{\lambda_2} = \text{Vect}(-3e_1 + e_2).$$

On peut vérifier que dans la base  $\mathcal{B} = (-3e_1 + e_2, 3e_1 + e_2)$ ,  $f$  a pour représentation matricielle

$$M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} 3 & 0 \\ 0 & -7 \end{pmatrix}.$$

### 3.1.2 Théorème de Cayley-Hamilton

Une matrice carrée  $A$  annule son polynôme caractéristique.

**Théorème 3.1.2 Cayley-Hamilton.** Pour le polynôme caractéristique d'une matrice  $A$ , si on substitue  $\lambda$  par la matrice  $A$ , on obtient une expression matricielle ("un polynôme en  $A$ ") qui est la matrice des zéros.

**Exemple 3.1.4.** Soit  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ . Alors,

$$P(\lambda) = \det(A - \lambda I) = \lambda^2 - \text{tr}(A)\lambda + \det(A) = \lambda^2 - 5\lambda - 2.$$

Le Théorème de Cayley-Hamilton affirme que  $P(A) = A^2 - 5A - 2I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ . Applications

1. Si on dispose d'un polynôme annulateur de  $A$ , on sait que les valeurs propres sont racines de ce polynôme.
2. Calcul de l'inverse d'une matrice. Dans notre cas :

$$A^2 - 5A = 2I \iff A \frac{A - 5I}{2} = \frac{A - 5I}{2} A = I \iff A^{-1} = \frac{A - 5I}{2}$$

## 3.2 Diagonalisation

On dit qu'un polynôme  $P$  est **scindé** s'il se factorise sous la forme

$$P(x) = C(x - a_1)(x - a_2) \cdots (x - a_n)$$

où tous les  $a_i \in \mathbb{R}$  sont les racines de  $P$  et  $C \in \mathbb{R}$ . Par ailleurs, le polynôme  $P$  est dit **scindé à racines simples** si dans la factorisation précédente tous les  $a_i$  sont tous distincts (c'est le cas que l'on considère dans la Section 3.2.1).

### 3.2.1 Condition suffisante

**Proposition 3.2.1.** Soit  $A$  une matrice carrée de taille  $n \times n$ . Si le polynôme caractéristique de  $A$  admet  $n$  racines distinctes 2 à 2 alors  $A$  est diagonalisable.

*Démonstration.* Soit  $\lambda_1, \dots, \lambda_n$  les  $n$  racines distinctes du polynôme caractéristique et  $v_1, \dots, v_n$  les  $n$  vecteurs propres associés à ces valeurs propres. Il suffit de montrer que  $\mathcal{B} = (v_1, \dots, v_n)$  est une base de  $E$ .

Montrons par récurrence sur  $k$  que  $(v_1, \dots, v_k)$  est une famille libre de  $E$ . C'est clairement vrai pour  $k = 1$ . Supposons que ce soit vrai pour un entier  $k$  avec  $1 \leq k < n$  et montrons que  $(v_1, \dots, v_{k+1})$  est une famille libre de  $E$ . Soit  $\alpha_1, \dots, \alpha_{k+1}$  tels que :

$$\alpha_1 v_1 + \dots + \alpha_{k+1} v_{k+1} = 0, \quad (3.1)$$

on a, en prenant l'image par  $f$  :

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_{k+1} \lambda_{k+1} v_{k+1} = 0. \quad (3.2)$$

La combinaison (3.2)  $- \lambda_{k+1}$ (3.1) s'écrit :

$$\alpha_1 (\lambda_1 - \lambda_{k+1}) v_1 + \dots + \alpha_k (\lambda_k - \lambda_{k+1}) v_k = 0$$

Par hypothèse de récurrence, chacun des coefficients de cette combinaison linéaire est nul et comme les  $\lambda_k$  sont deux à deux distincts, cela implique  $\alpha_1 = \dots = \alpha_k = 0$ , ce qui reporté dans (3.1), donne  $\alpha_{k+1} = 0$ . Par conséquent  $\mathcal{B}$  est une famille libre de  $n$  éléments de  $E$ , c'est donc une base de  $E$ .  $\square$

**En pratique** Pour diagonaliser  $f : E \rightarrow E$  dont le polynôme caractéristique de la matrice carrée  $A = \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f)$  admet  $n$  racines distinctes, il faut :

1. Calculer le polynôme caractéristique  $\lambda \mapsto \det(A - \lambda I)$ .
2. Trouver les  $n$  racines de ce polynôme  $\lambda_1, \dots, \lambda_n$ .
3. Déterminer un vecteur  $v_i$  qui engendre l'espace propre  $E_{\lambda_i}$  pour tout  $i = 1, \dots, n$ .
4. L'ensemble des vecteurs  $v_i$  forment une base  $\mathcal{B}' = (v_1, \dots, v_n)$  de  $E$  et la matrice de  $f$  dans cette base est

$$D = \mathcal{M}_{\mathcal{B}', \mathcal{B}'}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

**Question :** si les  $n$  valeurs propres du polynôme caractéristique ne sont pas distinctes, peut-on diagonaliser  $A$ ?

### 3.2.2 Condition nécessaire et suffisante

**Définition 3.2.1.** Soit  $\lambda_1$  une racine d'un polynôme  $P(\lambda)$ . L'ordre de multiplicité de la racine  $\lambda_1$  est égal au plus grand entier  $r \in \mathbb{N}$  tel qu'il existe un polynôme  $Q(\lambda)$  avec  $P(\lambda) = (\lambda - \lambda_1)^r Q(\lambda)$ .

**Exemple 3.2.1.** 1.  $P(x) = x^2 + 2x + 1 = (x + 1)^2$ , alors  $-1$  est racine d'ordre 2.  
2.  $P(x) = x^3 + x^2 - 2x + 12 = (x - 2)^2(x + 3)$  alors 2 est racine d'ordre 2 et  $-3$  est racine d'ordre 1.

**Théorème 3.2.1.** Soit  $A$  une matrice  $n \times n$  admettant  $p$  valeurs propres distinctes.

1. Pour  $1 \leq k \leq p$ , la dimension de l'espace propre associé à  $\lambda_k$  est inférieure ou égale à l'ordre de multiplicité de la valeur propre  $\lambda_k$ .
2. La matrice  $A$  est diagonalisable si et seulement si la somme des dimensions des sous-espaces propres est égale à  $n$ .

**Corollaire 3.2.1.** Une matrice  $A$  est diagonalisable si et seulement si la dimension des espaces propres associés à chaque valeur propre  $\lambda_k$  est égal à l'ordre de multiplicité de  $\lambda_k$ .

**Exemple 3.2.2.** Considérons l'application linéaire  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  définie dans la base canonique de  $\mathbb{R}^3$  par la matrice :

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

Le polynôme caractéristique de  $A$  est :

$$\det \begin{pmatrix} -\lambda & 1 & -1 \\ -1 & 2 - \lambda & -1 \\ -1 & 1 & -\lambda \end{pmatrix} = -\lambda(\lambda - 1)^2$$

L'espace propre  $E_0$  associé à la valeur propre 0 est défini par  $\{v \in \mathbb{R}^3 : A(v) = 0\}$ , il faut donc résoudre le système :

$$\begin{cases} y - z = 0 \\ -x + 2y - z = 0 \\ -x + y = 0 \end{cases} \iff \begin{cases} x = y \\ y = y \\ z = y \end{cases} .$$

Donc

$$E_0 = \left\{ y \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} : y \in \mathbb{R} \right\} = \text{Vect}(e_1 + e_2 + e_3).$$

L'espace propre  $E_1$  associé à la valeur propre 1 est défini par  $\{v \in \mathbb{R}^3 : A(v) = v\}$ , il faut donc résoudre le système :

$$\begin{cases} y - z = x \\ -x + 2y - z = y \\ -x + y = z \end{cases} \iff -x + y - z = 0 .$$

L'espace propre  $E_1$  est de dimension 2, la matrice  $A$  est donc diagonalisable. Il reste à trouver une base de  $E_1$ , par exemple  $v_2 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$  et  $v_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$  (on peut en choisir d'autres). La matrice

de  $f$  par rapport à la base  $\mathcal{B} = (v_1, v_2, v_3)$  est donc :

$$D = \mathcal{M}_{\mathcal{B}, \mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On peut vérifier que  $A = PDP^{-1}$ ,  $D = P^{-1}AP$  où  $P$  est la matrice de passage de la base canonique à la base  $\mathcal{B}$  :

$$P = \mathcal{M}_{\mathcal{B}, \text{can}}(\text{Id}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

**Exemple 3.2.3.** Considérons maintenant l'application linéaire  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  définie dans la base canonique de  $\mathbb{R}^3$  par la matrice :

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{pmatrix}.$$

Le polynôme caractéristique de  $A$  est encore  $-\lambda(\lambda - 1)^2$ . L'espace propre associé à la valeur propre 1 est défini par les équations :

$$\begin{cases} x + 2y - 3z = x \\ 2x + 5y - 7z = y \\ x + 3y - 4z = z \end{cases} \iff \begin{cases} 2y - 3z = 0 \\ 2x + 4y - 7z = 0 \end{cases} \iff \begin{cases} x = \frac{1}{3}y \\ y = y \\ z = \frac{2}{3}y \end{cases}.$$

Par conséquent  $E_1 = \text{Vect}(e_1 + 3e_2 + 2e_3)$  est de dimension 1. La dimension de  $E_1$  n'est pas égale à l'ordre de multiplicité de la valeur propre 1 donc  $A$  n'est pas diagonalisable.

**Question :** Que faire quand on ne peut diagonaliser un endomorphisme  $f$  ?

### 3.3 Triangularisation

Une réduction qui paraît intéressante est la réduction d'une matrice à la forme triangulaire. On a le résultat suivant : les endomorphismes dont le polynôme caractéristique est scindé sont trigonalisables.

**Proposition 3.3.1.** Soit  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension finie et  $f : E \rightarrow E$  un endomorphisme de  $E$ . Si toutes les racines du polynôme caractéristique sont dans  $\mathbb{R}$ , alors  $f$  est triangularisable.

**Remarque 18.** Si  $K = \mathbb{C}$ , tout endomorphisme de  $E$  est triangularisable (car  $\mathbb{C}$  est algébriquement clos).

**En pratique** En pratique pour triangulariser un endomorphisme on procède de la manière suivante :

1. On calcule le polynôme caractéristique de  $f$ .
2. On cherche les racines de ce polynôme, *i.e.*, les valeurs propres de  $f$ . Si toutes les racines sont dans  $K$  on peut triangulariser.
3. On recherche les espaces propres associés à chaque valeur propre.
4. On complète les vecteurs propres en une base de  $E$ .

5. On détermine la matrice de  $f$  par rapport à cette base.
6. On recommence la méthode sur la matrice d'ordre inférieur.

**Exemple 3.3.1.** Triangulariser la matrice

$$A = \mathcal{M}_{\text{can,can}}(f) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

Les valeurs propres sont  $\lambda_1 = 3$  d'ordre 1 et  $\lambda_2 = 2$  d'ordre 2. Les espaces propres sont

$$E_{\lambda_1} = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad E_{\lambda_2} = \text{Vect} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Le sev  $E_{\lambda_2}$  est de dimension 1 donc  $A$  n'est pas diagonalisable. Néanmoins comme toutes les valeurs propres sont dans  $\mathbb{R}$ , elle est triangularisable dans  $\mathbb{R}$ . On pose  $u_1 = e_1 + e_2 - 2e_3$ ,  $u_2 = e_1$  et on complète la famille  $(u_1, u_2)$  en une base de  $\mathbb{R}^3$  en prenant  $u_3 = e_3$ . Soit  $P$  la matrice de passage de la base canonique à  $\mathcal{B} = (u_1, u_2, u_3)$  :

$$P = \mathcal{M}_{\mathcal{B},\text{can}}(\text{Id}) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

La matrice  $D = \mathcal{M}_{\mathcal{B},\mathcal{B}}(f)$  est triangulaire. Pour la calculer, il suffit de déterminer les coordonnées de  $f(u_3) = f(e_3) = -e_2 + 4e_3$  dans  $\mathcal{B}$ , c'est à dire

$$P^{-1} \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

On a donc trouvé une base  $\mathcal{B}$  par rapport à laquelle la matrice  $D$  associée à l'application linéaire  $f$  est triangulaire,

$$D = \mathcal{M}_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = P^{-1}AP$$

## 3.4 Projecteurs

**Définition 3.4.1.** Soient  $E_1$  et  $E_2$  deux sev supplémentaires d'un espace vectoriel  $E = E_1 \oplus E_2$ . Pour tout  $x \in E$ , il existe donc un unique couple  $(x_1, x_2) \in E_1 \times E_2$  tel que  $x = x_1 + x_2$ . Soit

$$\begin{aligned} E &\longrightarrow E \\ x = x_1 + x_2 &\longmapsto x_1 \end{aligned}$$

L'application  $p$  est linéaire et est appelée **projecteur de  $E$  sur  $E_1$  parallèlement à  $E_2$** .

**Proposition 3.4.1.** Soient  $E_1$  et  $E_2$  deux sev supplémentaires d'un espace vectoriel  $E = E_1 \oplus E_2$ . Soit  $p$  la projection sur  $E_1$  parallèlement à  $E_2$ , alors :

1.  $\text{Ker}(p) = E_2$ .
2.  $\text{Im}(p) = E_1$ .

3.  $p(x) = x$  si et seulement si  $x \in E_1$ .
4.  $p \circ p(x) = p(x)$  pour tout  $x \in E$ .

*Démonstration.* 1. Soient  $x = x_1 + x_2 \in E_1 \oplus E_2 = E$ . On a :

$$p(x) = 0 \iff x_1 = 0 \iff x = x_2 \iff x \in E_2.$$

Par conséquent  $\text{Ker}(p) = E_2$ .

2. Soient  $x = x_1 + x_2 \in E_1 \oplus E_2 = E$ . On a :

$$\begin{aligned} x \in \text{Im}(p) &\iff \exists x' = x'_1 + x'_2 \in E_1 \oplus E_2 : p(x') = x \\ &\iff x = x'_1 \in E_1 \\ &\iff x \in E_1. \end{aligned}$$

Donc  $\text{Im}(p) = E_1$ .

3. Soient  $x = x_1 + x_2 \in E_1 \oplus E_2 = E$ . On a :

$$p(x) = x \iff x = x_1 \iff x \in E_1.$$

4. Si  $p$  est un projecteur et si  $x \in E$ , alors  $p(x) \in E_1$  et comme  $E_1$  est stable par  $p$ ,  $p(p(x)) = p(x)$ . Donc  $p \circ p = p$ .

Réciproquement, si  $p$  est un endomorphisme de  $E$  vérifiant  $p \circ p = p$  : posons  $E_1 = \text{Im}(p)$  et  $E_2 = \text{Ker}(p)$  et montrons que ces deux sous-espaces vectoriels de  $E$  sont supplémentaires dans  $E$ .

- Soit  $x \in E_1 \cap E_2$ . Alors, on a à la fois  $p(x) = 0$  car  $x \in E_2 = \text{ker}(p)$  et  $x = p(x')$  où  $x' \in E$  car  $x \in E_1 = \text{Im}(p)$ . Par conséquent :  $0 = p(x) = p(p(x'))$ . Mais comme  $p \circ p = p$ , on a  $p(p(x')) = p(x') = x$  et donc  $x = 0$ , ce qui prouve que  $E_1$  et  $E_2$  sont en somme directe.
- Soit  $x \in E$ . On a :

$$x = \underbrace{p(x)}_{=x_1} + \underbrace{(x - p(x))}_{=x_2}$$

et  $x_1 \in E_1$  (car  $x_1 = p(x) \in \text{Im}(p)$ ) et  $x_2 \in E_2$  (car  $p(x_2) = p(x - p(x)) = p(x) - p^2(x) = p(x) - p(x) = 0$ ). Par conséquent :  $E = E_1 + E_2$ .

Donc  $E = E_1 \oplus E_2$ . Si  $x = x_1 + x_2 \in E_1 \oplus E_2 = E$ , comme  $x_1 \in E_1$ ,  $p(x_1) = x_1$  et comme  $x_2 \in E_2$ ,  $p(x_2) = 0$ . Par linéarité de  $p$ ,  $p(x) = p(x_1) + p(x_2) = x_1$ . Et  $p$  est donc bien le projecteur de  $E$  sur  $E_1$  parallèlement à  $E_2$ . □

**Proposition 3.4.2.** Soient  $E_1$  et  $E_2$  deux sev supplémentaires d'un espace vectoriel  $E = E_1 \oplus E_2$ . Soit  $p$  la projection sur  $E_1$  parallèlement à  $E_2$ , alors  $\text{Id} - p$  est le projecteur de  $E$  sur  $E_2$  parallèlement à  $E_1$ .

**Proposition 3.4.3.** Soient  $E_1$  et  $E_2$  deux sev supplémentaires d'un espace vectoriel  $E = E_1 \oplus E_2$ . Soit  $p$  la projection sur  $E_1$  parallèlement à  $E_2$ , alors les valeurs propres de  $p$  sont 0 ou 1. Autrement dit, il existe une base dans laquelle la matrice de  $p$  s'écrit :

$$\begin{pmatrix} 0_{d_2} & 0 \\ 0 & I_{d_1} \end{pmatrix}$$

où  $d_i = \dim(E_i)$  pour  $i = 1, 2$ .

-

Conséquence, si  $E$  est un espace euclidien, alors  $p$  est une contraction :  $\|p(x)\| \leq \|x\|$ .





# Chapitre 4

## ACP

### 4.1 Matrices symétriques

**Théorème 4.1.1 Théorème spectral.** Une matrice symétrique  $S \in \mathbb{R}^{n \times n}$  est diagonalisable en base orthonormée, *i.e.* il existe  $\lambda_1 \geq \dots \geq \lambda_n$  et une matrice orthogonale  $U \in \mathbb{R}^{n \times n}$  telle que :

$$S = U \text{Diag}(\lambda_1, \dots, \lambda_n) U^\top \text{ ou } SU = U \text{Diag}(\lambda_1, \dots, \lambda_n)$$

On peut donc écrire la matrice  $S$  comme la somme pondérée de matrice de projection. On a  $U = [\mathbf{u}_1, \dots, \mathbf{u}_n]$  qui donne :

$$S = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^\top, \quad \text{avec } \forall i \in \llbracket 1, n \rrbracket, \quad S \mathbf{u}_i = \lambda_i \mathbf{u}_i.$$

**Remarque 19.** 1. une matrice  $U \in \mathbb{R}^{n \times n}$  est dite orthogonale si elle vérifie  $U^\top U = U U^\top = I_n$  ou  $\forall (i, j) \in \llbracket 1, n \rrbracket^2 \quad \mathbf{u}_i^\top \mathbf{u}_j = \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j}$ .  
2. les  $\lambda_i$  sont les valeurs propres de  $S$  et les  $\mathbf{u}_i \in \mathbb{R}^n$  sont les vecteurs propres associés.  
3. les matrices  $\mathbf{u}_i \mathbf{u}_i^\top$  sont de rang 1 et correspondent aux projections sur  $\text{Vect}(\mathbf{u}_i)$ .

### 4.2 Décomposition en valeurs singulières (SVD)

#### 4.2.1 Valeurs singulières et vecteurs singuliers

**Théorème 4.2.1.** Pour toute matrice  $X \in \mathbb{R}^{n \times p}$ , il existe une matrice orthogonale  $U \in \mathbb{R}^{n \times n}$  et une matrice orthogonale  $V \in \mathbb{R}^{p \times p}$ , telles que

$$U^\top X V = \text{Diag}(s_1, \dots, s_{\min(n,p)}) = \Sigma \in \mathbb{R}^{n \times p}$$

avec  $s_1 \geq s_2 \geq \dots \geq s_{\min(n,p)} \geq 0$ , ou encore :

$$X = U \Sigma V^\top$$

avec  $U = [\mathbf{u}_1, \dots, \mathbf{u}_n]$  et  $V = [\mathbf{v}_1, \dots, \mathbf{v}_p]$ .

*Démonstration.* Diagonaliser  $X^\top X$  [?]

□

Les matrices  $U$  et  $V$  étant orthogonales, on a bien

$$\begin{cases} \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j}, & \forall (i, j) \in \llbracket 1, n \rrbracket^2 \\ \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j}, & \forall (i, j) \in \llbracket 1, p \rrbracket^2 \end{cases}$$

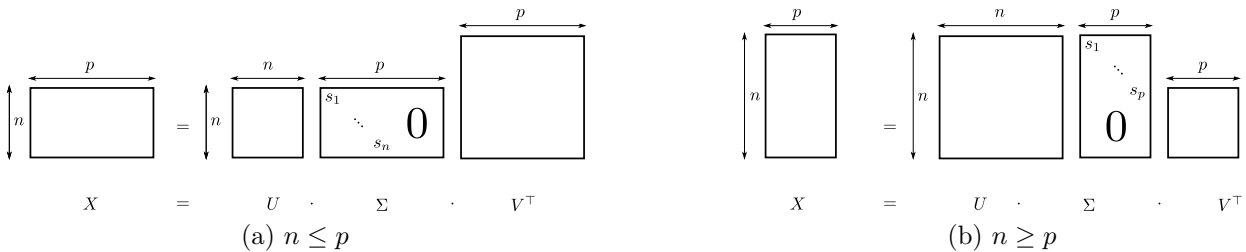


FIGURE 4.1 – Visualisation de la SVD.

Les nombres réels  $s_j$  sont les **valeurs singulières** de  $X$  et les  $\mathbf{u}_j$  (resp.  $\mathbf{v}_j$ ) sont les **vecteurs singuliers** à gauche (resp. droite).

**Remarque 20.** Attention, les matrices  $U$  et  $V$  ne sont pas unique. En particulier, les vecteurs  $\mathbf{u}_j$  (resp.  $\mathbf{v}_j$ ) sont définis à orientation près (multiplication par  $-1$ ).

**Propriété variationnelle de la plus grande valeur singulière** La plus grande valeur singulière satisfait :

$$s_1 = \begin{cases} \max_{\mathbf{u} \in \mathbb{R}^n, \mathbf{v} \in \mathbb{R}^p} \mathbf{u}^\top X \mathbf{v} \\ \text{s.c. } \|\mathbf{u}\|^2 = 1 \text{ et } \|\mathbf{v}\|^2 = 1 \end{cases}$$

Pour résoudre ce problème d’optimisation sous contrainte, on écrit le Lagrangien :

$$\mathcal{L}(\mathbf{u}, \mathbf{v}, \lambda_1, \lambda_2) = \mathbf{u}^\top X \mathbf{v} - \lambda_1(\|\mathbf{u}\|^2 - 1) - \lambda_2(\|\mathbf{v}\|^2 - 1)$$

qui donne les conditions nécessaire d’optimalité suivante :

$$\begin{cases} \nabla_{\mathbf{u}} \mathcal{L} = X \mathbf{v} - 2\lambda_1 \mathbf{u} = 0 \\ \nabla_{\mathbf{v}} \mathcal{L} = X^\top \mathbf{u} - 2\lambda_2 \mathbf{v} = 0 \end{cases} \iff \begin{cases} X \mathbf{v} = 2\lambda_1 \mathbf{u} \\ X^\top \mathbf{u} = 2\lambda_2 \mathbf{v} \end{cases} \Rightarrow \begin{cases} X^\top X \mathbf{v} = \alpha \mathbf{v} \\ X X^\top \mathbf{u} = \alpha \mathbf{u} \end{cases}$$

avec  $\alpha = 4\lambda_1\lambda_2$ , et donc  $\mathbf{v}$  est vecteur propre de  $X^\top X$  et  $\mathbf{u}$  est vecteur propre de  $X X^\top$ .

### 4.2.2 La SVD réduite

On part de la SVD  $X = U \Sigma V^\top$ . On ne garde que les éléments utiles avec  $r = \min(n, p)$  :

$$X = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^\top = U_r \text{Diag}(s_1, \dots, s_r) V_r^\top$$

avec  $s_i > 0, \forall i \in \llbracket 1, r \rrbracket$  et  $U_r = [\mathbf{u}_1, \dots, \mathbf{u}_r], V_r = [\mathbf{v}_1, \dots, \mathbf{v}_r]$ . Quand on en garde que les  $r = \text{rank}(X)$  valeurs singulières non-nulles, on parle alors de **SVD compacte**.

**Remarque 21.** 1. Les matrices  $\mathbf{u}_i \mathbf{v}_i^\top$  sont toutes de rang 1.  
2. Les  $\mathbf{u}_i$  (resp. les  $\mathbf{v}_i^\top$ ) sont orthonormés et engendrent le même espace que celui engendré par les colonnes (resp. les lignes) de  $X$  :

$$\text{Vect}(\mathbf{x}_1, \dots, \mathbf{x}_p) = \text{Vect}(\mathbf{u}_1, \dots, \mathbf{u}_r)$$

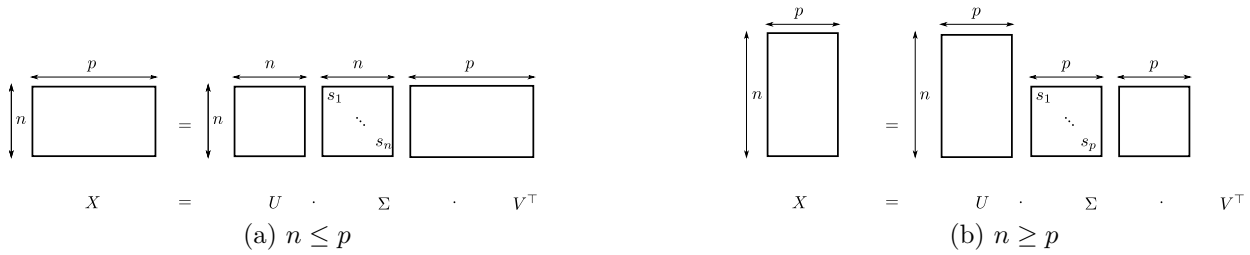


FIGURE 4.2 – Visualisation de la SVD réduite.

### 4.3 Meilleur approximation

Etant donnée une matrice  $X$  que l'on suppose de rang  $r \in \mathbb{N}^*$ , on se pose la question de trouver une matrice de rang  $1 \leq k \leq r$  qui approxime le mieux  $X$ . On a donc besoin d'un critère de proximité (en fait, une norme). On sait résoudre le problème pour la norme spectrale :

**Définition 4.3.1.** La **norme spectrale** de  $X \in \mathbb{R}^{n \times p}$  est définie par

$$\|X\|_2 = \sup_{u \in \mathbb{R}^p, \|u\|=1} \|Xu\| = |s_1(X)|$$

C'est donc le module de la plus grande valeur propre.

Dans ce cas, la réponse au problème de meilleur approximation est donnée par la SVD tronquée :

**Théorème 4.3.1 Meilleure approximation de rang  $k$ .** Soit  $X = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^\top$  la SVD compacte de  $X \in \mathbb{R}^{n \times p}$ . On note

$$X_k = \sum_{i=1}^k s_i \mathbf{u}_i \mathbf{v}_i^\top, \quad \text{pour tout } k \in \llbracket 1, r \rrbracket.$$

On a alors,

$$\min_{Z \in \mathbb{R}^{n \times p} : \text{rank}(Z)=k} \|X - Z\|_2 = \|X - X_k\|_2 = s_{k+1}$$

Cette propriété est cruciale pour l'analyse en composante principale (ACP).

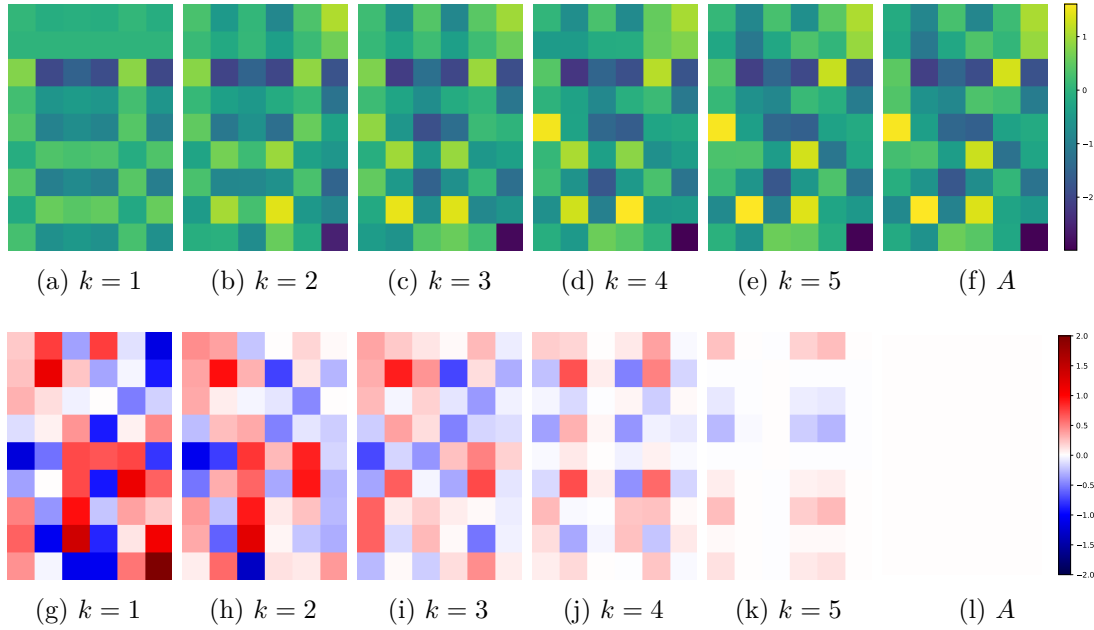


FIGURE 4.3 – Visualisation de l'approximation de rang  $k$ . Première ligne : les matrices  $A_k$ . Deuxième ligne : la différence  $A_k - A$ .

## 4.4 Pseudo-inverse

**Définition 4.4.1.** Si  $X \in \mathbb{R}^{n \times p}$  admet pour SVD  $X = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^\top$  avec  $r = \text{rank}(X)$ , alors sa **pseudo-inverse**  $X^+ \in \mathbb{R}^{p \times n}$  est définie par :

$$X^+ = \sum_{i=1}^r \frac{1}{s_i} \mathbf{v}_i \mathbf{u}_i^\top$$

Si  $X = \sum_{i=1}^n s_i \mathbf{u}_i \mathbf{v}_i^\top \in \mathbb{R}^{n \times n}$  est inversible alors  $X^+ = X^{-1}$

*Démonstration.*

$$\begin{aligned} XX^+ &= \sum_{j=1}^n s_j \mathbf{u}_j \mathbf{v}_j^\top \sum_{i=1}^n \frac{1}{s_i} \mathbf{v}_i \mathbf{u}_i^\top \\ &= \sum_{j=1}^n \sum_{i=1}^n s_j \frac{1}{s_i} \mathbf{u}_j \mathbf{v}_j^\top \mathbf{v}_i \mathbf{u}_i^\top \\ &= \sum_{j=1}^n \sum_{i=1}^n s_j \frac{1}{s_i} \delta_{i,j} \mathbf{u}_j \mathbf{u}_i^\top = \sum_{i=1}^n \mathbf{u}_i \mathbf{u}_i^\top = I_n. \end{aligned}$$

□

## 4.5 Analyse en composante principale (ACP ou PCA)

On observe  $n$  points  $x_1, \dots, x_n$  dans  $\mathbb{R}^p$ . On peut donc créer une matrice

$$X = \begin{pmatrix} x_1^\top \\ \vdots \\ x_n^\top \end{pmatrix} \in \mathbb{R}^{n \times p},$$

de  $n$  **observations** (lignes),  $p$  **features** (colonnes).

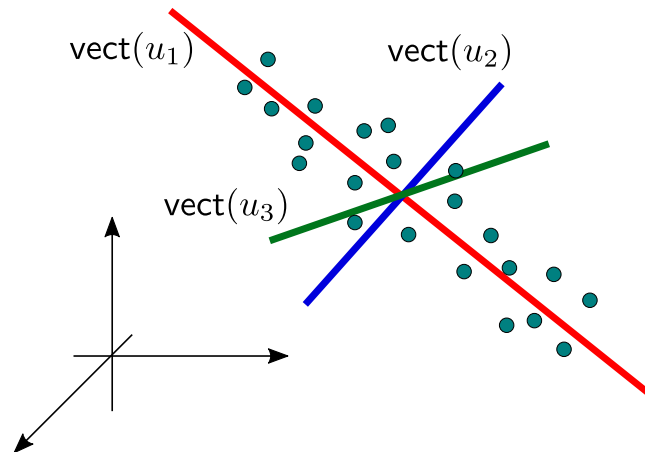


FIGURE 4.4 – Le but de l’ACP est de trouver les directions de plus grande dispersion d’un nuage de point (ici  $p = 3$ )

**Définition 4.5.1.** 1. La **moyenne** (ou le **barycentre**, ou encore le **centre de gravité**) d’un nuage de points  $X$  est

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \in \mathbb{R}^p.$$

2. La **variance** (ou l’**inertie**) d’un nuage de points est

$$\sigma_X^2 = \frac{1}{n} \sum_{i=1}^n \|x_i - \bar{x}\|^2 = \frac{1}{n} \sum_{i=1}^n \|x_i\|^2 - \left\| \frac{1}{n} \sum_{i=1}^n x_i \right\|^2.$$

3. L’**écart-type** d’un nuage de points est  $\sigma_X = \sqrt{\sigma_X^2}$ .

**Remarque 22.** La variance d’un nuage de point est nulle si et seulement si  $x_i = x_j$ , pour tout  $i, j = 1, \dots, n$ .

On considère ici des nuages de points centrés. À l’aide d’une translation, on peut toujours recentrer le nuage de points pour qu’il ait une moyenne nulle :

$$X \leftarrow \begin{pmatrix} x_1^\top - \bar{x}^\top \\ \vdots \\ x_n^\top - \bar{x}^\top \end{pmatrix} = X - \mathbb{1}_n \bar{x}^\top.$$

où l’on a noté  $\mathbb{1}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  et le vecteur colonne  $\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i \in \mathbb{R}^p$  contenant la moyenne de

chaque coordonnée des  $x_i, i = 1, \dots, n$ . On peut aussi mettre à l'échelle pour avoir un écart-type similaire par colonne (feature).

Soit le paramètre  $1 \leq k \leq p$ . On cherche à projeter le nuage de points  $X$  sur un sous espace vectoriel  $E_k \subset \mathbb{R}^p$  de dimension  $k$ . Cette méthode *compresse* donc le nuage de points de dimension  $p$  en un nuage de dimension  $k$ . L'idée est de choisir le sev  $E_k$  de manière à ce que le nuage projeté, noté  $X_k$  ressemble le plus à  $X$  : la compression doit faire perdre le moins possible d'information. On utilise un critère basée sur l'inertie.

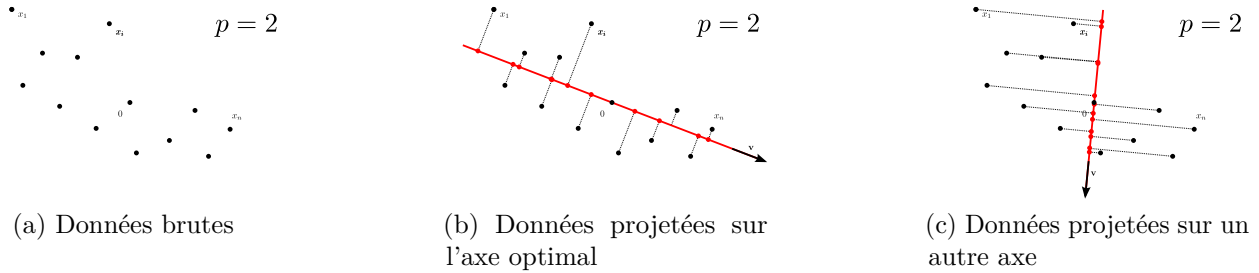


FIGURE 4.5 – Projection d'un nuage de points de dimension  $p = 2$  sur un sev de dimension  $k = 1$ .

**Définition 4.5.2.** L'Analyse en Composantes Principales (de niveau  $k$ ) consiste à effectuer la SVD de  $X$ , et à ne garder que les  $k$  axes principaux pour représenter le nuage.

$$X = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^\top \longrightarrow \sum_{i=1}^k s_i \mathbf{u}_i \mathbf{v}_i^\top.$$

Les sous espaces  $E_k = \text{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_k)$  sont appelés **sous espaces principaux**.

Cela nous donne une nouvelle manière de représenter les données. On a projeté le nuage de point original sur un sous-espaces vectoriel. Un peu de vocabulaire :

- Les axes (de direction)  $\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^p$  sont appelés **axes principaux** ou **axes factoriels**, les nouvelles variables

$$\mathbf{c}_j = X \mathbf{v}_j, j = 1, \dots, p$$

sont appelées **composantes principales** (coordonnées des  $x_i$  projetés sur l'axe  $\text{Vect}(\mathbf{v}_j)$ ).

- La matrice  $XV_k \in \mathbb{R}^{n \times k}$  (avec  $V_k = [\mathbf{v}_1, \dots, \mathbf{v}_k] \in \mathbb{R}^{p \times k}$ ) est la matrice représentant les coordonnées des données dans la base des  $k$  premiers axes principaux (nouvelle représentation d'ordre  $k$ ).

Souvent, on cherche à reconstruire les données dans l'espace original  $\mathbb{R}^p$  (e.g. pour débruiter) :

- Reconstruction "parfaite" pour  $x \in \mathbb{R}^p$  :

$$x = \sum_{j=1}^p \langle x, \mathbf{v}_j \rangle \mathbf{v}_j = \sum_{j=1}^p (x^\top \mathbf{v}_j) \mathbf{v}_j$$

qui donne  $X = XVV^\top$ .

- Reconstruction avec perte d'information pour  $x \in \mathbb{R}^p$  :

$$\hat{x} = \sum_{j=1}^k \langle x, \mathbf{v}_j \rangle \mathbf{v}_j = \sum_{j=1}^k (x^\top \mathbf{v}_j) \mathbf{v}_j \in \mathbb{R}^p$$

qui donne le nouveau nuage de points  $\hat{X}_k = XV_k V_k^\top$  projeté de  $X$  sur  $E_k$  dans  $\mathbb{R}^p$ .

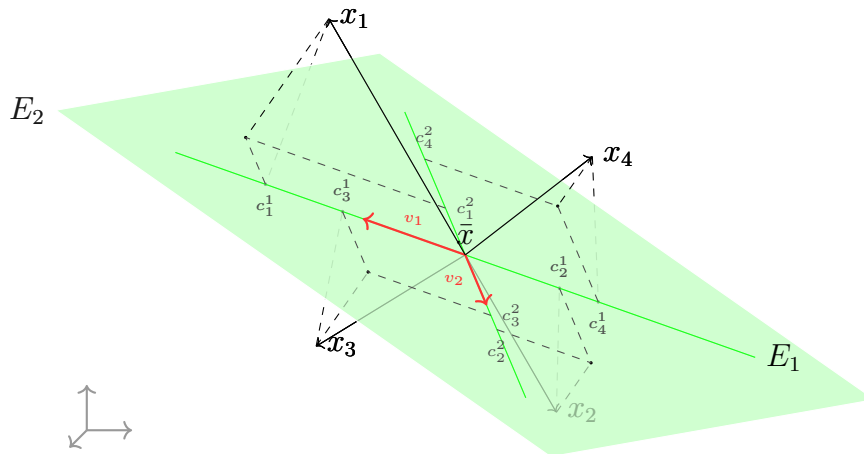


FIGURE 4.6 – Espaces factoriels et composantes principales ( $n = 4$  et  $p = 3$ )

### 4.5.1 Interprétation et récursion

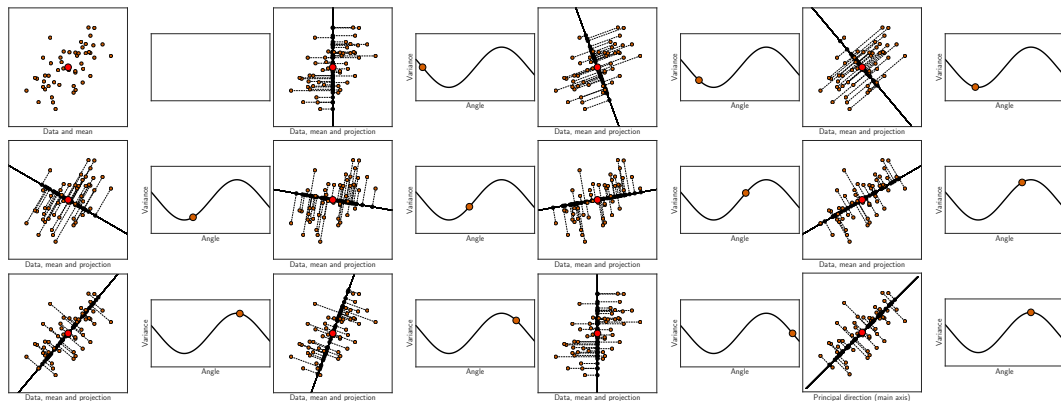


FIGURE 4.7 – Voir aussi la vidéo : <https://twitter.com/salmonjsph/status/1177564363261194240>

**Axe principal : maximisation de la variance projetée** L'axe principal (normalisé)  $\mathbf{v}_1$  est la solution du problème :

$$\mathbf{v}_1 \in \underset{\mathbf{v} \in \mathbb{R}^p, \|\mathbf{v}\|=1}{\text{Argmax}} \mathbf{v}^\top X^\top X \mathbf{v} = \underset{\mathbf{v} \in \mathbb{R}^p, \|\mathbf{v}\|=1}{\text{Argmax}} \|X\mathbf{v}\|^2 = \underset{\mathbf{v} \in \mathbb{R}^p, \|\mathbf{v}\|=1}{\text{Argmax}} \sum_{i=1}^n (x_i^\top \mathbf{v})^2$$

Après recentrage le dernier terme est la variance du nuage de points projeté sur l'axe  $\mathbf{v}$ . On résout donc un problème d'optimisation (une maximisation) sous contrainte convexe. On se ramène à un problème de maximisation global en considérant le Lagrangien. Cela revient à maximiser la fonction objectif suivante en  $\mathbf{v}$  :

$$\mathcal{L}(\mathbf{v}, \lambda) = (X\mathbf{v})^\top (X\mathbf{v}) - \lambda(\mathbf{v}^\top \mathbf{v} - 1) = \mathbf{v}^\top X^\top X \mathbf{v} - \lambda(\mathbf{v}^\top \mathbf{v} - 1)$$

où  $\lambda$  est le multiplicateur de Lagrange. Les conditions d'optimalité du premier ordre en un extremum sont

$$\frac{\partial \mathcal{L}(\mathbf{v}_1, \lambda)}{\partial \mathbf{v}} = 0 \Leftrightarrow X^\top X \mathbf{v}_1 = \lambda \mathbf{v}_1$$

La matrice de Gram  $X^\top X$  est diagonalisable (symétrique) donc si  $\mathbf{v}_1$  est un extremum alors c'est un vecteur propre.

On normalise  $\mathbf{v}_1$  pour que  $\|\mathbf{v}_1\| = 1$ , ainsi  $\lambda = \mathbf{v}_1^\top X^\top X \mathbf{v}_1$  et  $\mathbf{v}_1$  est un vecteur propre, de valeur propre  $\lambda$  maximale.

**Proposition 4.5.1.** Dans le cas où le nuage de points est centrée, on appelle **matrice de covariance** (ou **matrice de Gram**) de  $X \in \mathbb{R}^{n \times p}$  la matrice  $\frac{1}{n} X^\top X \in \mathbb{R}^{p \times p}$ . On a alors

$$\text{tr}\left(\frac{1}{n} X^\top X\right) = \sigma_X^2 = \underbrace{\sigma_{\pi_{\text{Vect}(u_1)}(X)}^2}_{=\frac{s_1}{n}} + \cdots + \underbrace{\sigma_{\pi_{\text{Vect}(u_p)}(X)}^2}_{=\frac{s_p}{n}}$$

où  $\pi_E$  est la projection orthogonale sur le sev  $E$ .

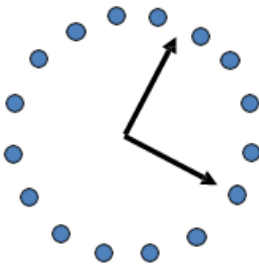
**Qualité de la représentation** La part d'inertie "expliquée" par le sous espace  $E_k$  est alors

$$\frac{s_1 + \cdots + s_k}{s_1 + \cdots + s_p}$$

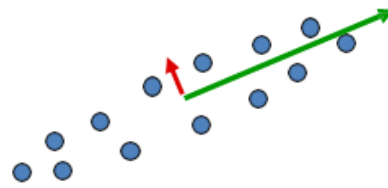
Si  $k = p$  la totalité de l'inertie est retrouvée. Le choix du nombre de dimension  $k$  se fait sur des critères empiriques. Cela dépend aussi de la finalité de l'ACP : pour de la description de données on a souvent  $q \leq 3$  (au delà difficultés d'interprétation), pour de la compression  $k$  peut être beaucoup plus grand.

Exemple de méthodes pour l'analyse descriptive :

- Méthode de Kaiser : Ne retenir que les directions principales qui expliquent une proportion de l'inertie supérieure à  $\frac{1}{p}$ .
- Méthode du coude : sur "l'éboulis" des valeurs propres, on observe un décrochement (coude) suivi d'une décroissance régulière : sélectionner les axes avant le décrochement.
- Méthode pratique : ne retenir que les dimensions que l'on peut interpréter...



(a) Les valeurs singulières sont égales. Pas de direction privilégiée. Tout vecteur est vecteur principal.



(b) Les valeurs singulières sont très déséquilibrées. Il y a une direction privilégiée.

**Exemple 4.5.1.** Soit  $X = \begin{pmatrix} -1 & -2 & 1 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \\ 2 & -1 & 2 \end{pmatrix}$ . On a donc  $n = 4$  et  $p = 3$ . Le barycentre des données

est  $\bar{X} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$  et la matrice de covariance est  $G = \frac{1}{4} \begin{pmatrix} 6 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ . Le polynôme caractéristique



de  $G$  est

$$p_G(\lambda) = -\lambda^3 + \frac{5}{2}\lambda^2 - \frac{7}{8}\lambda + \frac{1}{16}.$$

On trouve numériquement

$$s_1 \simeq 2.097$$

$$s_2 \simeq 0.3055$$

$$s_3 \simeq 0.09756$$

On trouve alors les vecteurs principaux  $v_1 = \begin{pmatrix} -0.833 \\ -0.330 \\ -0.443 \end{pmatrix}$ ,  $v_2 = \begin{pmatrix} -0.257 \\ 0.941 \\ -0.218 \end{pmatrix}$ ,  $v_3 = \begin{pmatrix} -0.489 \\ -0.0675 \\ -0.870 \end{pmatrix}$ . La part d'inertie expliquée par le premier axe principal est 0.839 et celle expliquée par le premier plan principal est 0.96.



# Bibliographie

- [1] E. Lengyel. *Foundations of Game Engine Development. Volume 1 : Mathematics*. Terathon Software LLC, 2016.