

Notes de cours d'algèbre linéaire
pour les groupes de maths renforcées
du S2 PCST

Michel Rumin

2025-2026

Table des matières

1	L'espace vectoriel \mathbb{R}^n	4
1.1	Qu'est-ce que c'est?	4
1.2	Les combinaisons linéaires	5
1.3	Droites et plans vectoriels	5
1.4	La base canonique de \mathbb{R}^n	5
2	Systèmes linéaires	6
2.1	Présentation du problème	6
2.2	Systèmes linéaires et combinaisons linéaires	7
2.3	Codage matriciel d'un système linéaire	7
2.4	Systèmes triangulaires	8
2.5	Systèmes échelonnés	9
2.6	Systèmes équivalents et transformations élémentaires	11
2.7	L'algorithme du pivot de Gauss–Jordan	12
2.8	Quelques exemples	13
2.9	La notion de rang	14
3	Matrices et applications linéaires	15
3.1	Applications linéaires de \mathbb{R}^p dans \mathbb{R}^n	15
3.2	Matrice d'une application linéaire	16
3.3	Calcul matriciel de l'image d'un vecteur	18
3.4	Exemples d'applications linéaires du plan	19
4	Opérations sur les matrices et les applications linéaires	22
4.1	Structure d'espace vectoriel	22
4.2	Produit de matrices et composition d'applications linéaires	23
5	Isomorphismes et matrices inversibles	26
5.1	Les isomorphismes	26
5.2	Matrices inversibles	27
5.3	Exemple et méthode de calcul	28
5.4	Le cas des matrices 2×2	29
6	Sous-espaces vectoriels, bases et dimension	31
6.1	Les sous-espaces vectoriels de \mathbb{R}^n	31

6.2	Bases et coordonnées	32
6.3	Notion d'indépendance linéaire	34
6.4	Dimension d'un sous-espace vectoriel	35
6.5	Calculs de dimensions et de bases	36
6.6	Conséquences utiles	40
7	Retour sur l'équation $AX=Y$	42
7.1	Généralités, noyau, image	42
7.2	Le théorème du rang	43
7.3	Exemple	43
7.4	Conséquences du théorème du rang	45
8	Espaces supplémentaires, somme directe, projections et symétries	45
8.1	Espaces supplémentaires et somme directe	45
8.2	Projections et symétries	46
8.3	Exemples	47
9	Calcul matriciel dans des bases quelconques	48
9.1	Matrices et applications linéaires dans des bases générales	49
9.2	Calcul matriciel général	51
9.3	Changements de bases et matrices de passage	52
9.4	Changement de base pour les endomorphismes	54
9.5	La trace d'une matrice	56
10	Évolution de systèmes et matrices de transition	58
10.1	Un exemple : les effectifs dans un cycle d'étude	58
10.2	Matrices de transition et évolution de systèmes	60
10.3	Existence des états mélangés stationnaires	62
10.4	Le théorème de Perron-Frobenius.	64
10.5	Sur le PageRank de Google	65
11	Diagonalisation en dimension 2	67
11.1	Motivations. Exemples de problèmes linéaires en dimension 2	67
11.2	Matrices diagonales et matrices diagonalisables	70
11.3	Valeurs propres et vecteurs propres	72
11.4	Le polynôme caractéristique	74
11.5	Synthèse	75
11.6	Applications aux problèmes d'évolution	77

Ce document présente en détail les résultats qui seront abordés en cours d'algèbre linéaire des groupes de « maths renforcées » du S2 PCST. Il peut servir d'aide-mémoire à condition d'avoir soi-même réfléchi et mis en pratique ces notions. La science n'étant toujours pas infuse, la seule possession de ces notes ne vous dispensera sans doute pas de faire quelques efforts pour assimiler cette matière, ô combien abstraite, mais aussi utile, qu'est l'algèbre linéaire !

1 L'espace vectoriel \mathbb{R}^n

1.1 Qu'est-ce que c'est ?

Définition 1.1. Soit $n \geq 1$, un entier donné. On note \mathbb{R}^n l'ensemble des n-uplets de réels, c'est-à-dire

$$\mathbb{R}^n = \{ \vec{v} = (x_1, x_2, \dots, x_n) \text{ avec } x_1, x_2, \dots, x_n \text{ nombres réels} \}.$$

Les éléments \vec{v} de \mathbb{R}^n s'appellent des **vecteurs**. Les différents nombres réels x_i le constituant sont ses **composantes**.

Le cas $n = 1$ donne la droite réelle \mathbb{R} , $n = 2$ le plan \mathbb{R}^2 , et $n = 3$ l'espace \mathbb{R}^3 .

On peut faire deux opérations principales dans ces ensembles.

- On peut **additionner** deux vecteurs **de même taille**.

Si $\vec{u} = (x_1, x_2, \dots, x_n)$ et $\vec{v} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$, on pose

$$\vec{u} + \vec{v} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \in \mathbb{R}^n.$$

On additionne deux vecteurs composante par composante.

- On peut **multiplier** un vecteur de \mathbb{R}^n par un **nombre réel** λ .

Si $\vec{u} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, on pose

$$\lambda \cdot \vec{u} = (\lambda x_1, \lambda x_2, \dots, \lambda x_n) \in \mathbb{R}^n.$$

On multiplie toutes les composantes de \vec{u} par le même nombre λ .

Muni de ces opérations, \mathbb{R}^n est ce que l'on appelle un **espace vectoriel**.

On note $\vec{0} = (0, 0, \dots, 0) \in \mathbb{R}^n$ le **vecteur nul** de \mathbb{R}^n . Attention, lorsque $n \geq 2$, de ne pas le confondre avec le nombre zéro, qui lui appartient à \mathbb{R} !

Les opérations sont définies algébriquement mais ont un sens géométrique. La multiplication d'un vecteur \vec{u} par un nombre réel λ , appelé **scalaire**, permet de changer son échelle en

gardant la même direction. La somme $\vec{u} + \vec{v}$ est la diagonale du parallélogramme s'appuyant sur \vec{u} et \vec{v} .

1.2 Les combinaisons linéaires

En combinant les deux opérations on fabrique les combinaisons linéaires de vecteurs.

Définition 1.2. Soient $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ p vecteurs d'un espace vectoriel E .

Une **combinaison linéaire** de $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ est une expression du type

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_p \vec{v}_p \quad \text{avec } \lambda_1, \lambda_2, \dots, \lambda_p \text{ réels.}$$

1.3 Droites et plans vectoriels

- Partant d'un seul vecteur non nul \vec{v} de \mathbb{R}^n , les multiples $\lambda \vec{v}$ décrivent la **droite vectorielle** (passant par $\vec{0}$) de direction \vec{v} (on dit aussi la droite **engendrée par \vec{v}**)

$$D = \{\lambda \vec{v} \mid \lambda \in \mathbb{R}\}.$$

- Si on se donne deux vecteurs **non colinéaires** \vec{v}_1 et \vec{v}_2 de \mathbb{R}^n , c'est-à-dire tel qu'aucun des deux ne soit multiple de l'autre, alors l'ensemble des combinaisons linéaires de \vec{v}_1 et \vec{v}_2 décrivent un **plan vectoriel**

$$P = \{\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \mid \lambda_1, \lambda_2 \in \mathbb{R}\}.$$

On dit aussi que P est engendré par \vec{v}_1 et \vec{v}_2 et on note $P = \text{Vect}(\vec{v}_1, \vec{v}_2)$.

Ainsi, droites et plans sont définis simplement en toute dimension n par des vecteurs générateurs, plutôt que par des équations.

1.4 La base canonique de \mathbb{R}^n

Dans \mathbb{R}^n , il y a n axes de coordonnées engendrés par les vecteurs

$$\begin{cases} \vec{e}_1 = (1, 0, 0, \dots, 0) \\ \vec{e}_2 = (0, 1, 0, \dots, 0) \\ \dots \\ \vec{e}_n = (0, 0, 0, \dots, 1) \end{cases}.$$

Tout vecteur $\vec{v} = (x_1, x_2, \dots, x_n)$ de \mathbb{R}^n se décompose de manière unique comme combinaison linéaire de ces vecteurs élémentaires. On a

$$\begin{aligned}\vec{v} &= (x_1, x_2, \dots, x_n) \\ &= (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n) \\ &= x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n.\end{aligned}$$

La famille $B_{can} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ s'appelle la **base canonique de \mathbb{R}^n** , et les composantes x_1, x_2, \dots, x_n s'appellent aussi les **coordonnées** de \vec{v} dans la base B_{can} .

2 Systèmes linéaires

2.1 Présentation du problème

L'objectif de ce chapitre est de donner une méthode générale de résolution de systèmes d'équations linéaires. En voici des exemples :

$$(S_1) \begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = -1 \\ x_1 - x_2 + 2x_3 = 1 \end{cases}, \quad (S_2) \begin{cases} x_1 + 2x_2 + x_3 = 1 \\ 2x_1 + 3x_2 - x_3 = 2 \end{cases}, \quad (S_3) \begin{cases} x_1 + 2x_2 = 1 \\ x_1 + 3x_2 = 2 \\ x_1 + 4x_2 = a \end{cases}.$$

avec a paramètre (donnée variable) dans (S_3) .

Le problème est de déterminer si ce type de systèmes a des solutions, et si oui, comment **toutes** les obtenir.

Pour fixer les notations en général, on s'intéresse aux systèmes linéaires (S) à **n équations** et **p inconnues** : x_1, x_2, \dots, x_p . Ils se présentent sous la forme

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = y_1 & (L_1) \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p = y_2 & (L_2) \\ \vdots & \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p = y_i & (L_i) \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p = y_n & (L_n) \end{cases},$$

où les nombres a_{ij} sont les **coefficients** (donnés) du système, et les variables y_1, \dots, y_n constituent le **second membre**, a priori connu, du système.

Il est important de se souvenir de la convention d'indexation des coefficients du système, car c'est celle adoptée dans tout le calcul matriciel :

a_{ij} désigne le j -ème coefficient, celui de l'inconnue x_j , dans la i -ème équation L_i . Autrement dit, le **premier** indice i est celui de la **ligne** et le **second** j de la **colonne**.

2.2 Systèmes linéaires et combinaisons linéaires

Tous les systèmes linéaires s'interprètent en termes de combinaisons linéaires dans \mathbb{R}^n , et inversement. Le système (S) ci-dessus se lit comme l'égalité de deux vecteurs de \mathbb{R}^n .

Proposition 2.1. *On a*

$$(S) \iff x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p = \vec{y}$$

avec

$$\vec{v}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \dots, \vec{v}_p = \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix} \text{ et } \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Les coordonnées des vecteurs de la combinaison linéaire se retrouvent **en colonnes** dans le système associé.

2.3 Codage matriciel d'un système linéaire

Il est pratique en informatique, mais aussi pour les calculs à la main, de coder le système à résoudre à l'aide de matrices. On peut ranger les coefficients du système (S) dans un tableau A à n lignes et p colonnes, appelée **matrice**

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

où a_{ij} est l'élément de A situé sur la i -ème ligne et la j -ème colonne de A .

Le système complet, avec son second membre, peut se coder avec une colonne de plus, séparée par un trait vertical : une matrice « augmentée ». Par exemple, on a

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & -1 & 2 & 1 \end{array} \right) \iff \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + x_3 = -1 \\ x_1 - x_2 + 2x_3 = 1 \end{cases}$$

2.4 Systèmes triangulaires

Définition 2.2. Un système linéaire (S) est dit **triangulaire** si

- le nombre d'équations est égal au nombre d'inconnues : $n = p$,
- et les coefficients a_{ij} sont nuls sous la diagonale : $a_{ij} = 0$ si $i > j$.

Autrement dit, la matrice des coefficients de (S) est carrée avec des zéros sous la diagonale.

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \quad \quad \quad \ddots \quad \quad \quad \vdots \\ a_{nn}x_n = y_n \end{cases} \iff \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & y_1 \\ 0 & a_{22} & \cdots & a_{2n} & y_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{nn} & y_n \end{array} \right)$$

Exemple. Soit a un **paramètre réel**, c'est-à-dire une **donnée variable** du problème (pas une inconnue). Le système

$$(S) \begin{cases} x + y + z = 0 \\ ay - 2z = 1 \\ z = 2 \end{cases} \text{ est triangulaire.}$$

On le résout en partant de la dernière équation, et en remontant de proche en proche. On obtient

$$(S) \iff \begin{cases} x + y + z = 0 \\ ay = 2z + 1 = 5 \\ z = 2 \end{cases} \iff \begin{cases} x = -\frac{5}{a} - 2 \\ y = \frac{5}{a} \\ z = 2 \end{cases} \text{ si } a \neq 0.$$

Le système triangulaire (S) a donc une **unique solution** si le paramètre $a \neq 0$, et **n'en a pas** si $a = 0$ (deuxième ligne incompatible $0.y = 5$).

Proposition 2.3. On dit qu'un système triangulaire est **de Cramer** si **tous ses coefficients diagonaux** a_{ii} sont non nuls. Dans ce cas il possède une **unique solution**.

Il n'y a alors aucune obstruction à le résoudre de proche en proche en partant de la dernière équation. On obtient d'abord x_n que l'on injecte au dessus, etc...

On peut aussi travailler avec les matrices en utilisant deux sortes d'opérations. Multiplier une ligne par une constante non nulle, pour obtenir des 1 sur la diagonale, et ajouter un multiple de cette ligne au dessus pour faire disparaître les coefficients hors diagonale.

Dans l'exemple ci-dessus, pour $a \neq 0$,

$$\begin{aligned} (S) &\Leftrightarrow \begin{pmatrix} 1 & 1 & 1 & | & 0 \\ 0 & a & -2 & | & 1 \\ 0 & 0 & 1 & | & 2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 1 & 0 & | & -2 \\ 0 & a & 0 & | & 5 \\ 0 & 0 & 1 & | & 2 \end{pmatrix} \begin{array}{l} \leftarrow L_1 - L_3 \\ \leftarrow L_2 + 2L_3 \end{array} \\ &\Leftrightarrow \begin{pmatrix} 1 & 1 & 0 & | & -2 \\ 0 & 1 & 0 & | & 5/a \\ 0 & 0 & 1 & | & 2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & | & -2 - 5/a \\ 0 & 1 & 0 & | & 5/a \\ 0 & 0 & 1 & | & 2 \end{pmatrix} \leftarrow L_1 - L_2 \end{aligned}$$

La solution est dans la dernière colonne, puisque le système associé s'écrit $x = -2 - 5/a$, $y = 5/a$ et $z = 2$.

2.5 Systèmes échelonnés

Définition 2.4. On dit qu'un système linéaire (S) de n équations et p inconnues est **échelonné** si

- chaque ligne de coefficients a_{ij} non nulle commence par davantage de zéros que la précédente,
- si la i -ème ligne de coefficients a_{ij} est nulle, alors les lignes suivantes le sont aussi.

Le **premier** coefficient non nul a_{ij} d'une ligne i donnée s'appelle un **pivot** du système.

Par exemple • Les systèmes triangulaires de Cramer sont échelonnés, et les coefficients diagonaux a_{ii} (tous non nuls par hypothèse) sont des pivots.

- Les systèmes suivants :

$$\begin{aligned} (S_1) \begin{cases} \boxed{x} - y + z = 1 \\ \boxed{2y} + z = 2 \end{cases}, & (S_2) \begin{cases} \boxed{x} + y + z = 0 \\ \boxed{z} = 1 \end{cases}, \\ (S_3) \begin{cases} \boxed{x} + 2y = 0 \\ \boxed{3y} = 1 \\ 0 = a \end{cases} \end{aligned}$$

sont échelonnés.

Les coefficients et inconnues pivots seront toujours encadrés par soucis de lisibilité et leur importance dans toute la suite.

- Par contre, le système

$$\begin{cases} x + 2y + 3z + t = 1 \\ z - t = 2 \\ 2z + t = 3 \end{cases}$$

n'est pas échelonné, à cause de la présence d'un coefficient en z dans la 3-ème équation.

Inconnues principales ou non Les systèmes échelonnés possèdent deux sortes d'inconnues.

Définition 2.5. — Les inconnues x_{j_i} correspondant aux coefficients pivots a_{ij_i} d'un système échelonné s'appellent les **inconnues principales**.
— Les autres inconnues s'appellent les **inconnues non principales** (ou secondaires ou libres).

Par exemple • Toutes les inconnues des systèmes triangulaires de Cramer sont principales.

• Dans les exemples précédents :

- (S_1) : inconnues principales x et y , inconnue non principale z ,
- (S_2) : inconnues principales x et z , inconnue non principale y ,
- (S_3) : x et y principales.

Équations de compatibilité. Dans un système échelonné général, il peut apparaître des lignes d'équations du type

$$\begin{cases} 0 = & y_i & (L_i) \\ 0 = & y_{i+1} & (L_{i+1}) \\ & \dots & \\ 0 = & y_n & (L_n) \end{cases} .$$

Ces équations ne portent pas sur les **inconnues** x_i mais sur les **données** y_i du second membre. Elles sont des **contraintes** sur les données pour que le système puisse avoir des solutions. Elles s'appellent les **équations de compatibilité** de (S) . Si elles sont satisfaites (ou si il n'y en a pas), on dit que le système est **compatible**.

Par exemple, les systèmes **homogènes**, c'es-à-dire dont le second membre est nul, sont compatibles, puisque le vecteur nul est solution. Le système (S_3) ci-dessus n'a pas de solution si le paramètre a n'est pas nul.

Résolution. La technique de résolution est alors la suivante.

Pour résoudre un système échelonné compatible, il suffit de passer les inconnues non principales dans le second membre. Ces inconnues non principales ne peuvent être déterminées, elles sont libres, et deviennent des **paramètres**. Le système devient alors un système triangulaire de Cramer par rapport aux inconnues principales, dont les solutions sont déterminées par la valeur des inconnues non principales.

On traite les exemples précédents.

$$(S_1) \iff \begin{cases} x - y = 1 - z \\ y = 1 - \frac{z}{2} \\ z \text{ quelconque} \end{cases} \iff \begin{cases} x = 2 - \frac{3}{2}z \\ y = 1 - \frac{z}{2} \\ z \text{ quelconque} \end{cases}$$

(S_1) possède donc **une infinité** de solutions paramétrée par l'inconnue non principale z . Si on travaille avec les matrices, cela donne :

$$\begin{aligned} (S_1) &\Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 1 \\ 0 & \boxed{2} & 1 & 2 \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 1 \\ 0 & \boxed{1} & 1/2 & 1 \end{array} \right) \\ &\Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 0 & 3/2 & 2 \\ 0 & \boxed{1} & 1/2 & 1 \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 0 & 2 - 3z/2 & 2 \\ 0 & \boxed{1} & 1 - z/2 & 1 \end{array} \right) \end{aligned}$$

- De la même façon, (S_2) a une infinité de solutions avec y quelconque, $x = -1 - y$ et $z = 1$.

- Enfin, (S_3) n'a pas de solution si $a \neq 0$, et possède une solution unique $y = 1/3$ et $x = -2/3$ si $a = 0$.

Proposition 2.6. — *Un système échelonné possède des solutions si et seulement si ses équations de compatibilité sont satisfaites.*

— *Une solution d'un système échelonné compatible est déterminée de manière unique par la donnée des valeurs des inconnues non principales.*

2.6 Systèmes équivalents et transformations élémentaires

Définition 2.7. On dit que deux systèmes d'équations (S) et (S') sont **équivalents** s'ils ont même espace de solutions.

Cela signifie que soit aucun des deux systèmes n'a de solution, soit que les deux systèmes ont des solutions et que ce sont les mêmes.

On peut transformer un système en un système équivalent en appliquant une suite de **transformations élémentaires** suivantes :

1. Intervertir deux équations $L_i \leftrightarrow L_j$.
2. Ajouter à une équation L_i un multiple d'une **autre équation** L_j . Autrement dit, on laisse toutes les équations inchangées, sauf la i -ème, $L_i \rightarrow L'_i = L_i + \lambda L_j$ avec $i \neq j$.
3. Multiplier une équation L_i par un nombre **non nul** : $L_i \rightarrow L'_i = \lambda L_i$ avec $\lambda \neq 0$.

Toutes ces opérations donnent des systèmes (S) et (S') équivalents car on peut revenir au système initial. C'est clair pour 1 et 3 ($L'_i \rightarrow L_i = L'_i/\lambda$ car $\lambda \neq 0$). Pour 2, partant de $L'_i = L_i + \lambda L_j$ et $L'_j = L_j$ (inchangée), on retrouve $L_i = L'_i - \lambda L'_j$.

Attention. Il est essentiel de réaliser ces opérations élémentaires de **manière séquentielle**, et non simultanément, pour avoir des systèmes équivalents et conserver l'espace des solutions. Par exemple, un système à trois équations

$$(S) : L_1, L_2, L_3 \text{ implique } (S') : L'_1 = L_1 - L_2, L'_2 = L_2 - L_3 \text{ et } L'_3 = L_3 - L_1,$$

mais (S') **n'implique pas** (S) en général car on ne peut retrouver les équations de départ à l'aide des L'_i . Ces nouvelles équations L'_i sont liées par la relation $L'_1 + L'_2 + L'_3 = 0$, on n'a plus que deux équations indépendantes !

2.7 L'algorithme du pivot de Gauss–Jordan

On peut discuter le résultat et la technique clé du chapitre.

Théorème 2.8. *Tout système linéaire (S) est équivalent à un système échelonné (S') obtenu par transformations élémentaires.*

En particulier (S) a des solutions si et seulement si les équations de compatibilité de (S') sont satisfaites, auquel cas, ces solutions sont uniquement déterminées (paramétrées) par la donnée des inconnues non principales de (S') .

Démonstration. La méthode consiste à éliminer progressivement le nombre d'inconnues dans les équations successives par la **méthode du pivot**. Le système initial est de la forme

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = y_1 & L_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = y_2 & L_2 \\ \cdots & \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ip}x_p = y_i & L_i \\ \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p = y_n & L_n \end{cases}$$

- Si $a_{11} \neq 0$, on garde L_1 et on l'utilise pour éliminer l'inconnue x_1 des équations L_2, \dots, L_n par des transformations élémentaires de type 2,

$$L_2 \rightarrow L'_2 = L_2 - \frac{a_{21}}{a_{11}}L_1, \quad L_i \rightarrow L'_i = L_i - \frac{a_{i1}}{a_{11}}L_1, \quad L_n \rightarrow L'_n = L_n - \frac{a_{n1}}{a_{11}}L_1.$$

Le nouveau système équivalent (S') est de la forme

$$(S') \left\{ \begin{array}{l} \boxed{a_{11}x_1} + a_{12}x_2 + \cdots + a_{1p}x_p = y_1 \quad L_1 \\ a'_{22}x_2 + \cdots + a'_{2p}x_p = y'_2 \quad L'_2 \\ \cdots \\ a'_{i2}x_2 + \cdots + a'_{ip}x_p = y'_i \quad L'_i \\ \cdots \\ a'_{n2}x_2 + \cdots + a'_{np}x_p = y'_n \quad L'_n \end{array} \right.$$

L'inconnue x_1 est principale, on poursuit l'algorithme avec le sous-système d'équations L'_2, \dots, L'_n , en essayant d'éliminer x_2 de L'_3 , etc.

• **Problème** : il peut se produire au démarrage que $\boxed{a_{11} = 0}$, auquel cas il ne peut pas servir de pivot. On peut alors chercher un coefficient $a_{i1} \neq 0$ dans la première colonne du système (S) , et l'amener en haut à gauche à l'aide d'une permutation des lignes $L_1 \leftrightarrow L_i$ pour démarrer l'élimination. Si tous les coefficients a_{i1} de la colonne sont nuls, il n'y a rien à faire par rapport à x_1 , et on passe à l'inconnue x_2 .

• L'algorithme s'arrête sur un système échelonné quand il n'y a plus d'inconnues à traiter dans les sous-systèmes. □

2.8 Quelques exemples

- On reprend les systèmes de l'introduction.

Pour la rédaction, il est pratique d'encadrer au fur et à mesure les inconnues qui servent de pivot, car alors les transformations élémentaires de type 2 utilisées sont déterminées.

On les rappelle ici dans le premier exemple.

$$(S_1) \left\{ \begin{array}{l} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = -1 \\ x_1 - x_2 + 2x_3 = 1 \end{array} \right. \iff \left\{ \begin{array}{l} \boxed{x_1} + x_2 + x_3 = 1 \quad L_1 \\ -x_2 - x_3 = -3 \quad L_2 \leftarrow L_2 - 2L_1 \\ -2x_2 + x_3 = 0 \quad L_3 \leftarrow L_3 - L_1 \end{array} \right.$$

$$\iff \left\{ \begin{array}{l} \boxed{x_1} + x_2 + x_3 = 1 \quad L_1 \\ -\boxed{x_2} - x_3 = -3 \quad L_2 \\ \boxed{3x_3} = 6 \quad L_3 \leftarrow L_3 - 2L_2 \end{array} \right.$$

$$\iff \left\{ \begin{array}{l} x_1 = 1 - x_2 - x_3 = -2 \\ x_2 = 3 - x_3 = 1 \\ x_3 = 2 \end{array} \right.$$

Le système (S_1) a donc une unique solution $(x_1, x_2, x_3) = (-2, 1, 2)$.

- On passe à $(S_2) \iff$

$$\begin{aligned} \begin{cases} \boxed{x_1} + 2x_2 + x_3 = 1 \\ 2x_1 + 3x_2 - x_3 = 2 \end{cases} &\iff \begin{cases} \boxed{x_1} + 2x_2 + x_3 = 1 \\ \boxed{-x_2} - 3x_3 = 0 \end{cases} \\ &\iff \begin{cases} x_3 = \text{quelconque} \\ x_2 = -3x_3 \\ x_1 = 1 - 2x_2 - x_3 = 1 + 5x_3 \end{cases} \end{aligned}$$

Le système (S_2) a une infinité de solutions paramétrées par l'inconnue non principale x_3 . Notez que l'ensemble des solutions est une droite passant par le point $(1, 0, 0)$ et de vecteur directeur $\vec{v} = (5, -3, 1)$, puisque

$$(x_1, x_2, x_3) = (1 + 5x_3, -3x_3, x_3) = (1, 0, 0) + x_3 \vec{v}.$$

- On a $(S_3) \iff$

$$\begin{aligned} \begin{cases} \boxed{x_1} + 2x_2 = 1 \\ x_1 + 3x_2 = 2 \\ x_1 + 4x_2 = a \end{cases} &\iff \begin{cases} \boxed{x_1} + 2x_2 = 1 \\ x_2 = 2 - 1 = 1 \\ 2x_2 = a - 1 \end{cases} \\ &\iff \begin{cases} \boxed{x_1} + 2x_2 = 1 \\ \boxed{x_2} = 1 \\ 0 = a - 3 \end{cases} \end{aligned}$$

Ainsi, (S_3) ne possède pas de solution si le **paramètre** $a \neq 3$. Si $a = 3$, (S_3) est compatible et possède une solution unique $x_1 = -1$ et $x_2 = 1$.

2.9 La notion de rang

Définition 2.9. Le **rang** d'un système linéaire est le **nombre d'inconnues principales** d'un système échelonné équivalent. Il ne dépend que de la matrice A des coefficients du système, pas du second membre. Par extension, le rang d'une matrice A est le rang d'un système associé, par exemple du système homogène $(A \mid 0)$.

Comme il y a plusieurs façons d'échelonner un système (on pourrait changer l'ordre des équations ou même des inconnues), conduisant en général à différents choix d'inconnues principales, il nous faudra comprendre pourquoi ce nombre est bien défini indépendamment des choix.

En attendant, nous pouvons constater son importance dans la discussion de l'existence ou non, et de l'unicité ou non des solutions d'un système.

Théorème 2.10. Soit (S) un système linéaire à n équations, p inconnues et de rang r . Alors on a nécessairement $r \leq n$ et $r \leq p$, et un des cas suivants

	$r = p$	$r < p$
$r = n$	solution unique	infinité de solutions
$r < n$	au plus une solution	pas de solution ou une infinité

- Remarques 2.11.*
1. Lorsque $r = n = p$, le système s'appelle un **système de Cramer**. Il possède une solution unique.
 2. Dans le cas $r < n$, il y a des équations de compatibilité. Le système possède des solutions si et seulement si elles sont satisfaites.
 3. Il n'est pas nécessaire de résoudre le système pour connaître son rang, et connaître ainsi le nombre de solutions **sans** les calculer ! Il suffit d'échelonner le système et de compter les pivots.

Démonstration. Le rang r est le nombre d'inconnues principales. Comme il y a p inconnues en tout, on a $r \leq p$, et de plus $p - r$ est le nombre d'inconnues non principales.

De plus, il y a au plus une inconnue principale par équation d'un système échelonné. On a donc aussi $r \leq n$ et $n - r$ est le nombre d'équations de compatibilité.

Si $r = n$, il n'y a pas d'équations de compatibilité, le système possède donc au moins une solution.

Si de plus $r < p$, il y a une infinité de solution paramétrées par les $p - r$ inconnues non principales. Si $r = p$, il n'y a pas d'inconnue non principale et la solution est unique. \square

3 Matrices et applications linéaires

3.1 Applications linéaires de \mathbb{R}^p dans \mathbb{R}^n

Les applications $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ que nous allons étudier font « bon ménage » avec les combinaisons linéaires. Voici une première définition abstraite.

Définition 3.1. Une application $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est dite **linéaire** si pour tout $\vec{u}, \vec{v} \in \mathbb{R}^p$, et tout $\lambda \in \mathbb{R}$, on a :

- a) $f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$ et
- b) $f(\lambda \vec{v}) = \lambda f(\vec{v})$.

On peut aussi regrouper les deux propriétés en disant que :

$$f(\vec{u} + \lambda \vec{v}) = f(\vec{u}) + \lambda f(\vec{v}).$$

Écrit de cette façon, on voit une première propriété des applications linéaires :

- Une application linéaire satisfait toujours $f(\vec{0}) = \vec{0}$ (vient de b).
- Une application linéaire transforme une droite vectorielle en une droite vectorielle, ou $\vec{0}$, et une droite affine en une droite affine, ou un point.

Démonstration. Soit D la droite passant par \vec{u}_0 de direction \vec{u} . Alors \vec{v} est dans D s'il s'écrit $\vec{v} = \vec{u}_0 + \lambda \vec{u}$ pour un réel λ .

D'où $f(\vec{v}) = f(\vec{u}_0) + \lambda f(\vec{u})$ est sur la droite passant par $f(\vec{u}_0)$ et de direction $f(\vec{u})$ si $f(\vec{u}) \neq \vec{0}$. Notez que si $f(\vec{u}) = \vec{0}$ alors l'image $f(D)$ se réduit au point $f(\vec{u}_0)$. \square

On rappelle (cf §1.4) que l'on note $B_c = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ la base canonique de \mathbb{R}^p . Une application linéaire est entièrement déterminée par l'image de cette base dans \mathbb{R}^n .

Théorème 3.2. Soient $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ des vecteurs **quelconques** de \mathbb{R}^n . Alors il existe une **unique** application linéaire $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ telle que

$$f(\vec{e}_1) = \vec{v}_1, f(\vec{e}_2) = \vec{v}_2, \dots, f(\vec{e}_p) = \vec{v}_p.$$

Si $\vec{u} = (x_1, x_2, \dots, x_p)$ alors on a

$$f(\vec{u}) = x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p.$$

Démonstration. On a $\vec{u} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_p \vec{e}_p$, d'où nécessairement si f est linéaire

$$f(\vec{u}) = x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) + \dots + x_p f(\vec{e}_p) = x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p.$$

Le fait que cette formule définisse une application linéaire provient de la « linéarité » des composantes des vecteurs par somme et multiplication scalaire. \square

Par exemple dans \mathbb{R}^2 , avec sa base canonique $\vec{e}_1 = (1, 0)$ et $\vec{e}_2 = (0, 1)$, la connaissance de $f(\vec{e}_1)$ et $f(\vec{e}_2)$ détermine f tout entière. Géométriquement, $f(x, y) = xf(\vec{e}_1) + yf(\vec{e}_2)$ est la diagonale du parallélogramme de côtés $xf(\vec{e}_1)$ et $yf(\vec{e}_2)$.

3.2 Matrice d'une application linéaire

Une application linéaire f de \mathbb{R}^p dans \mathbb{R}^n dépend donc de la donnée de p vecteurs de \mathbb{R}^n et se code en pratique avec une matrice, c'est-à-dire un tableau de nombres. Concrètement,

dans l'espace d'arrivée \mathbb{R}^n , les vecteurs image de la base s'écrivent

$$\left\{ \begin{array}{l} \vec{v}_1 = f(\vec{e}_1) = (a_{11}, a_{21}, \dots, a_{n1}) \\ \vec{v}_2 = f(\vec{e}_2) = (a_{12}, a_{22}, \dots, a_{n2}) \\ \vdots \\ \vec{v}_j = f(\vec{e}_j) = (a_{1j}, a_{2j}, \dots, a_{nj}) \\ \vdots \\ \vec{v}_p = f(\vec{e}_p) = (a_{1p}, a_{2p}, \dots, a_{np}) \end{array} \right.$$

Si $\vec{u} = (x_1, x_2, \dots, x_p)$ et $\vec{v} = f(\vec{u}) = (y_1, y_2, \dots, y_n)$ alors d'après la proposition 2.1, l'égalité

$$\vec{v} = (y_1, y_2, \dots, y_n) = x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p = f(\vec{u}) \quad (1)$$

se traduit de la façon suivante :

$$f : X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \mapsto Y = \begin{pmatrix} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ \vdots \\ y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p \\ \vdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p \end{pmatrix} \quad (2)$$

Le passage de l'égalité vectorielle (1) au système (2) se fait en mettant **en colonnes** les coordonnées des vecteurs $\vec{v}_j = f(\vec{e}_j)$ de l'image de la base.

Définition 3.3. La **matrice** de l'application linéaire f (dans les bases canoniques) est

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

où a_{ij} est l'élément de A situé sur la i -ème ligne et la j -ème colonne de A . (C'est la même convention d'indice que pour les systèmes linéaires.)

Partant de A , on retrouve l'image de la base canonique de départ $f(\vec{e}_1), f(\vec{e}_2), \dots, f(\vec{e}_p)$ dans les différentes **colonnes** de A : les coordonnées de $f(\vec{e}_j)$ se trouvent dans la j -ème

colonne de A .

$$A = \text{Mat}(f) = \begin{array}{ccc} f(\vec{e}_1) & f(\vec{e}_2) & f(\vec{e}_p) \\ \downarrow & \downarrow & \downarrow \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix} \end{array}$$

Notez que la largeur p de A est la dimension de l'espace de départ, et la hauteur n la dimension de l'espace d'arrivée.

3.3 Calcul matriciel de l'image d'un vecteur

On voit dans la formule développée (2) que les coordonnées de l'image $\vec{v} = (y_1, y_2, \dots, y_n)$ de $\vec{u} = (x_1, x_2, \dots, x_p)$ par f apparaissent **en colonnes**.

On prend alors comme convention (déjà utilisée dans la manipulation des systèmes linéaires), **d'écrire systématiquement les vecteurs en colonnes en calcul matriciel**, et de les nommer par des majuscules (comme des matrices) :

$$\vec{u} = (x_1, x_2, \dots, x_p) \iff X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \text{ et } \vec{v} = (y_1, y_2, \dots, y_n) \iff Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Pour calculer l'image du vecteur $\vec{u} = (x_1, x_2, \dots, x_p)$ de \mathbb{R}^p à l'aide de la matrice A , on écrit

$$Y = AX = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} y_1 \\ \mathbf{y}_2 \\ \vdots \\ y_n \end{pmatrix}$$

La méthode de calcul, par exemple de y_2 , est de faire la somme

$$y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p$$

en **balayant simultanément** de gauche à droite la seconde ligne de A , à hauteur de y_2 , et de haut en bas la colonne X .

On peut aussi se souvenir que la colonne Y est $x_1 \times$ première colonne de A + $x_2 \times$ seconde colonne + etc, comme il vient de (2) ou de (1) :

$$f : \vec{u} \mapsto \vec{v} = f(\vec{u}) = x_1 f(\vec{u}_1) + x_2 f(\vec{u}_2) + \cdots + x_p f(\vec{u}_p)$$

\updownarrow passage en colonne

$$A : X \mapsto Y = AX = x_1 C_1 + x_2 C_2 + \cdots + x_p C_p,$$

| où C_j est la j -ème colonne de A donnant les coordonnées de $f(\vec{u}_j)$.

Exemple. La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ est celle d'une application linéaire

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \\ X \mapsto Y = AX$$

Pour calculer $f(1, -1, 2)$ on écrit

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \times 1 + 2 \times (-1) + 3 \times 2 \\ 4 \times 1 + 5 \times (-1) + 6 \times 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}.$$

Autrement dit : $f(1, -1, 2) = (5, 11)$.

Attention, à la compatibilité des longueurs en calcul matriciel. La largeur de A doit coïncider avec la hauteur du vecteur colonne X ; car c'est la dimension de l'espace de départ.

Par exemple le produit suivant

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ n'existe pas!}$$

Il ne faut pas compléter les « composantes fantômes » par des 0 s'il en manque. Ce calcul n'a pas de sens car A est associée à une application de \mathbb{R}^3 dans \mathbb{R}^2 mais $(1, -1)$ est un vecteur de \mathbb{R}^2 .

3.4 Exemples d'applications linéaires du plan

Comme premiers exemples, on va étudier une famille d'applications linéaires $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ d'origine géométrique.

En général, une application linéaire de $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est déterminée par une matrice 2×2

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

par la formule

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ X = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto Y = AX = \begin{pmatrix} ax + cy \\ bx + dy \end{pmatrix}.$$

Certaines de ces applications ont une importance géométrique : les **similitudes directes**.

Définition 3.4. Soit $\lambda = a + ib$ un nombre complexe. L'application

$$\begin{aligned} f_\lambda : \mathbb{C} &\rightarrow \mathbb{C} \\ z = x + iy &\mapsto \lambda z \end{aligned}$$

s'appelle une *similitude directe* du plan.

Comme on peut assimiler \mathbb{C} à \mathbb{R}^2 par $z = x + iy \leftrightarrow (x, y)$, on voit que

$$\lambda z = (a + ib)(x + iy) = ax - by + i(bx + ay) \leftrightarrow (ax - by, bx + ay)$$

Autrement dit, la similitude f_λ est l'application linéaire de \mathbb{R}^2 dans lui-même associée à la matrice

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Voici leur interprétation géométrique.

Les homothéties. Si λ est un réel a , alors $f_\lambda(z) = az$ est l'**homothétie** de rapport a . Sa matrice est

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

Elle effectue un « zoom » de rapport a dans le plan. Si $a < 0$, f_a effectue une symétrie centrale $z \mapsto -z$, suivi d'une homothétie de rapport $-a$.

Pour $a = 1$, $f_1 : (x, y) \mapsto (x, y)$ laisse tous les vecteurs inchangés. Cette application s'appelle l'**identité**, et on note sa matrice

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Les rotations. Si $\lambda = e^{i\theta} = \cos \theta + i \sin \theta$ alors $f_\lambda(z) = e^{i\theta}z$ est la rotation d'angle θ . C'est une application linéaire de matrice

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (3)$$

Notez que $R_\theta(\vec{e}_1) = (\cos \theta, \sin \theta)$ est bien la première colonne de A , tandis que $R_\theta(\vec{e}_2) = (-\sin \theta, \cos \theta)$ est la deuxième colonne. On a

$$R_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta).$$

Cela donne par exemple les formules d'addition des angles pour \cos et \sin .

En effet, on a

$$\begin{aligned} \begin{pmatrix} \cos(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) \end{pmatrix} &= R_{\theta_1} \begin{pmatrix} \cos \theta_2 \\ \sin \theta_2 \end{pmatrix} = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix} \begin{pmatrix} \cos \theta_2 \\ \sin \theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 \\ \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2 \end{pmatrix}. \end{aligned}$$

Rotation dans \mathbb{R}^3 . Un exemple classique lié est celui de la rotation dans l'espace \mathbb{R}^3 , d'angle θ et d'axe vertical (Oz) engendré par $\vec{e}_3 = (0, 0, 1)$.

On a ici,

$$R_\theta(\vec{e}_1) = (\cos \theta, \sin \theta, 0), \quad R_\theta(\vec{e}_2) = (-\sin \theta, \cos \theta, 0) \quad \text{et} \quad R_\theta(\vec{e}_3) = \vec{e}_3.$$

La matrice de R_θ est donc

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

à ne pas confondre avec la matrice (3) de la rotation *dans le plan*. On a à faire ici à une application linéaire de \mathbb{R}^3 dans \mathbb{R}^3 , et donc à une matrice de taille 3×3 .

Similitude générale. On revient sur le cas général des similitudes du plan.

On peut écrire λ sous la forme $\rho e^{i\theta}$, avec $\rho = |\lambda| = \sqrt{a^2 + b^2}$ et θ l'argument de λ . Alors $f_\lambda(z) = |\lambda| e^{i\theta} z$ est la composée d'une rotation d'angle θ et d'une homothétie de rapport $|\lambda|$. Les dimensions sont multipliées par $|\lambda|$ mais les formes des figures sont préservées (un carré devient un carré), car les angles sont préservés. En effet si $z_2 = e^{i\varphi} z_1$ alors $f_\lambda(z_2) = |\lambda| e^{i\theta} e^{i\varphi} z_1 = e^{i\varphi} f_\lambda(z_1)$.

Autres exemples. On rappelle qu'en général, les applications linéaires de \mathbb{R}^2 ne respectent pas les angles. En fait, on peut transformer le repère orthonormé $\vec{e}_1 = (1, 0)$, $\vec{e}_2 = (0, 1)$ en n'importe quel autre couple de vecteurs $f(\vec{e}_1) = (a, b)$ et $f(\vec{e}_2) = (c, d)$ avec la matrice $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. On peut même « écraser » le repère. Par exemple, les matrices

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad Q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

sont respectivement associées aux applications linéaires $p : (x, y) \mapsto (x, 0)$ et $q : (x, y) \mapsto (0, y)$ qui sont les **projections orthogonales** sur les axes des x et des y .

Les **symétries orthogonales** suivant les axes sont aussi des applications linéaires de matrices

$$S_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{et} \quad S_y = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

4 Opérations sur les matrices et les applications linéaires

4.1 Structure d'espace vectoriel

Notations. On note $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ l'ensemble des applications linéaires $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$, et $M_{n,p}(\mathbb{R})$ l'ensemble des matrices réelles à n lignes et p colonnes.

Somme. On peut additionner deux applications linéaires f et $g \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ en écrivant pour tout $\vec{u} \in \mathbb{R}^p$,

$$(f + g)(\vec{u}) = f(\vec{u}) + g(\vec{u}).$$

De la même façon, on peut additionner deux matrices **de même taille**, en ajoutant les coefficients terme à terme. Si $A = (a_{ij})$ et $B = (b_{ij}) \in M_{n,p}(\mathbb{R})$ alors on pose

$$A + B = (a_{ij} + b_{ij})$$

Exemple. Si $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 8 & 7 \\ 6 & 5 \end{pmatrix}$ alors $A + B = \begin{pmatrix} 9 & 9 \\ 9 & 9 \end{pmatrix}$.

Ces deux opérations sont compatibles.

Proposition 4.1. Si $A = \text{Mat}(f)$ et $B = \text{Mat}(g)$, alors $A + B = \text{Mat}(f + g)$.

Multiplication par un réel. Si $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ et $\lambda \in \mathbb{R}$ alors on définit λf par

$$(\lambda f)(\vec{u}) = \lambda f(\vec{u}).$$

De la même façon, si $A = (a_{ij}) \in M_{n,p}(\mathbb{R})$ et $\lambda \in \mathbb{R}$, alors on définit

$$\lambda A = (\lambda a_{ij}) \in M_{n,p}(\mathbb{R})$$

Exemple. Si $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ alors $2A = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$.

Ces deux opérations sont compatibles.

Proposition 4.2. Si $A = \text{Mat}(f)$ alors $\lambda A = \text{Mat}(\lambda f)$.

Structure d'espaces vectoriels. Munis de ces deux opérations de somme et de multiplication scalaire, on voit que $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ et $M_{n,p}(\mathbb{R})$ sont eux-mêmes des espaces vectoriels. En fait de ce point de vue, on peut identifier

$$\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n) \simeq M_{n,p}(\mathbb{R}) \simeq \mathbb{R}^{np}.$$

4.2 Produit de matrices et composition d'applications linéaires

Composition d'applications linéaires. Quand on a deux applications

$$f : \mathbb{R}^q \rightarrow \mathbb{R}^n \quad \text{et} \quad g : \mathbb{R}^p \rightarrow \mathbb{R}^q,$$

alors on peut définir l'**application composée** $f \circ g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ en posant

$$(f \circ g)(\vec{u}) = f(g(\vec{u})).$$

Notez qu'il faut que l'espace d'arrivée de g coïncide avec celui du départ de f , égale à \mathbb{R}^q ici.

Proposition 4.3. *Si f et g sont des applications linéaires composables, alors $f \circ g$ est aussi une application **linéaire**.*

Démonstration. Pour $\vec{u}, \vec{v} \in \mathbb{R}^q$ et $\lambda \in \mathbb{R}$, on a

$$\begin{aligned} (f \circ g)(\vec{u} + \lambda \vec{v}) &= f(g(\vec{u} + \lambda \vec{v})) \\ &= f(g(\vec{u}) + \lambda g(\vec{v})) && \text{par linéarité de } g, \\ &= f(g(\vec{u})) + \lambda f(g(\vec{v})) && \text{par linéarité de } f, \\ &= (f \circ g)(\vec{u}) + \lambda (f \circ g)(\vec{v}). \end{aligned}$$

□

Produit de matrices. L'opération de composition d'applications linéaires est naturelle, et définie sans faire appel aux composantes des vecteurs. Il reste à exprimer la matrice de la composée $f \circ g$ à l'aide des matrices de f et g dans les bases canoniques.

Soient $f \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^n)$ et $g \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$. On note

$$A = \text{Mat}(f), \quad B = \text{Mat}(g) \quad \text{et} \quad C = \text{Mat}(f \circ g).$$

Le problème est de calculer C à l'aide de A et B .

On revient à la définition de $C = \text{Mat}(f \circ g)$. Sa j -ème colonne C_j sont les coordonnées de $(f \circ g)(\vec{e}_j) = f(g(\vec{e}_j))$, et donc

$$\begin{aligned} C_j &= \text{Mat}(f)(j\text{-ème colonne de } \text{Mat}(g)) \\ &= A(B_j) \end{aligned}$$

où B_j est la j -ème colonne de B .

1. $(AB)C = A(BC)$ si les produits existent,
2. $(A + B)C = AC + BC$ et $A(B + C) = AB + AC$.

Attention, On ne pas toujours composer deux applications linéaires f et g . Pour définir $f \circ g$, il faut pour cela que l'espace d'arrivée de g soit celui de départ de f . De la même façon, la produit AB de deux matrices n'est défini que **si la largeur de A coïncide avec la hauteur de B**.

Par exemple, les produits $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}$ n'existent pas.

Définition 4.5. - Une application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ s'appelle un **endomorphisme**. On note $\text{End}(\mathbb{R}^n) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ cet espace.

- Les matrices associées aux endomorphismes sont des matrices carrées. On note $M_n(\mathbb{R})$ cet espace.

On peut toujours composer deux endomorphismes de $\text{End}(\mathbb{R}^n)$, et on peut donc toujours multiplier deux matrices carrées **de même taille**.

Ce produit possède un **élément neutre** : l'application **identité**

$$\begin{aligned} \text{Id} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \vec{v} &\mapsto \vec{v} \end{aligned}$$

a pour matrice la **matrice identité** de taille n ,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

On a

$$f \circ \text{Id} = \text{Id} \circ f = f \text{ et } AI_n = I_nA = A \text{ pour toute } M \in M_n(\mathbb{R}).$$

Cependant, le produit obtenu sur $M_n(\mathbb{R})$ n'est :

1. ni commutatif : on a n'a pas $AB = BA$ pour $A, B \in M_n(\mathbb{R})$ en général,
2. ni intègre. C'est à dire que $AB = AC$ et $A \neq 0$ n'implique pas $B = C$ en général.

Exemple. Si on prend $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. On trouve que

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

$$A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \text{ avec } A \neq 0!$$

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $X \mapsto AX$ l'application linéaire associée à A . L'égalité $A^2 = 0$ signifie que $\text{Mat}(f)^2 = \text{Mat}(f^2) = 0$, c'est-à-dire $f \circ f = 0$.

5 Isomorphismes et matrices inversibles

5.1 Les isomorphismes

On rappelle quelques propriétés générales des applications, linéaires ou pas. Elles portent sur le nombre de solutions de l'équation $f(x) = y$ avec y donné et x inconnue.

Définition 5.1. Soit $f : X \rightarrow Y$ une application $x \mapsto f(x)$ quelconque.

- On dit que $x \in X$ est un **antécédent** de $y \in Y$ si $f(x) = y$.
- On dit que f est **injective** si tout $y \in Y$ possède **au plus** un antécédent.
- f est **surjective** si tout $y \in Y$ possède **au moins** un antécédent.
- f est **bijective** si elle est injective et surjective, c'est-à-dire si tout $y \in Y$ possède un **unique** antécédent x .
- Si f est une bijection, on note $f^{-1} : Y \rightarrow X$ l'application qui à $y \in Y$ associe son unique antécédent x , c'est-à-dire l'unique solution de l'équation $f(x) = y$. On appelle f^{-1} l'application **réciproque** de f .
- Une application linéaire bijective s'appelle un **isomorphisme**.

Proposition 5.2. Si f est un isomorphisme, alors son application réciproque f^{-1} est aussi une application linéaire.

Démonstration. Supposons que $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est linéaire et bijective. Soient $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$. On a

$$\begin{aligned} f(f^{-1}(\vec{v}_1) + \lambda f^{-1}(\vec{v}_2)) &= f(f^{-1}(\vec{v}_1)) + \lambda f(f^{-1}(\vec{v}_2)) \text{ car } f \text{ est linéaire} \\ &= \vec{v}_1 + \lambda \vec{v}_2, \end{aligned}$$

d'où $f^{-1}(\vec{v}_1) + \lambda f^{-1}(\vec{v}_2) = f^{-1}(\vec{v}_1 + \lambda \vec{v}_2)$ et la linéarité de f^{-1} . □

Les résultats suivants sont importants dans la suite et sont spécifiques aux applications linéaires.

Théorème 5.3.

1. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application linéaire injective, alors $p \leq n$.
2. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ un isomorphisme, alors $p = n$.

Démonstration. 1. L'équation $f(\vec{v}) = \vec{0}$ conduit à un système linéaire homogène (second membre nul) à p équations et n équations. Ce système est compatible puisque $f(\vec{0}) = \vec{0}$.

Si $p > n$, il reste d'après le théorème 2.10 au moins une inconnue non principale après échelonnage. Dans ce cas, l'équation $f(\vec{v}) = \vec{0}$ possède une infinité de solutions, et f ne peut pas être injective.

2. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est un isomorphisme, alors f est injective et $p \leq n$. Il en est de même pour son application réciproque $f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^p$, d'où $n \leq p$ et $n = p$ finalement. \square

5.2 Matrices inversibles

Définition 5.4. Une matrice carré $A \in M_n(\mathbb{R})$ est dite inversible si l'application linéaire associée $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $X \mapsto AX$ est un isomorphisme, c'est-à-dire si pour tout $Y \in \mathbb{R}^n$, l'équation

$$AX = Y$$

possède une unique solution.

Si A est inversible, on note $A^{-1} = \text{Mat}(f^{-1})$ la matrice de l'application réciproque.

Remarque 5.5. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est un isomorphisme, on a $n = p$ nécessairement d'après le théorème 5.3, et donc la matrice associée est carrée. C'est pourquoi les matrices inversibles sont nécessairement carrées.

Proposition 5.6. Si $A \in M_n(\mathbb{R})$ est inversible, alors on a

1. $AX = Y \Leftrightarrow X = A^{-1}Y$,
2. $AA^{-1} = A^{-1}A = I_n$ matrice identité de taille n .

Démonstration. 1. Par définition l'équation $f(X) = AX = Y$ possède une unique solution $X = f^{-1}(Y)$, c'est-à-dire $X = \text{Mat}(f^{-1})Y = A^{-1}Y$ en notation matricielle.

2. On a donc pour tout $Y \in \mathbb{R}^n$, $AA^{-1}Y = AX = Y$, d'où $AA^{-1} = I_n$. De même, pour tout $X \in \mathbb{R}^n$, $A^{-1}AX = A^{-1}Y = X$, d'où $A^{-1}A = I_n$. On peut aussi dire que cela vient des propriétés générales

$$f \circ f^{-1} = \text{Id} = f^{-1} \circ f \Rightarrow \text{Mat}(f) \times \text{Mat}(f^{-1}) = I_n = \text{Mat}(f^{-1}) \times \text{Mat}(f).$$

\square

Critères d'inversibilité.

D'après la discussion précédente et le théorème 2.10, une matrice A est inversible ssi le système $AX = Y$ associé est de Cramer, c'est-à-dire un système carré dont toutes les inconnues sont principales.

Théorème 5.7. Une matrice carrée $A \in M_n(\mathbb{R})$ est inversible si et seulement si son rang vaut n .

Le rang se calcule en échelonnant A et en comptant les pivots. Il est donc relativement rapide de savoir si A est inversible. Il est plus long de calculer son inverse, car il faut alors résoudre le système $AX = Y$, et exprimer le résultat sous la forme $X = A^{-1}Y$.

5.3 Exemple et méthode de calcul

Matrice inversible et inverse. On considère la matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix}$.

1. A est-elle inversible ?
 2. Si oui, calculer son inverse.
- Pour 1, il suffit de calculer le rang de A . On a

$$\text{rang}(A) = \text{rang} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix} = \text{rang} \begin{pmatrix} \boxed{1} & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \text{rang} \begin{pmatrix} \boxed{1} & 1 & 0 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & \boxed{1} \end{pmatrix} = 3,$$

et donc $A \in M_3(\mathbb{R})$ est inversible.

- Pour 2, il faut résoudre le système $AX = Y$. On a

$$\begin{aligned} AX = Y &\Leftrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 + x_2 = y_1 \\ x_1 + 2x_2 + x_3 = y_2 \\ 2x_1 + 3x_2 + 2x_3 = y_3 \end{cases} \\ &\Leftrightarrow \begin{cases} \boxed{x_1} + x_2 = y_1 \\ x_2 + x_3 = y_2 - y_1 \\ x_2 + 2x_3 = y_3 - 2y_1 \end{cases} \Leftrightarrow \begin{cases} \boxed{x_1} + x_2 = y_1 \\ \boxed{x_2} + x_3 = y_2 - y_1 \\ \boxed{x_3} = y_3 - y_2 - y_1 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = y_1 - x_2 = y_1 - 2y_2 + y_3 \\ x_2 = y_2 - y_1 - x_3 = 2y_2 - y_3 \\ x_3 = -y_1 - y_2 + y_3 \end{cases} \end{aligned}$$

On retrouve donc que A est inversible, puisque le système $AX = Y$ possède une unique solution, et de plus on a $X = A^{-1}Y$ avec

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

Conseil. À la fin d'un calcul d'inverse, il est bon de faire quelques vérifications, par exemple en calculant quelques termes du produit $A^{-1}A = I_3$. Les erreurs ont tendance à se propager vers x_1 , c'est-à-dire en première ligne de A^{-1} .

Méthode d'inversion avec matrice augmentée. On peut inverser une matrice A sans poser explicitement le système $AX = Y$, en restant au niveau des codages par matrices augmentées.

Si A est de taille n , on part de la matrice augmentée $(A | I_n)$. On l'échelonne et résout en lignes par la méthode du pivot. On arrive à la forme $(I_n | A^{-1})$ à la fin. Magique non?!

Exemple avec la matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix}$ précédente. On écrit

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 2 & 3 & 2 & 0 & 0 & 1 \end{array} \right) &\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & -2 & 0 & 1 \end{array} \right) &\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right) \\ &\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right) &\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right) \end{aligned}$$

On retrouve bien A^{-1} dans la partie droite. Ce n'est pas si étonnant car l'on fait exactement les mêmes opérations que dans la méthode précédente, mais sans écrire les x_i et les y_i . En fait on est en train de faire la résolution simultanée des 3 équations

$$AX = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

d'où l'apparition à la fin du calcul de $A^{-1}(e_1)$, $A^{-1}(e_2)$ et $A^{-1}(e_3)$ dans les 3 colonnes de droite, c'est-à-dire de A^{-1} elle-même en fait.

5.4 Le cas des matrices 2×2

Il n'y a pas de formule simple de l'inverse d'une matrice carrée en dimension $n \geq 4$.¹ Ce n'est pas le cas pour les matrices 2×2 .

Théorème 5.8. Une matrice $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M_2(\mathbb{R})$ est inversible si et seulement si son **déterminant** $\det A = ad - bc \neq 0$, auquel cas on a :

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

1. Heureusement, les programmes comme Python accomplissent très efficacement ces calculs!

Démonstration. - Soient $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. On a $AX = Y$ ssi

$$\begin{cases} ax_1 + cx_2 = y_1 \\ bx_1 + dx_2 = y_2 \end{cases} \Rightarrow \begin{cases} (ad - bc)x_1 = dy_1 - cy_2 & (dL_1 - cL_2) \\ (ad - bc)x_2 = -by_1 + ay_2 & (aL_2 - bL_1) \end{cases}$$

$$\Rightarrow \begin{cases} x_1 = \frac{1}{\det A}(dy_1 - cy_2) \\ x_2 = \frac{1}{\det A}(-by_1 + ay_2) \end{cases} \quad \text{si } \det A \neq 0.$$

On vérifie facilement que la solution convient et donc A est une matrice inversible, et A^{-1} est la matrice souhaitée.

- Si $\det A = 0$, on observe que $A \begin{pmatrix} d \\ -b \end{pmatrix} = A \begin{pmatrix} -c \\ a \end{pmatrix} = \vec{0} = A(\vec{0})$, et donc A n'est pas injective. □

Exemple. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ est inversible car $\det A = 4 - 6 = -2 \neq 0$, et on a

$$A^{-1} = -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$$

Le déterminant d'une matrice 2×2 a une interprétation géométrique importante.

Théorème 5.9. *Si $A \in M_2(\mathbb{R})$, alors $\det A$ est l'aire orientée du parallélogramme de côtés $A(\vec{e}_1)$ et $A(\vec{e}_2)$.*

Ainsi, $\det A$ mesure la façon dont l'application linéaire associée dilate ou contracte les aires. Par aire « orientée » on entend qu'elle est comptée positivement si $A(\vec{e}_2)$ s'obtient en tournant $A(\vec{e}_1)$ dans le sens trigonométrique, négativement sinon. Le déterminant est nul lorsque le parallélogramme $(A(\vec{e}_1), A(\vec{e}_2))$ est plat, c'est-à-dire lorsque $A(\vec{e}_1)$ et $A(\vec{e}_2)$ sont colinéaires.

Démonstration. On écrit les décompositions polaires de $A(e_1)$ et $A(e_2)$:

$$A(e_1) = \|A(e_1)\|(\cos \theta_1, \sin \theta_1) \quad \text{et} \quad A(e_2) = \|A(e_2)\|(\cos \theta_2, \sin \theta_2).$$

D'où

$$\begin{aligned} \det A &= \det \begin{pmatrix} \|A(e_1)\| \cos \theta_1 & \|A(e_2)\| \cos \theta_2 \\ \|A(e_1)\| \sin \theta_1 & \|A(e_2)\| \sin \theta_2 \end{pmatrix} \\ &= \|A(e_1)\| \|A(e_2)\| (\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2) \\ &= \|A(e_1)\| \|A(e_2)\| \sin(\theta_2 - \theta_1), \end{aligned}$$

en calculant par exemple la partie imaginaire de $e^{i(\theta_2 - \theta_1)} = e^{i\theta_2} e^{-i\theta_1}$.

□

6 Sous-espaces vectoriels, bases et dimension

6.1 Les sous-espaces vectoriels de \mathbb{R}^n

Il y a dans l'espace vectoriel \mathbb{R}^n des parties jouant un grand rôle : les sous-espaces vectoriels. Cette notion généralise celle de droite et de plan vectoriel.

Définition 6.1. On dit qu'une partie E de \mathbb{R}^n est un **sous-espace vectoriel** (seu en abrégé), si

- i) $\vec{0} \in E$,
- ii) pour tout $\vec{v} \in E$ et $\lambda \in \mathbb{R}$, on a $\lambda \cdot \vec{v} \in E$,
- iii) pour tout \vec{u} et $\vec{v} \in E$, on a $\vec{u} + \vec{v} \in E$.

Les droites et plans vectoriels de \mathbb{R}^n sont des seu de \mathbb{R}^n . Ils sont bien stables par dilatation et somme. Dans un seu général, on dispose des opérations vectorielles minimales pour faire les calculs. On peut par exemple parler d'une application linéaire f entre deux seu E et F . Ce sont les applications $f : E \rightarrow F$ telles que $f(\vec{u} + \lambda \vec{v}) = f(\vec{u}) + \lambda f(\vec{v})$.

On rencontre des seu dans deux situations typiques : les seu définis par des systèmes linéaires homogènes, et ceux définis par combinaisons linéaires d'une collection donnée de vecteurs. Ces deux situations sont naturellement associées aux applications linéaires, c'est-à-dire aussi aux matrices.

Définition 6.2. Soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application linéaire de matrice A .

1. Le **noyau** de f est l'ensemble, noté $\ker f$, des vecteurs $\vec{v} \in \mathbb{R}^p$ tels que $f(\vec{v}) = \vec{0}$.
Concrètement $\ker f$ est l'ensemble des solutions du système homogène $AX = 0$.
2. **L'image** de f est l'ensemble des \vec{v} de \mathbb{R}^n qui s'écrivent $f(\vec{u})$ avec $\vec{u} \in \mathbb{R}^p$.
Concrètement $\text{Im } f$ est l'ensemble des combinaisons linéaires des colonnes de A .

Pour le 2, on sait d'après le paragraphe 3.2 que si $\vec{u} = (x_1, x_2, \dots, x_p)$ alors

$$f(\vec{u}) = x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) + \dots + x_p f(\vec{e}_p)$$

est une combinaison linéaire des colonnes de $A = \text{Mat}(f)$. On dit aussi que $\text{Im } f$ est l'espace **engendré** par ces vecteurs. De manière équivalente, on utilise aussi la notation suivante.

Définition 6.3. Si $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ est une famille de vecteurs de \mathbb{R}^n , on note

$$\text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$$

l'ensemble de toutes les combinaisons linéaires de ces vecteurs, on dit aussi l'espace **engendré** par ces vecteurs.

Proposition 6.4. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est une application linéaire, alors $\ker f$ est un sev de \mathbb{R}^p et $\text{Im } f$ est un sev de \mathbb{R}^n .

Démonstration. 1. On a bien $\vec{0} \in \ker f$ puisque $f(\vec{0}) = \vec{0}$ si f est linéaire. Si $\vec{u}, \vec{v} \in \ker f$ et $\lambda \in \mathbb{R}$, on a par linéarité de f , $f(\vec{u} + \lambda\vec{v}) = f(\vec{u}) + \lambda f(\vec{v}) = \vec{0}$ et donc $\vec{u} + \lambda\vec{v} \in \ker f$.

2. On a $\vec{0} = f(\vec{0})$ et donc $\vec{0} \in \text{Im } f$. De plus si $\vec{v}_1 = f(\vec{u}_1)$ et $\vec{v}_2 = f(\vec{u}_2)$ sont deux vecteurs de $\text{Im } f$, alors $\vec{v}_1 + \lambda\vec{v}_2 = f(\vec{u}_1 + \lambda\vec{u}_2)$ appartient aussi à $\text{Im } f$. \square

6.2 Bases et coordonnées

On veut associer à tout sev E de \mathbb{R}^n des systèmes de coordonnées. Cela est possible lorsqu'on dispose d'une **base** de E . Cette notion généralise celle de repère de \mathbb{R}^2 ou \mathbb{R}^3 , que vous avez sûrement déjà utilisée en physique notamment. Il s'agit ici de représenter les vecteurs de E de manière optimale, en le déterminant par un nombre minimal de données.

Définition 6.5. Soit E un sev de \mathbb{R}^n .

- On dit qu'une famille de vecteurs $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ de E est une base de E si et seulement si tout vecteur \vec{v} de E s'écrit **de manière unique** comme combinaison linéaire des \vec{v}_i , c'est-à-dire qu'il existe un unique choix de réels $\lambda_1, \lambda_2, \dots, \lambda_p$ tels que

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_p \vec{v}_p.$$

- Les nombres $\lambda_1, \lambda_2, \dots, \lambda_p$ s'appellent les **coordonnées** de \vec{v} dans la base B . On notera cela

$$\vec{v} = (\lambda_1, \lambda_2, \dots, \lambda_p)_B,$$

avec le B en indice pour préciser la base de travail.

Ainsi, à chaque base d'un sev E est associé un système de coordonnées, une carte de E en quelque sorte.

Composantes et coordonnées.

La base canonique de \mathbb{R}^n , $B_{can} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ est bien une base de \mathbb{R}^n en ce sens puisque tout $\vec{v} = (x_1, x_2, \dots, x_n)$ s'écrit

$$\vec{v} = (x_1, x_2, \dots, x_n) = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n$$

de manière unique. Dans **ce cas**, les coordonnées de \vec{v} dans B_{can} sont simplement ses composantes x_1, x_2, \dots, x_n .

Attention, les coordonnées d'un vecteur donné **changent** lorsque l'on change de base (de repère). Il ne faut pas confondre les **composantes** x_i d'un vecteur $\vec{v} = (x_1, x_2, \dots, x_n)$ de \mathbb{R}^n , et ses **coordonnées** dans une base en général. Les deux notions sont différentes si on travaille dans une autre base que la base canonique.

Par exemple, considérons dans \mathbb{R}^2 les deux vecteurs $\vec{v}_1 = (1, 1)$ et $\vec{v}_2 = (1, -1)$. Ces vecteurs forment une base de \mathbb{R}^2 .

En effet, soit $\vec{v} = (x, y)$ un vecteur quelconque de \mathbb{R}^2 . On a

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \Leftrightarrow \begin{cases} \lambda_1 + \lambda_2 = x \\ \lambda_1 - \lambda_2 = y \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = \frac{x+y}{2} \\ \lambda_2 = \frac{x-y}{2} \end{cases}$$

Ainsi, tout vecteur s'écrit de manière unique comme combinaison linéaire de \vec{v}_1 et \vec{v}_2 , et $B = (\vec{v}_1, \vec{v}_2)$ est une base de \mathbb{R}^2 . De plus les coordonnées de $\vec{v} = (x, y)$ dans la base B sont les nombres

$$\lambda_1 = \frac{x+y}{2} \quad \text{et} \quad \lambda_2 = \frac{x-y}{2}.$$

Autrement dit, avec les notations de la proposition 6.5, on a

$$\vec{v} = (x, y) = \frac{x+y}{2} \vec{v}_1 + \frac{x-y}{2} \vec{v}_2 = \left(\frac{x+y}{2}, \frac{x-y}{2} \right)_B.$$

Par exemple, $\vec{v}_1 = (1, 1)$ a pour coordonnées $(1, 0)$ dans la base B et $\vec{v}_2 = (1, -1) = (0, 1)_B$.

Exemple des droites et plans.

- Tout vecteur \vec{v} non nul est une base de la droite $D = \text{Vect}(\vec{v})$ engendrée par \vec{v} . Par exemple le vecteur $\vec{v} = (1, 2, 3)$ a pour coordonnée unique $\vec{v} = (1)_B$, et $(-2, -4, -6) = (-2)_B$ pour $B = \{\vec{v}\}$.

- De la même façon, tout couple de vecteurs **non colinéaires** \vec{v}_1, \vec{v}_2 est une base du plan $P = \text{Vect}(\vec{v}_1, \vec{v}_2)$ engendré par ces vecteurs.

Démonstration. Montrons l'*unicité* de la décomposition d'un vecteur \vec{v} de P à l'aide de \vec{v}_1 et \vec{v}_2 . Si

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \lambda'_1 \vec{v}_1 + \lambda'_2 \vec{v}_2$$

alors $(\lambda_1 - \lambda'_1) \vec{v}_1 = (\lambda'_2 - \lambda_2) \vec{v}_2$ et donc $\lambda_1 = \lambda'_1$ et $\lambda_2 = \lambda'_2$ si aucun des vecteurs n'est multiple de l'autre. \square

Par exemple, si $\vec{v}_1 = (1, 1, 1)$ et $\vec{v}_2 = (1, 2, 3)$ les *coordonnées* de

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = (\lambda_1 + \lambda_2, \lambda_1 + 2\lambda_2, \lambda_1 + 3\lambda_2)$$

dans la base $B = (\vec{v}_1, \vec{v}_2)$ du plan $P = \text{Vect}(\vec{v}_1, \vec{v}_2)$ sont bien λ_1 et λ_2 , et pas du tout $\lambda_1 + \lambda_2$, $\lambda_1 + 2\lambda_2$ et $\lambda_1 + 3\lambda_2$. Vu dans P , le vecteur $\vec{v} = (\lambda_1, \lambda_2)_B$ est déterminé par deux coordonnées et non trois.

6.3 Notion d'indépendance linéaire

Dans une base $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ d'un sev E de \mathbb{R}^n , **aucun** des vecteurs \vec{v}_i ne peut s'écrire comme combinaison linéaire **des autres**. Sinon, cela contredirait l'unicité de la décomposition d'un vecteur de E dans B . On dit que les vecteurs de B sont **linéairement indépendants**. Il n'y a aucune relation linéaire entre eux, autre que

$$0\vec{v}_1 + 0\vec{v}_2 + \dots + 0\vec{v}_p = \vec{0} !$$

Définition 6.6. On dit qu'une famille de vecteurs $\mathcal{F} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ est **libre** ou **linéairement indépendante** si l'équation

$$\lambda_1 \vec{v}_1 + \dots + \lambda_p \vec{v}_p = \vec{0} \text{ implique } \lambda_1 = \dots = \lambda_p = 0.$$

Dans le cas contraire, on dit que la famille \mathcal{F} est **liée**, ou que les vecteurs \vec{v}_i sont linéairement dépendants.

Quand une famille génératrice est liée, elle n'est pas optimale au sens qu'un des vecteurs peut-être enlevé de la famille sans qu'elle cesse d'être génératrice. En effet si par exemple \vec{v}_p est combinaison linéaire de $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{p-1}$ alors

$$\text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p) = \text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{p-1}).$$

C'est la conjonction des deux propriétés, libre et génératrice, qui donne des bases.

Proposition 6.7. Une famille de vecteurs $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ de E est une base de E ssi B est libre et génératrice de E .

Démonstration. En fait B libre équivaut à ce que tout vecteur de \vec{v} de E se décompose **au plus** d'une manière comme combinaison linéaire des \vec{v}_i . En effet, si

$$\vec{v} = \lambda_1 \vec{v}_1 + \dots + \lambda_p \vec{v}_p = \lambda'_1 \vec{v}_1 + \dots + \lambda'_p \vec{v}_p$$

alors

$$(\lambda_1 - \lambda'_1) \vec{v}_1 + \dots + (\lambda_p - \lambda'_p) \vec{v}_p = \vec{0}$$

d'où $\lambda_1 - \lambda'_1 = 0 = \dots = \lambda_p - \lambda'_p$ si B est libre.

Inversement, si on a unicité de l'écriture de $\vec{0}$ à l'aide B , alors l'équation

$$\lambda_1 \vec{v}_1 + \dots + \lambda_p \vec{v}_p = \vec{0} = 0\vec{v}_1 + \dots + 0\vec{v}_p$$

implique $\lambda_1 = \dots = \lambda_p = 0$ et la famille est libre. □

6.4 Dimension d'un sous-espace vectoriel

On admet pour l'instant que tout sev non nul de \mathbb{R}^n possède des bases finies. On verra concrètement comment en construire plus loin. Le résultat suivant sur le nombre de vecteurs dans une base est fondamental en algèbre linéaire.

Théorème 6.8. *Soit E un sev non nul de \mathbb{R}^n . Alors toutes les bases de E ont le même nombre de vecteurs. Ce nombre entier s'appelle la **dimension** de E et se note $\dim E$. On a de plus ici $\dim E \leq n$.*

Par convention utile, on définit $\dim\{\vec{0}\} = 0$. (En fait le vecteur $\vec{0}$ ne peut faire partie d'aucune base puisque $\vec{0} = 2\vec{0}$ et il ne peut y avoir unicité des coordonnées.)

Le nombre d'éléments d'un ensemble fini B s'appelle son **cardinal**, et se note $\text{Card}(B)$. On a donc ici

$$\text{Card}(B) = \dim E \text{ pour toute base } B \text{ de } E.$$

Démonstration. Pour la preuve, nous allons utiliser des notions utiles dans la suite pour les changements de base.

Définition 6.9. Si $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ est une famille de vecteurs d'un sev E de \mathbb{R}^n , alors on peut considérer l'application « **combinaison linéaire suivant B** »

$$\begin{aligned} L_B : \mathbb{R}^p &\rightarrow E \\ X = (x_1, x_2, \dots, x_p) &\mapsto x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p. \end{aligned}$$

L_B est un isomorphisme lorsque B est une base de E .

Dans ce cas, l'application réciproque $C_B = L_B^{-1} : E \rightarrow \mathbb{R}^p$ est l'application « **coordonnées dans la base B** » puisque $C_B(\vec{v})$ donne les nombres x_i tels que

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p = \vec{v}.$$

- Si B et B' sont deux bases de E de cardinaux respectifs p et k , alors l'application composée $C_B L_{B'}$ est un isomorphisme entre \mathbb{R}^k et \mathbb{R}^p . On a donc $p = k$ d'après le théorème 5.3.

Remarque 6.10. L'application $P_{B,B'} = C_B L_{B'} : \mathbb{R}^p \rightarrow \mathbb{R}^p$ donne les coordonnées d'un vecteur de E dans B les connaissant dans B' . En effet, si $X = C_B(\vec{v})$ et $X' = C_{B'}(\vec{v})$ alors

$$P_{B,B'} X' = C_B L_{B'} C_{B'}(\vec{v}) = C_B(\vec{v}) = X.$$

La matrice de $P_{B,B'}$ s'appelle la **matrice de changement de base** de B à B' .

- Si E est un sev de dimension d de \mathbb{R}^n et B est une base de E , alors L_B induit une **injection** de \mathbb{R}^d dans \mathbb{R}^n . On a donc $d \leq n$ par le théorème 5.3. \square

6.5 Calculs de dimensions et de bases

Le cas de \mathbb{R}^n .

On a $\dim \mathbb{R}^n = n$ puisque la base canonique $B_{can} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ à n vecteurs.
Toutes les bases de \mathbb{R}^n ont donc n vecteurs.

Soient $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ n vecteurs donnés de \mathbb{R}^n . La famille $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ est une base de \mathbb{R}^n ssi l'application L_B de combinaison linéaire suivant B est un isomorphisme de \mathbb{R}^n dans \mathbb{R}^n . Par définition la matrice de L_B est constituée des composantes des \vec{v}_i écrits en colonnes. D'après le théorème 5.7 on a le résultat pratique suivant.

Proposition 6.11. *Une famille de n vecteurs de \mathbb{R}^n est une base de \mathbb{R}^n ssi la matrice des colonnes des composantes de ces vecteurs est de rang n .*

Il suffit donc d'échelonner cette matrice pour le savoir. Par contre, pour calculer les nouvelles coordonnées d'un vecteur \vec{v} dans cette base B , il faut **résoudre** le système de Cramer $L_B(X) = \vec{v}$ dont la solution est bien $X = L_B^{-1}(\vec{v}) = C_B(\vec{v})$.

Droites et plans vectoriels.

D'après leur définition et §6.2, les droites vectorielles sont les sous-espaces vectoriels de dimension 1, tandis que les plans vectoriels sont les sous-espaces vectoriels de dimension 2.

Intuitivement, on peut dire que la dimension d'un sous-espace vectoriel E de \mathbb{R}^n est le nombre de « paramètres libres » dont dépend la position d'un vecteur de E , à ne pas confondre avec le nombre n éventuellement beaucoup plus grand de composantes de ce vecteur. Par exemple, la détermination d'un vecteur dans un plan P dépend de deux nombres, ses coordonnées dans une base de P , que P soit plongé dans \mathbb{R}^3 , \mathbb{R}^4 ou \mathbb{R}^{100} .

Base et dimension d'un sev engendré par une famille de vecteurs.

Théorème 6.12. *Soit $E = \text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ un sev de \mathbb{R}^n engendré par p vecteurs.*

On considère le système homogène (S) (à n équations et p inconnues)

$$x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p = \vec{0}.$$

Concrètement, si A est la matrice constituée des vecteurs colonnes \vec{v}_i , alors E est l'image de A , et (S) s'écrit $AX = 0$.

On note $I = (i_1, i_2, \dots, i_r)$ les indices des inconnues principales d'un système échelonné équivalent et r son rang.

Alors la famille B des r vecteurs $\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_r}$ correspondant aux inconnues **principales** est une base de E . En particulier, on a

$$\dim E = \text{rang}(S) = \text{rang}(A)$$

Démonstration. On montre que la famille B est libre et génératrice de E .

- B est-elle génératrice ?

On pose $F = \text{Vect}(B)$. On a $B \subset \mathcal{F} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ d'où $F = \text{Vect}(B) \subset E = \text{Vect}(\mathcal{F})$. Nous allons voir que les vecteurs « manquants » \vec{v}_j correspondant aux inconnues *non principales* sont aussi dans F , d'où l'on tire l'inclusion opposée.

Si x_j est une inconnue non principale, on sait par le pivot de Gauss-Jordan qu'il existe une (unique) solution de (S) avec $x_j = 1$ et les autres inconnues non principales nulles. Cela s'écrit

$$\sum_{i \in I} x_i \vec{v}_i + \vec{v}_j = \vec{0} \iff \vec{v}_j = - \sum_{i \in I} x_i \vec{v}_i \in F = \text{Vect}(B).$$

- B est-elle libre ?

Supposons que l'on ait une relation de dépendance linéaire entre les vecteurs de B , c'est-à-dire que

$$\sum_{i \in I} x_i \vec{v}_i = \vec{0}.$$

Alors cela s'interprète comme une solution du système (S) dont les inconnues non principales sont toutes nulles. D'après le pivot de Gauss-Jordan, la donnée des inconnues non principales détermine la solution. Comme la solution nulle convient, c'est la seule possible. On a donc $x_i = 0$ et la famille B est libre. \square

On voit donc que le rang du système homogène (S) associé à la famille de vecteurs s'interprète en fait comme une dimension et est une donnée purement géométrique de la famille.

Définition 6.13. On appelle **rang d'une famille de vecteurs** $\mathcal{F} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$, la dimension de l'espace qu'elle engendre. D'après le résultat précédent, on a :

$$\text{rang}(\mathcal{F}) = \dim \text{Vect}(\mathcal{F}) = \text{rang}(S)$$

où (S) est le système $x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p = \vec{0}$.

Dimension d'un sev présenté par un système homogène.

Théorème 6.14. Soit $E = \text{Sol}(S)$ l'espace des solutions d'un système homogène. Alors la dimension de E est le nombre d'inconnues **non principales** d'un système échelonné équivalent à (S) . Une base de E est donnée dans la preuve.

Il suffit donc d'échelonner (S) , **sans le résoudre**, pour connaître la dimension de $E = \text{Sol}(S)$. Par exemple la dimension de $E = \{\vec{v} = (x, y, z, t, u) \in \mathbb{R}^5, x + 2y - 3z + 4t = 0\}$ est 4 (x inconnue principale, y, z, t, u non principales).

Attention, de ne pas confondre ce résultat avec le précédent sur la dimension d'un sev engendré par une famille de vecteurs. La dimension est donnée ici par le nombre **d'inconnues non principales**. Ce sont en effet elles qui paramètrent les solutions, et non les inconnues principales.

Démonstration du théorème 6.14. On note r le rang d'un système échelonné équivalent, $I = (i_1, i_2, \dots, i_r)$ ses inconnues principales et $J = {}^c I = (j_1, j_2, \dots, j_{p-r})$ ses inconnues non principales. Si $r = p$, il n'y a pas d'inconnues non principales, et E est réduit au vecteur nul, donc de dimension 0. Sinon, nous allons construire une base de E .

Si $r < p$, une base de E est donnée par $p - r$ vecteurs

$$B = (\vec{v}_{j_1}, \vec{v}_{j_2}, \dots, \vec{v}_{j_r}),$$

un par inconnue non principale, où chaque \vec{v}_{j_k} est l'unique solution de (S) correspondant à la donnée des inconnues non principales

$$x_{j_k} = 1 \quad \text{et} \quad x_{j_l} = 0 \quad \text{pour} \quad l \neq k.$$

Par linéarité des équations du système, le vecteur

$$\vec{v} = \sum_{j \in J} x_j \vec{v}_j$$

est une solution de (S) dont les inconnues non principales sont les x_j . D'après le théorème 2.8 de Gauss-Jordan, c'est l'unique solution correspondant à ces valeurs. On a donc que tout vecteur de $E = \text{Sol}(S)$ est de cette forme, et ainsi que la famille $B = (\vec{v}_j, j \in J)$ est génératrice de E .

Par ailleurs, la famille B est libre car les composantes « non principales » des \vec{v}_j décrivent par définition la base canonique de \mathbb{R}^{p-r} . □

Exemple. On traite un exemple pour fixer les idées. Soit

$$E = \{\vec{v} = (x, y, z, t) \in \mathbb{R}^4 \mid x + y - z + t = 0, 2x + y + 2z - t = 0\}.$$

On échelonne et résout le système. On a $\vec{v} = (x, y, z, t) \in E \Leftrightarrow$

$$\begin{cases} x + y - z + t = 0 \\ 2x + y + 2z - t = 0 \end{cases} \Leftrightarrow \begin{cases} \boxed{x} + y - z + t = 0 \\ \boxed{-y} + 4z - 3t = 0 \end{cases} \\ \Leftrightarrow \begin{cases} x = -y + z - t = -4z + 3t + z - t = -3z + 2t \\ y = 4z - 3t \end{cases}$$

On a donc que $\vec{v} \in E$ est de la forme

$$\vec{v} = (-3z + 2t, 4z - 3t, z, t) = z(-3, 4, \boxed{1, 0}) + t(2, -3, \boxed{0, 1}).$$

Autrement dit E est le plan engendré par les deux vecteurs

$$\vec{v}_1 = (-3, 4, 1, 0) \quad \text{et} \quad \vec{v}_2 = (2, -3, 0, 1),$$

où \vec{v}_1 est la solution de (S) correspondant aux inconnues non principales $z = 1$ et $t = 0$, tandis que \vec{v}_2 correspond à $z = 0$ et $t = 1$. Notez que ces vecteurs sont non colinéaires, en particulier car les deux dernières composantes encadrées (des inconnues non principales) sont celles de la base canonique de \mathbb{R}^2 .

Une conséquence intéressante du théorème 6.14 précédent est la suivante.

Corollaire 6.15. *Un système (S) de n équations homogènes à p inconnues définit un sous-espace vectoriel $E = \text{Sol}(S)$ de \mathbb{R}^p de dimension $\geq p - n$.*

Démonstration. En effet, d'après le théorème 6.14, on a $\dim E =$ nombre d'inconnues non-principales. Or, il y a au plus un pivot par équation, et donc au plus n inconnues principales. Il reste donc au moins $p - n$ inconnues non-principales. \square

En particulier, dans \mathbb{R}^n , il faut (au moins) $n - 1$ équations pour définir une droite vectorielle, et $n - 2$ pour définir un plan, d'où l'intérêt de les définir plutôt par vecteurs directeurs en général!

Extraction et complétion de bases.

Pour conclure, nous pouvons énoncer deux techniques générales de construction de bases. L'idée est d'essayer de faire grossir une famille libre, en ajoutant des directions « manquantes », jusqu'à ce qu'elle devienne génératrice. On peut aussi faire maigrir une famille génératrice, en enlevant les vecteurs qui s'expriment à l'aide des autres, jusque ce que la famille devienne libre.

Théorème 6.16 (Thm d'extraction de base). *Soit \mathcal{F} une famille finie et **génératrice** d'un espace vectoriel non nul E . Alors on peut **extraire** de \mathcal{F} une sous-famille $\mathcal{F}' \subset \mathcal{F}$ qui est une base de E .*

Dans l'autre direction, en partant d'une famille libre, on a le résultat.

Théorème 6.17 (Théorème de la base incomplète).

Soit $\mathcal{F} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q)$ une famille **libre** d'un espace vectoriel E muni d'une base $B = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$.

Alors, on peut **compléter la famille libre** \mathcal{F} à l'aide de certains vecteurs $\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_r}$ de la base B pour obtenir une base

$$B' = \mathcal{F} \cup (\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_r}) \text{ de } E.$$

En pratique dans $E = \mathbb{R}^n$, on peut compléter une famille libre donnée à l'aide de certaines directions \vec{e}_i de la base canonique. Par exemple dans \mathbb{R}^3 , on peut compléter deux vecteurs non colinéaires \vec{u} et \vec{v} par un des trois vecteurs $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$ ou $\vec{e}_3 = (0, 0, 1)$. Il suffit de prendre un axe qui n'est pas inclus dans le plan $P = \text{Vect}(\vec{u}, \vec{v})$. En effet, toutes les bases de \mathbb{R}^3 ont 3 vecteurs et que donc il n'y a qu'un des \vec{e}_i à ajouter pour compléter.

Démonstration. Les deux théorèmes sont des conséquences du théorème 6.12. Il énonce bien que l'on peut extraire une base d'une famille génératrice, en précisant même que les vecteurs à garder sont ceux correspondant aux inconnues principales du système homogène $AX = 0$ associé.

Pour compléter une famille libre \mathcal{F} , on peut lui rajouter (derrière) tout B . Elle devient génératrice. Puis on applique le théorème 6.12. Comme la famille \mathcal{F} est libre, ses vecteurs sont associés à des inconnues principales du système associé à $\mathcal{F} \sqcup B$ et on les garde bien dans le procédé. \square

6.6 Conséquences utiles

La notion de dimension est un outil de raisonnement efficace dans les problèmes de familles, de bases et de comparaisons de sev. On énonce quelques propriétés utiles.

Cardinal des familles et dimension.

Proposition 6.18. Soit E un espace vectoriel de dimension n et \mathcal{F} une famille de vecteurs de E . Alors

- a) $\text{Card}(\mathcal{F}) \leq n$ si \mathcal{F} est libre.
- b) $\text{Card}(\mathcal{F}) \geq n$ si \mathcal{F} est génératrice.
- c) Si \mathcal{F} est génératrice et de cardinal n , alors c'est une base de E .
- d) Si \mathcal{F} est libre et de cardinal n , alors c'est une base de E .

Démonstration. a) On peut compléter \mathcal{F} en une base de E . On a donc $d = \text{Card}(\mathcal{F}) \leq \dim E$. On peut aussi utiliser le fait que l'application composée $C_B \circ L_{\mathcal{F}}$ de \mathbb{R}^d dans \mathbb{R}^n est injective. D'où $d \leq n$ par le théorème 5.3 a.

b) Par extraction d'une base de cardinal $n = \dim E$ de la famille génératrice.

c) De nouveau par extraction de base. Si \mathcal{F} est génératrice et de « bon » cardinal $\dim E$, alors il n'a rien à en extraire pour avoir une base !

d) Si \mathcal{F} est une famille libre de cardinal $\dim E$, alors il n'a rien à compléter. C'est déjà une base.

□

Remarques 6.19. • Les énoncés a) et b) s'utilisent souvent en contraposée, c'est-à-dire qu'une famille de cardinal $> \dim E$ est automatiquement liée, et de cardinal $< \dim E$ ne peut être génératrice.

• Les résultats c) et d) permettent de réduire les calculs. Pour montrer qu'une famille \mathcal{F} est une base, il suffit de vérifier que $\text{Card}(\mathcal{F}) = \dim E$ et que \mathcal{F} est libre **ou** génératrice !

Croissance de la dimension.

Proposition 6.20. *Soit F un sev d'un ev E (de dimension finie). Alors on a*

a) $\dim F \leq \dim E$.

b) *Si de plus $\dim F = \dim E$ alors $F = E$.*

Démonstration. a) Tout d'abord, soit $F = \{\vec{0}\}$ et $\dim F = 0$ par convention. Sinon, une base de F est une famille libre de E et la proposition 6.18a s'applique.

b) Si on a de plus $\dim F = \dim E$ alors une base B de F est une famille libre de E de cardinal $\dim E$. C'est donc une base de E d'après la proposition 6.18d. On a donc

$$F = \text{Vect}(B) = E.$$

□

Cette proposition est utile pour montrer que deux espaces E et F sont égaux. Il suffit de vérifier que $\dim E = \dim F$ et **qu'une** des inclusions $E \subset F$ **ou** $F \subset E$ est satisfaite. Dans \mathbb{R}^n , il y a une infinité de droites vectorielles différentes (si $n > 1$), ou de plans différents (si $n > 2$), mais **un seul** sev de dimension maximale n : \mathbb{R}^n lui-même.

7 Retour sur l'équation $AX=Y$

7.1 Généralités, noyau, image

Soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application linéaire de matrice A . On revient sur les propriétés générales de l'équation $f(\vec{u}) = \vec{v}$, c'est-à-dire concrètement du système $AX = Y$.

On rappelle (voir § 6.1 et Prop. 6.2) que l'on peut associer à f (ou A) deux sous espaces vectoriels.

— Le **noyau** de f est le sev de \mathbb{R}^p défini par $\ker f = \{\vec{u} \in \mathbb{R}^p, f(\vec{u}) = \vec{0}\}$.

— **L'image** de f est l'ensemble noté $\text{Im } f$ des $\vec{v} \in \mathbb{R}^n$ qui peuvent s'écrire $f(\vec{u})$.

Concrètement si f a pour matrice A , $\ker f$ est donné par les solutions $X \in \mathbb{R}^p$ du système homogène $AX = 0$ et $\text{Im } f$ est l'espace engendré par les colonnes de A .

Ces deux espaces donnent la structure générale de l'ensemble des solutions de l'équation linéaire $f(\vec{u}) = \vec{v}$ lorsque \vec{v} est donné et \vec{u} inconnu.

Proposition 7.1. Soient $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ et $\vec{v} \in \mathbb{R}^n$ donnés. Alors :

1. l'équation $(E) : f(\vec{u}) = \vec{v}$ possède au moins une solution \vec{u}_0 ssi $\vec{v} \in \text{Im } f$.
2. Si $\vec{v} \in \text{Im } f$ et \vec{u}_0 est une **solution particulière** de (E) , alors toute solution de (E) est de la forme $\vec{u}_0 + \vec{u}$ avec $\vec{u} \in \ker f$. Autrement dit, l'espace des solutions de (E) est **l'espace affine** passant par \vec{u}_0 et de direction vectorielle $\ker f$.

Corollaire 7.2. En particulier, on a que :

1. f est surjective ssi $\text{Im } f = \mathbb{R}^n$ (une évidence vraie en général),
2. f est injective ssi $\ker f = \{\vec{0}\}$ (vrai lorsque f est linéaire).

Démonstration de la Proposition 7.1. 1 est vrai par définition de $\text{Im } f$.

2. Si \vec{u}_0 est une solution particulière de (E) et \vec{w} un vecteur quelconque de \mathbb{R}^n , alors

$$\begin{aligned} f(\vec{w}) = \vec{v} &\Leftrightarrow f(\vec{w}) = f(\vec{u}_0) \Leftrightarrow f(\vec{w} - \vec{u}_0) = \vec{0} \text{ car } f \text{ est linéaire} \\ &\Leftrightarrow \vec{w} - \vec{u}_0 = \vec{v} \in \ker f. \end{aligned}$$

□

Par définition, f ne varie pas dans la direction de son noyau. Dans le cas où f n'est pas injective, elle écrase donc les directions de $\ker f$ (un peu comme une projection), et l'espace de départ \mathbb{R}^p est feuilleté en sous-espaces affines parallèles à $\ker f$ et le long desquels f prend des valeurs constantes.

7.2 Le théorème du rang

Il y a une relation importante et simple entre les dimensions du noyau et de l'image d'une application linéaire. On appelle rang de f la dimension de l'image de f .

Théorème 7.3 (Théorème du rang). *Soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ linéaire, alors on a :*

$$p = \text{rang } f + \dim \ker f.$$

Démonstration. Concrètement d'après la définition 6.2, l'image de f est engendrée par les vecteurs colonnes de $A (= f(\vec{e}_i))$. De plus, d'après le théorème d'extraction 6.16, une base de $\text{Im } f$ est constituée des colonnes correspondantes aux inconnues principales de A . On a donc

$$\text{rang } f = \text{rang } A = \text{nombre d'inconnues principales du système } AX = 0.$$

Par ailleurs, d'après le théorème 6.14,

$$\dim \ker f = \text{nombre d'inconnues non principales du système } AX = 0.$$

On a bien

$$\text{rang } f + \dim \ker f = p = \text{nombre d'inconnues de } AX = 0$$

□

Ainsi, pour une dimension de l'espace de départ p donnée, plus le noyau grossit (en dimension) plus l'image rétrécit. f est en effet constante suivant de plus en plus de directions et prend donc moins de valeurs différentes.

7.3 Exemple

On illustre les résultats précédents en étudiant un exemple.

Soit $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ et $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $X \mapsto AX$ l'application linéaire associée. On résout

l'équation $AX = Y$ par la méthode du pivot. On a $AX = Y \Leftrightarrow$

$$\begin{aligned} \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} \boxed{1} & 4 & 7 & | & y_1 \\ 2 & 5 & 8 & | & y_2 \\ 3 & 6 & 9 & | & y_3 \end{pmatrix} \Leftrightarrow \begin{pmatrix} \boxed{1} & 4 & 7 & | & y_1 \\ 0 & -3 & -6 & | & y_2 - 2y_1 \\ 0 & -6 & -12 & | & y_3 - 3y_1 \end{pmatrix} \\ &\Leftrightarrow \begin{pmatrix} \boxed{1} & 4 & 7 & | & y_1 \\ 0 & \boxed{-3} & -6 & | & y_2 - 2y_1 \\ 0 & 0 & 0 & | & y_1 - 2y_2 + y_3 \end{pmatrix} \end{aligned}$$

À ce niveau, on voit que $Y \in \text{Im } f$ ssi l'équation de compatibilité $y_1 - 2y_2 + y_3 = 0$ est satisfaite. C'est l'équation de l'image de f . En particulier f n'est pas surjective car $\text{Im } f$ est un plan. On a bien $\text{rang } f = 2$ car il y a deux pivots (inconnues principales). Une base de l'image de f est engendrée par les colonnes de A correspondant aux inconnues principales, c'est-à-dire $C_1 = Ae_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ et $C_2 = Ae_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.

On poursuit la résolution du système, lorsqu'il est compatible. On a alors pour $y_1 - 2y_2 + y_3 = 0$, $AX = Y$ ssi

$$\begin{aligned} \left(\begin{array}{ccc|c} \boxed{1} & 4 & 7 & y_1 \\ 0 & \boxed{1} & 2 & 2y_1/3 - y_2/3 \end{array} \right) &\Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 4 & & y_1 - 7x_3 \\ 0 & \boxed{1} & & 2y_1/3 - y_2/3 - 2x_3 \end{array} \right) \\ &\Leftrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 0 & & -5y_1/3 + 4y_2/3 + x_3 \\ 0 & \boxed{1} & & 2y_1/3 - y_2/3 - 2x_3 \end{array} \right) \\ &\Leftrightarrow \begin{cases} x_1 = -5y_1/3 + 4y_2/3 + x_3 \\ x_2 = 2y_1/3 - y_2/3 - 2x_3 \\ x_3 \text{ paramètre réel quelconque.} \end{cases} \end{aligned}$$

Lorsque $Y = 0$, on trouve les vecteurs du noyau de A . Ils sont de la forme $X = x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. $\ker A$ est donc la droite vectorielle engendrée par le vecteur $\vec{v} = (1, -2, 1)$. On note que $V \in \ker A$ se traduit par une relation de dépendance linéaire entre les colonnes de A . On a

$$C_1 - 2C_2 + C_3 = 0.$$

C'est une propriété générale utile.

Proposition 7.4. *Soit $A \in M_{n,p}(\mathbb{R})$. On a*

$$X = (x_1, x_2, \dots, x_p) \in \ker A \Leftrightarrow AX = x_1C_1 + x_2C_2 + \dots + x_pC_p = 0,$$

avec les C_j vecteurs colonnes de A .

On a en effet $C_j = Ae_j$ avec e_j base canonique.

Pour terminer, lorsqu'il est compatible, les solutions de $AX = Y$ sont bien de la forme

$$X = X_0 + x_3V \text{ avec } X_0 = \begin{pmatrix} -5y_1/3 + 4y_2/3 \\ 2y_1/3 - y_2/3 \\ 0 \end{pmatrix} \text{ et } V = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

X_0 est une solution particulière (correspondant à $x_3 = 0$) et V engendre le noyau de f .

7.4 Conséquences du théorème du rang

Les critères suivants permettent d'obtenir des propriétés sur f à partir des dimensions des espaces de départ et d'arrivée et de son rang.

Proposition 7.5. *Soit $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$. Alors :*

- a) *on a toujours $\text{rang } f \leq p$ et n . Avec $\text{rang } f = p$ si et seulement si f est injective et $\text{rang } f = n$ ssi f est surjective.*
- b) *Si f est injective, alors $p \leq n$.*
- c) *Si f est surjective, alors $p \geq n$.*
- d) *Si f est bijective, alors $p = n$.*
- e) *Inversement, si $p = n$, alors $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est bijective ssi f est injective **ou** surjective, c'est-à-dire ssi $\ker f = \{\vec{0}\}$ **ou** $\text{rang } f = n$.*

Démonstration.

- a) On a toujours $\text{rang } f = p - \dim \ker f \leq p$, avec égalité ssi $\ker f = \{\vec{0}\}$. On a aussi $\text{rang } f \leq n$ puisque $\text{rang } f = \dim \text{Im } f$ et $\text{Im } f \subset \mathbb{R}^n$. On a $\text{rang } f = n$ ssi $\text{Im } f = \mathbb{R}^n$ d'après la proposition 6.20.
- b) Si f est injective alors $\text{rang } f = p \leq n$ d'après a). En fait, on sait cela depuis le théorème 5.3 !
- c) Si f est surjective alors $\text{rang } f = n \leq p$ d'après a).
- d) Synthèse de b) et c). Aussi connu depuis le théorème 5.3.
- e) Cas d'égalité du a).

□

8 Espaces supplémentaires, somme directe, projections et symétries

Les projections et les symétries sont des applications linéaires qui se définissent (et sont utiles) dans tous les espaces \mathbb{R}^n , dès que l'on dispose de deux sous-espaces E et F *supplémentaires*.

8.1 Espaces supplémentaires et somme directe

Définition 8.1. On dit que 2 sev E et F de \mathbb{R}^n sont *supplémentaires* si tout vecteur \vec{u} de \mathbb{R}^n se décompose de manière unique en

$$\vec{u} = \vec{v} + \vec{w} \quad \text{avec} \quad \vec{v} \in E \text{ et } \vec{w} \in F.$$

On dit aussi que \mathbb{R}^n est *somme directe* de E et F et on écrit

$$\mathbb{R}^n = E \oplus F.$$

Un critère pratique pour reconnaître cette situation est le suivant.

Théorème 8.2. On a $\mathbb{R}^n = E \oplus F$ si et seulement si

$$\dim E + \dim F = n \text{ et } E \cap F = \{\vec{0}\}.$$

Par exemple, 2 droites vectorielles distinctes sont supplémentaires dans \mathbb{R}^2 . Dans \mathbb{R}^3 , un plan P et une droite D non contenue dans P sont supplémentaires.

Démonstration. • \Leftarrow Soient $B_E = (\vec{e}_i)$ et $B_F = (\vec{f}_j)$ des bases de E et F . On considère la famille $B = B_E \sqcup B_F$. Montrons que B reste libre si $E \cap F = \{\vec{0}\}$.

En effet, si $\sum_i \lambda_i \vec{e}_i + \sum_j \mu_j \vec{f}_j = \vec{0}$ alors $\sum_i \lambda_i \vec{e}_i = -\sum_j \mu_j \vec{f}_j \in E \cap F = \{\vec{0}\}$, d'où $\sum_i \lambda_i \vec{e}_i = \sum_j \mu_j \vec{f}_j = \vec{0}$ et finalement $\lambda_i = \mu_j = 0$ car B_E et B_F sont libres.

Si de plus $\dim E + \dim F = n$ alors $\text{Card}(B) = \text{Card}(B_E) + \text{Card}(B_F) = n$ et B est une famille libre de cardinal n de \mathbb{R}^n . C'est donc une base de \mathbb{R}^n . On a alors que tout vecteur de \mathbb{R}^n se décompose de manière unique dans la base B

$$\vec{u} = \sum_i \lambda_i \vec{e}_i + \sum_j \mu_j \vec{f}_j = \vec{v} + \vec{w},$$

avec $\vec{v} = \sum_i \lambda_i \vec{e}_i \in E$ et $\vec{w} = \sum_j \mu_j \vec{f}_j \in F$.

• \Rightarrow Si $\mathbb{R}^n = E \oplus F$, alors $E \cap F = \{\vec{0}\}$. En effet, si $\vec{u} \in E \cap F$, alors on peut décomposer $\vec{u} = \vec{u} + \vec{0} = \vec{0} + \vec{u}$ suivant $E \oplus F$. Par unicité de la décomposition, on a bien $\vec{u} = \vec{0}$. D'après le point précédent, la famille $B = B_E \sqcup B_F$ reste donc libre. Comme elle doit être génératrice de \mathbb{R}^n si $\mathbb{R}^n = E \oplus F$, c'est donc une base de \mathbb{R}^n et on doit avoir $n = \text{Card}(B) = \dim E + \dim F$. \square

8.2 Projections et symétries

Définition 8.3. Soient E et F deux sev supplémentaires dans \mathbb{R}^n .

1. La *projection* $p_{E/F}$ de \mathbb{R}^n sur E parallèlement à F (ou le long de F) est l'application définie par

$$\begin{aligned} p_{E/F} : \mathbb{R}^n = E \oplus F &\longrightarrow \mathbb{R}^n \\ \vec{u} = \vec{v} + \vec{w} &\longmapsto \vec{v} = p_{E/F}(\vec{u}) \end{aligned}$$

2. La *symétrie* $s_{E/F}$ de \mathbb{R}^n par rapport à E et de direction F est l'application définie par

$$\begin{aligned} s_{E/F} : \mathbb{R}^n = E \oplus F &\longrightarrow \mathbb{R}^n \\ \vec{u} = \vec{v} + \vec{w} &\longmapsto \vec{v} - \vec{w} = s_{E/F}(\vec{u}) \end{aligned}$$

On vérifie facilement (exercice) que $p_{E/F}$ et $s_{E/F}$ sont des applications linéaires.

La projection $p_{E/F}$ est en fait *l'unique* application linéaire qui laisse invariant les vecteurs de E : $p_{E/F}(\vec{v}) = \vec{v}$ si $\vec{v} \in E$, et qui envoie F sur $\vec{0}$.

De la même façon, la symétrie $s_{E/F}$ est *l'unique* application linéaire qui laisse invariant les vecteurs de E : $s_{E/F}(\vec{v}) = \vec{v}$ si $\vec{v} \in E$, et transforme les vecteurs de F en leur opposé : $s_{E/F}(\vec{w}) = -\vec{w}$ si $\vec{w} \in F$.

Les projections $p_{E/F}$ et $p_{F/E}$ donnent les 2 *composantes* d'un vecteur donné suivant E et F . Elles permettent de reconstruire le vecteur \vec{u} d'origine par la formule $\vec{u} = \vec{v} + \vec{w} = p_{E/F}(\vec{u}) + p_{F/E}(\vec{u})$. On écrit

$$\text{Id} = p_{E/F} + p_{F/E} \tag{5}$$

où Id désigne **l'application identité**, qui ne déplace aucun vecteur : $\text{Id}(\vec{u}) = \vec{u}$ pour tout $\vec{u} \in \mathbb{R}^n$.

On remarque également que les symétries s'expriment facilement à l'aide des projections. En effet, on a par définition et (5)

$$s_{E/F} = p_{E/F} - p_{F/E} = 2p_{E/F} - \text{Id}. \tag{6}$$

8.3 Exemples

Dans le plan \mathbb{R}^2 .

Soient D_1 et D_2 deux droites vectorielles distinctes de \mathbb{R}^2 , **pas nécessairement orthogonales**.

On a alors $\mathbb{R}^2 = D_1 \oplus D_2$. Si D_1 est engendrée par \vec{u}_1 , et D_2 a pour équation $ax + by = 0$, on peut écrire explicitement la projection p_{D_1/D_2} sur D_1 le long de D_2 .

Si $\vec{u} = (x, y)$ et $\ell(\vec{u}) = ax + by$, on a

$$p_{D_1/D_2}(\vec{u}) = \frac{\ell(\vec{u})}{\ell(\vec{u}_1)} \vec{u}_1 \quad \text{et} \quad s_{D_1/D_2}(\vec{u}) = 2 \frac{\ell(\vec{u})}{\ell(\vec{u}_1)} \vec{u}_1 - \vec{u}. \quad (7)$$

Démonstration. La première formule est clairement une application linéaire. De plus, on a

$$p_{D_1/D_2}(\vec{u}_1) = \vec{u}_1 \quad \text{et} \quad p_{D_1/D_2}(\vec{u}_2) = \vec{0} \quad \text{si} \quad \vec{u}_2 \in D_2,$$

et donc $p(\lambda_1 \vec{u}_1 + \vec{u}_2) = \lambda_1 \vec{u}_1$ est bien l'application cherchée. On note que le dénominateur $\ell(\vec{u}_1)$ de la formule est non nul, car par hypothèse $\vec{u}_1 \notin D_2$ et donc ne satisfait l'équation de D_2 . Cette formule s'applique, que les droites D_1 et D_2 soient orthogonales ou non. La seconde formule pour la symétrie provient de (6). \square

Projection orthogonale. Si de plus D_2 est la droite orthogonale à D_1 , alors on peut prendre $\ell(\vec{u}) = \langle \vec{u}_1, \vec{u} \rangle$, car alors $\ell(\vec{u}) = \langle \vec{u}_1, \vec{u} \rangle = 0$ est bien l'équation définissant les vecteurs de D_2 , orthogonaux à \vec{u}_1 . Dans ce cas la formule (7) s'écrit

$$p_{D_1/D_2}(\vec{u}) = \frac{\langle \vec{u}_1, \vec{u} \rangle}{\|\vec{u}_1\|^2} \vec{u}_1.$$

Cette expression est utile, par exemple en physique, car elle donne la composante du vecteur \vec{u} suivant \vec{u}_1 , lorsque \vec{u}_1 fait partie d'une base orthogonale.

Exemples dans \mathbb{R}^3 .

Dans \mathbb{R}^3 , un plan P et une droite D (pas nécessairement orthogonale à P) telle que $P \cap D = \{\vec{0}\}$ sont supplémentaires :

$$\mathbb{R}^3 = D \oplus P.$$

Si D est engendré par \vec{u}_1 et P a pour équation $\ell(x, y, z) = ax + by + cz = 0$, alors la formule (7)

$$p_{D/P}(\vec{u}) = \frac{\ell(\vec{u})}{\ell(\vec{u}_1)} \vec{u}_1$$

définit de nouveau la projection sur D , parallèlement à P .

9 Calcul matriciel dans des bases quelconques

Pour l'instant, nous n'avons travaillé que dans les bases canoniques de \mathbb{R}^n aux niveau des matrices. Pourtant il est utile de pouvoir choisir une base de travail adapté à chaque problème. Par exemple dans le cas d'une projection sur E par rapport à F , on préfère travailler dans une base du type $B = B_E \cup B_F$ car l'action de $p_{E/F}$ est simple sur B_E et B_F . En physique également, on cherche souvent à utiliser un repère où les forces ont des coordonnées simples.

9.1 Matrices et applications linéaires dans des bases générales

Soit $f : E \rightarrow F$ une application linéaire, entre deux (s)ev E et F munis de bases **quelconques** $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ de E et $B' = (\vec{u}'_1, \vec{u}'_2, \dots, \vec{u}'_n)$ de F .

Comme dans le théorème 3.2, f est déterminée par ses valeurs prises sur B . En effet, si

$$\vec{u} = x_1 \vec{u}_1 + x_2 \vec{u}_2 + \dots + x_p \vec{u}_p = (x_1, x_2, \dots, x_p)_B$$

alors nécessairement

$$f(\vec{u}) = x_1 f(\vec{u}_1) + x_2 f(\vec{u}_2) + \dots + x_p f(\vec{u}_p).$$

On peut comme dans le §3.2 écrire cette relation en coordonnées dans la base B' , ce qui donne en colonne une expression similaire

$$\vec{u} = (x_1, x_2, \dots, x_p)_B \mapsto f(\vec{u}) = (y_1, y_2, \dots, y_n)_{B'}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \mapsto Y = AX$$

avec

$$A = \begin{array}{cccc} f(\vec{u}_1) & f(\vec{u}_2) & \dots & f(\vec{u}_p) \\ \downarrow & \downarrow & & \downarrow \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} & \rightarrow & \vec{u}'_1 & \rightarrow \vec{u}'_2 & \rightarrow \vec{u}'_n \end{array}$$

mais où maintenant la j -ème colonne de A est constituée des coordonnées de $f(\vec{u}_j)$ dans la base B' de l'espace d'arrivée.

Définition 9.1. On appelle cette matrice A la matrice de l'application linéaire f **dans les bases B et B'** , et on la note

$$A = \text{Mat}_{B, B'}(f)$$

Les formules de calcul sont les mêmes que dans les bases canoniques, mais on manipule ici les coordonnées X de \vec{u} dans la base B de départ, et on obtient des coordonnées $Y = AX$ de $f(\vec{u})$ dans la base B' d'arrivée. On peut dire que l'on **parle** à la matrice A en base B et qu'elle **répond** en base B' .

Si $E = F$ c'est-à-dire que $f : E \rightarrow E$ et que l'on utilise la même base B au départ et à l'arrivée, on écrit simplement $A = \text{Mat}_B(f)$. Lorsque $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$ et que l'on utilise les bases canoniques, on écrit juste $\text{Mat}(f)$ comme précédemment.

Sinon, il faut garder les bases en notation, car **une** application linéaire donnée a en général **des** matrices associées très différentes en changeant les bases de travail.

Exemples. • Soit par exemple, l'application linéaire $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ associée à

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 4 & 4 \end{pmatrix}$$

dans les bases canoniques de B_3 de \mathbb{R}^3 et B_2 de \mathbb{R}^2 . Si on garde la base de départ $B = B_3$ mais que l'on utilise la base $B'_2 = (\vec{u}_1 = (1, 2), \vec{u}_2 = (3, 4))$ dans \mathbb{R}^2 , alors on a

$$A' = \text{Mat}_{B_3, B'_2}(f) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}$$

puisque

$$\begin{aligned} f(\vec{e}_1) &= (1, 2) = \vec{u}_1 = (1, 0)_{B'_2} &\Rightarrow & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ première colonne de } A', \\ f(\vec{e}_2) &= (3, 4) = \vec{u}_2 = (0, 1)_{B'_2} &\Rightarrow & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ deuxième colonne de } A', \\ f(\vec{e}_3) &= (2, 4) = 2\vec{u}_1 = (2, 0)_{B'_2} &\Rightarrow & \begin{pmatrix} 2 \\ 0 \end{pmatrix} \text{ troisième colonne de } A'. \end{aligned}$$

• Un exemple plus géométrique. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ associée à la matrice

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

On a $A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Dans la base $B' = (\vec{u}_1 = (1, 1), \vec{u}_2 = (1, -1))$ de \mathbb{R}^2 , on a donc

$$\text{Mat}_{B'}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Cela signifie que si on travaille dans la base B' , alors

$$\vec{u} = x'\vec{u}_1 + y'\vec{u}_2 \mapsto f(\vec{u}) = x'\vec{u}_1$$

L'application f est en fait la projection (orthogonale) sur la droite $y = x$ engendrée par \vec{u}_1 , le long de la droite $y = -x$ engendrée par \vec{u}_2 .

On voit sur cet exemple qu'une matrice d'apparence compliquée peut cacher une application linéaire d'interprétation claire, mais dans une autre base. Notez aussi que malgré la forme de $\text{Mat}_{B'}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, l'application f **n'est pas la projection sur l'axe des x**.

En fait, **toutes** les projections sur des droites dans \mathbb{R}^2 ont une matrice de cette forme, dans une base convenable ! En effet, celles-ci sont définies par deux droites distinctes : D_1 , la droite sur laquelle on projette, et D_2 la direction de projection (pas toujours orthogonale, penser à l'ombre du soleil). Si on prend une base $B' = (\vec{u}_1, \vec{u}_2)$ de \mathbb{R}^2 avec $\vec{u}_1 \in D_1$ et $\vec{u}_2 \in D_2$, alors la formule de projection sur D_1 le long de D_2 est

$$p_{D_1/D_2}(x'\vec{u}_1 + y'\vec{u}_2) = x'\vec{u}_1.$$

La matrice de p_{D_1/D_2} dans la base B' est donc $\text{Mat}_{B'}(p) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

9.2 Calcul matriciel général

Bases, matrices et applications coordonnées.

Pour comprendre les résultats à venir, on explique différemment la relation entre une application linéaire donnée $f : E \rightarrow F$ et ses matrices dans différentes bases.

Comme on l'a déjà remarqué dans la définition 6.9, le choix d'une base $B = (\vec{u}_1, \dots, \vec{u}_p)$ de E donne deux applications naturelles réciproques l'une de l'autre,

- La première, $L_B : \mathbb{R}^p \rightarrow E$ envoie des coordonnées $X = (x_1, \dots, x_p)$ dans \mathbb{R}^p sur le vecteur correspondant $\vec{v} = L_B(X) = x_1\vec{u}_1 + \dots + x_p\vec{u}_p$ de E .
- Réciproquement, « l'application coordonnées dans B » notée $C_B : E \rightarrow \mathbb{R}^p$ envoie $\vec{v} \in E$ sur ses coordonnées X dans B .

$$X \in \mathbb{R}^p \begin{array}{c} \xrightarrow{L_B} \\ \xleftarrow{C_B} \end{array} \vec{v} \in E$$

Si f est une application linéaire « abstraite » entre E et F , alors le choix de bases de E et F donnent des « cartes » de E vers \mathbb{R}^p et F vers \mathbb{R}^n . Cela permet de traduire concrètement f en une matrice. En fait, l'application

$$F = C_{B'} \circ f \circ L_B : \mathbb{R}^p \xrightarrow{L_B} E \xrightarrow{f} F \xrightarrow{C_{B'}} \mathbb{R}^n$$

envoie \vec{e}_i sur $C_{B'}(f(\vec{u}_i))$. C'est précisément la i ème colonne de $\text{Mat}_{B,B'}(f)$ définie précédemment. On a donc en fait,

$$\text{Mat}_{B,B'}(f) = \text{Mat}(F) \tag{8}$$

où la seconde matrice est celle de l'application linéaire induite $F : \mathbb{R}^p \rightarrow \mathbb{R}^n$ dans les bases canoniques.

Produit, composition, inversion.

Proposition 9.2. Soient E, F, G trois sev munis de bases B, B' et B'' . Soient $f : E \rightarrow F$ et $g : F \rightarrow G$ deux applications linéaires. Alors on a $g \circ f : E \rightarrow G$ et

$$\text{Mat}_{B, B''}(g \circ f) = \text{Mat}_{B', B''}(g) \times \text{Mat}_{B, B'}(f)$$

Démonstration. Avec (8) cela vient de la même règle vue dans le théorème 4.4 pour les bases canoniques. On utilise $L_{B'} = C_{B'}^{-1}$ pour écrire

$$C_{B''} \circ (g \circ f) \circ L_B = (C_{B''} \circ g \circ L_{B'}) \circ (C_{B'} \circ f \circ L_B) = G \circ F$$

et on prend les matrices de F et G dans les bases canoniques. □

On a de la même façon.

Proposition 9.3. Soient E et F deux ev de bases B et B' . Alors $f : E \rightarrow F$ est un isomorphisme ssi sa matrice $\text{Mat}_{B, B'}(f)$ est inversible et on a

$$\text{Mat}_{B', B}(f^{-1}) = [\text{Mat}_{B, B'}(f)]^{-1}.$$

Démonstration. On applique cette propriété déjà vue dans la base canonique à $F = C_{B'} \circ f \circ L_B$. On a F isomorphisme ssi f l'est et $F^{-1} = L_B^{-1} \circ f^{-1} \circ C_{B'}^{-1} = C_B \circ f^{-1} \circ L_{B'}$. □

9.3 Changements de bases et matrices de passage

Nous allons étudier l'effet d'un changement de base sur les coordonnées d'un vecteur donné. Soient $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $B' = (\vec{u}'_1, \vec{u}'_2, \dots, \vec{u}'_n)$ deux bases d'un ev E . Un vecteur \vec{v} donné a des coordonnées X dans la base B et X' dans la base B' . On a

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{L_{B'}} & E & \xrightarrow{C_B} & \mathbb{R}^n \\ X' & \longmapsto & \vec{v} & \longmapsto & X \end{array}$$

C'est-à-dire $X = C_B \circ L_{B'}(X')$.

Définition 9.4. On note $P_B^{B'}$ l'isomorphisme $C_B \circ L_{B'}$ de \mathbb{R}^n . Sa matrice (dans la base canonique de \mathbb{R}^n) s'appelle la **matrice de passage de B à B'** .

Par définition ses colonnes sont constituées des coordonnées des vecteurs \vec{u}'_j de la « nouvelle base » B' exprimés dans « l'ancienne base » B .

Théorème 9.5. 1. On a $X = P_B^{B'} X'$.
 2. Toute matrice de passage est inversible et $(P_B^{B'})^{-1} = P_B^B$.

Démonstration. On a déjà vu 1. Pour 2, $P_B^{B'} = C_B \circ L_{B'}$ est un isomorphisme car C_B et $L_{B'}$ le sont. Son inverse est

$$(P_B^{B'})^{-1} = (L_{B'})^{-1} \circ C_B^{-1} = C_{B'} \circ L_B = P_B^B.$$

□

Remarque 9.6. En fait, d'après (8), $P_{B,B'}$ est aussi la matrice de l'application identité

$$\begin{aligned} \text{Id} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \vec{v} &\mapsto \vec{v} \end{aligned}$$

écrite dans B' au départ et dans B à l'arrivée ! C'est-à-dire que

$$P_B^{B'} = \text{Mat}_{B',B}(\text{Id}). \tag{9}$$

On remarque inversement que toute matrice inversible P s'interprète comme une matrice de passage. En effet, cette matrice P est de rang n , c'est-à-dire que ses colonnes forment une famille génératrice de n vecteurs de \mathbb{R}^n , et donc une base. P est alors la matrice de passage $P_B^{B'}$ où B' est constituée des vecteurs colonnes dans la base initiale.

Attention à la terminologie !

Malgré son nom, la matrice de passage **de B à B'** permet de calculer les coordonnées d'un vecteur \vec{v} dans B (« l'ancienne base ») connaissant celles dans B' (la « nouvelle base »), et non l'inverse.

Si on veut calculer les **nouvelles** coordonnées à l'aide des **anciennes**, il faut échanger B et B' , c'est-à-dire utiliser

$$X' = P_{B'}^B X = (P_B^{B'})^{-1} X.$$

Il faut donc inverser la matrice $P_B^{B'}$.

L'application $P_B^{B'}$ pousse bien B sur B' mais fait le contraire au niveau des coordonnées. On dit que les coordonnées sont des quantités **contravariantes**.²

2. On pourrait dire contrariantes ;)

Un exemple. Dans \mathbb{R}^2 , on considère la base canonique $B = B_{can} = (\vec{e}_1, \vec{e}_2)$ et la base tournée d'un angle θ : $B' = (\vec{v}_1, \vec{v}_2)$ avec

$$\vec{v}_1 = (\cos \theta, \sin \theta) \quad \text{et} \quad \vec{v}_2 = (-\sin \theta, \cos \theta).$$

La matrice de passage de B à B' est $P_B^{B'} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ et si $\vec{v} = (x, y)$ dans la base canonique B , on a $\vec{v} = (x', y')_{B'}$ dans la base B' avec

$$\begin{aligned} X = \begin{pmatrix} x \\ y \end{pmatrix} &= P_B^{B'} X' = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &\Leftrightarrow \begin{cases} x = x' \cos \theta - y' \sin \theta \\ y = x' \sin \theta + y' \cos \theta \end{cases} \end{aligned}$$

qui exprime les anciennes coordonnées (que l'on connaît déjà, dans la base canonique) à l'aide des nouvelles (que l'on voudrait connaître, dans B') !

Pour avoir X' à l'aide de X , il faut inverser $P_B^{B'}$, soit en utilisant la relation générale du théorème 5.8, soit en remarquant que $P_B^{B'}$ est en fait la matrice de la rotation d'angle θ , et que donc son inverse est la matrice de rotation d'angle $-\theta$:

$$(P_B^{B'})^{-1} = P_{B'}^B = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

On a donc

$$X' = (P_B^{B'})^{-1} X \Leftrightarrow \begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases}$$

La transformation $X \mapsto X'$ est la rotation d'angle $-\theta$. Cela signifie que, **vu de la nouvelle base B'** , le vecteur \vec{v} **semble** avoir tourné en sens inverse de $-\theta$ par rapport à sa position dans la base de départ. En fait, ce n'est pas \vec{v} qui a bougé, mais le repère !

C'est un point important à bien saisir dans cette notion de changement de base : on s'intéresse à l'application identité, qui ne déplace aucun vecteur, mais exprimée dans des bases différentes. Les coordonnées changent, pas les vecteurs. La physique exploite ce genre d'idée, en privilégiant des lois vectorielles qui ne dépendent pas du repère choisi. Par exemple le principe fondamental de la dynamique de Newton :

$$\sum \overrightarrow{\text{forces}} = \frac{d\vec{p}}{dt}$$

est une loi **vectorielle** valable dans tous les repères galiléens. Cette loi ne s'appuie pas sur un choix explicite de coordonnées.

9.4 Changement de base pour les endomorphismes

Soit $f : E \rightarrow E$ un endomorphisme. Étant données deux bases $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $B' = (\vec{u}_1', \vec{u}_2', \dots, \vec{u}_n')$ de E , on peut associer deux matrices à f :

$$A = \text{Mat}_B(f) \quad \text{et} \quad A' = \text{Mat}_{B'}(f).$$

On note enfin $P = P_B^{B'}$ la matrice de passage de B à B' .

Théorème 9.7. *Avec les notations précédentes, on a*

$$A' = P^{-1}AP.$$

Démonstration. On donne deux preuves.

- La première est abstraite : on exprime $f = \text{Id} \circ f \circ \text{Id}$ (!) matriciellement dans les bases

$$\begin{aligned} {}_{B'}f_{B'} &= ({}_{B'}\text{Id}_B)({}_Bf_B)({}_B\text{Id}_{B'}) \\ \Leftrightarrow \text{Mat}_{B',B'}(f) &= \text{Mat}_{B',B'}(\text{Id}) \text{Mat}_{B,B}(f) \text{Mat}_{B',B}(\text{Id}) \quad \text{d'après (9),} \\ \Leftrightarrow A' &= P^{-1}AP. \end{aligned}$$

- La seconde preuve utilise la formule de changement de base pour les vecteurs. Soit \vec{u} quelconque et $\vec{v} = f(\vec{u})$. On note X et X' les coordonnées de \vec{u} , et Y et Y' celles de \vec{v} . On a, d'après le théorème 9.5,

$$\begin{aligned} AX &= Y \text{ avec } Y = PY' \text{ et } X = PX' \\ \Rightarrow APX' &= PY' \\ \Rightarrow (P^{-1}AP)X' &= Y' = A'X'. \end{aligned}$$

Comme cela est vrai pour tout $X' \in \mathbb{R}^n$, les deux matrices A' et $P^{-1}AP$ sont égales (en prenant par exemple pour X' les vecteurs de la base canonique de \mathbb{R}^n). \square

La formule précédente comporte l'inversion de la matrice de passage P , suivi du produit de trois matrices. Ces calculs ne posent pas de problèmes pour les ordinateurs, mais sont lourds et source d'erreurs quand ils sont fait à la main. Dans les exercices, il est souvent plus intéressant et rapide de revenir à la définition de $A' = \text{Mat}_{B'}(f)$ pour la calculer. On calcule l'image des vecteurs de B' par f dans la base B , en utilisant A . On exprime ensuite les résultats dans B' . Les colonnes obtenues sont celles de A' . C'est formellement équivalent au calcul général, mais souvent dans les exercices, l'image de la base B' proposée s'exprime simplement dans B' .

Exemple. Soit $A = \begin{pmatrix} -1 & 2 & -2 \\ -2 & 3 & -2 \\ -2 & 2 & -1 \end{pmatrix}$ et $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $X \mapsto AX$ l'application linéaire associée dans la base canonique B_{can} . On considère les vecteurs

$$\vec{u}_1 = (1, 1, 0), \quad \vec{u}_2 = (0, 1, 1) \text{ et } \vec{u}_3 = (1, 1, 1).$$

On montre que $B' = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est une base de \mathbb{R}^3 en calculant le rang de $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

On a

$$\text{rang}(P) = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{pmatrix} = 3,$$

et donc B' est une base, et P s'appelle alors la matrice de passage de B à B' . Pour calculer la matrice de f dans la base B' , on peut inverser P et calculer le produit $A' = P^{-1}AP$.

Ici, il est plus efficace de remarquer que

$$A \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{et} \quad A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Leftrightarrow f(\vec{u}_1) = \vec{u}_1, \quad f(\vec{u}_2) = \vec{u}_2 \quad \text{et} \quad f(\vec{u}_3) = -\vec{u}_3.$$

On a donc $A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Géométriquement, f est la symétrie (non orthogonale) par rapport au plan P engendré par \vec{u}_1 et \vec{u}_2 , et le long de la droite D engendrée par \vec{u}_3 .

Dans la base B' , l'action de f est simple

$$f(x'\vec{u}_1 + y'\vec{u}_2 + z'\vec{u}_3) = x'\vec{u}_1 + y'\vec{u}_2 - z'\vec{u}_3.$$

On peut remarquer que $f \circ f = \text{Id}$, comme c'est le cas de toutes les symétries. Cela se traduit matriciellement dans toutes les bases. On a la fois

$$A'^2 = \text{Mat}_{B'}(f)^2 = \text{Mat}_{B'}(f^2) = \text{Mat}_{B'}(\text{Id}) = I_3$$

ce qui est clair sur A' , mais aussi nécessairement

$$A^2 = \text{Mat}_B(f)^2 = \text{Mat}_B(f^2) = \text{Mat}_B(\text{Id}) = I_3,$$

ce qui ne saute pas aux yeux !

9.5 La trace d'une matrice

On voit sur l'exemple précédent que les matrices A et A' ont peu de points communs. Il y a quand même un nombre simple à calculer qui est préservé (à part le rang).

Définition 9.8. Soit

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in M_n(\mathbb{R})$$

une matrice carrée. La trace de A est la somme des coefficients diagonaux de A :

$$\text{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Dans l'exemple précédent, on a $\text{Tr}(A) = \text{Tr}(A') = 1$. Cette coïncidence est toujours vérifiée dans un changement de base d'un endomorphisme.

Théorème 9.9. *Si A et $A' \in M_n(\mathbb{R})$ satisfont $A' = P^{-1}AP$, alors*

$$\text{Tr}(A') = \text{Tr}(A).$$

Cela provient du lemme suivant.

Lemme 9.10. *Soient $A, B \in M_n(\mathbb{R})$, alors on a $\text{Tr}(AB) = \text{Tr}(BA)$.*

Démonstration du lemme. D'après (4), on a

$$(AB)_{ii} = \sum_{j=1}^n a_{ij}b_{ji}$$

d'où

$$\begin{aligned} \text{Tr}(AB) &= \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ji} \\ &= \sum_{j=1}^n \sum_{i=1}^n b_{ji}a_{ij} \\ &= \sum_{j=1}^n (BA)_{jj} = \text{Tr}(BA). \end{aligned}$$

□

Démonstration du théorème. On a

$$\text{Tr}(A') = \text{Tr}(P^{-1}AP) = \text{Tr}(APP^{-1})$$

par le lemme avec $A = P^{-1}$ et $B = AP$. Comme $PP^{-1} = I_3$, on obtient

$$\text{Tr}(A') = \text{Tr}(AI_3) = \text{Tr}(A).$$

□

10 Évolution de systèmes et matrices de transition

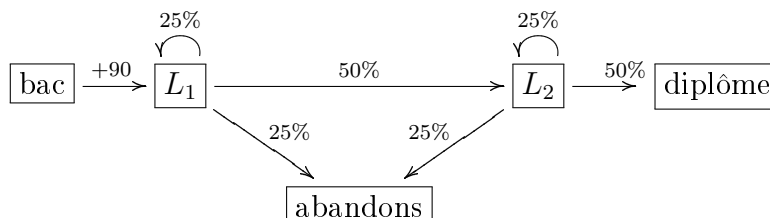
Nous allons utiliser le calcul matriciel pour modéliser l'évolution d'un système possédant un nombre fini « d'états ». Nous traitons un premier exemple pour se familiariser avec cette notion.

10.1 Un exemple : les effectifs dans un cycle d'étude

Le problème. Nous voulons étudier l'évolution des effectifs des étudiants dans un cycle d'étude de deux ans L_1 et L_2 . On a les données suivantes :

- chaque année, 90 nouveaux étudiants arrivent en L_1 ;
- 50% des étudiants de L_1 passent en L_2 , 25% redoublent et 25% abandonnent ;
- 50% des étudiants de L_2 ont le diplôme, 25% redoublent et 25% abandonnent.

On peut résumer les données sous forme d'un graphe :



On voudrait connaître :

1. l'évolution des effectifs, se stabilisent-ils, explosent-ils ?
2. le taux de réussite global du cycle.

Modélisation. On regroupe les effectifs de l'année n dans un vecteur

$$X(n) = \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix},$$

et on écrit une relation de récurrence entre $X(n+1)$ et $X(n)$. On a

$$X(n+1) = \begin{pmatrix} x_1(n+1) \\ x_2(n+1) \end{pmatrix} = \begin{pmatrix} \frac{1}{4}x_1(n) + 90 \\ \frac{1}{2}x_1(n) + \frac{1}{4}x_2(n) \end{pmatrix} = \begin{pmatrix} 1/4 & 0 \\ 1/2 & 1/4 \end{pmatrix} X(n) + \begin{pmatrix} 90 \\ 0 \end{pmatrix}$$

c'est-à-dire

$$\left| \begin{array}{l} X(n+1) = MX(n) + E, \\ \text{avec } M = \begin{pmatrix} 1/4 & 0 \\ 1/2 & 1/4 \end{pmatrix} \text{ et } E = \begin{pmatrix} 90 \\ 0 \end{pmatrix} \text{ est le vecteur des entrées.} \end{array} \right. \quad (10)$$

C'est une relation de récurrence sur le vecteur $X(n)$. Cette suite est déterminée par les effectifs initiaux $X(0)$.

Résolution dans le cas stable. On peut facilement implanter le calcul de la suite $X(n)$ avec Python pour différentes *conditions initiales* $X(0)$, et en étudier numériquement l'évolution. On peut aussi résoudre théoriquement le problème.

- On recherche d'abord les **effectifs stables possibles**. Dans ce cas, on doit avoir

$$X(n+1) = X(n) = X.$$

D'où, d'après (10)

$$X = MX + E \Leftrightarrow (I_2 - M)X = E.$$

La matrice $I_2 - M = \begin{pmatrix} 3/4 & 0 \\ -1/2 & 3/4 \end{pmatrix}$ est échelonnée et inversible, de déterminant $9/16$ et d'inverse

$$(I_2 - M)^{-1} = \begin{pmatrix} 4/3 & 0 \\ 8/9 & 4/3 \end{pmatrix},$$

d'après l'expression générale vue dans le chapitre précédent. On a donc

$$X = (I_2 - M)^{-1}(E) = \begin{pmatrix} 120 \\ 80 \end{pmatrix}, \quad (11)$$

c'est-à-dire que si les effectifs sont stables, on a

- 120 étudiants en L_1 ;
- et 80 en L_2 .
- Dans ce cas, on a de plus un taux de réussite globale

$$\frac{\text{diplômés}}{\text{entrants}} = \frac{50\% \times 80}{90} = \frac{40}{90} \simeq 44\%,$$

ce qui peut paraître élevé à première vue par rapport aux taux de réussite des années prisent séparément.²

Évolution des effectifs en général. Pour étudier l'évolution générale des effectifs, nous les comparons aux effectifs stables $X = \begin{pmatrix} 120 \\ 80 \end{pmatrix}$, en considérant la différence

$$\Delta(n) = X(n) - X.$$

Comme $E = (I_2 - M)X$, la récurrence (10) s'écrit

$$\begin{aligned} \Delta(n+1) &= X(n+1) - X = MX(n) + (I_2 - M)X - X \\ &= MX(n) - MX = M\Delta(n). \end{aligned}$$

On a donc par une récurrence immédiate

$$\Delta(n) = M^n \Delta(0).$$

2. Ce qui se passe, c'est que les étudiants de cette filière peuvent avoir leur diplôme « à l'usure », en redoublant plusieurs fois.

Il faut calculer les puissances de M . On trouve

$$M = \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad M^2 = \frac{1}{16} \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}, \quad M^3 = MM^2 = \frac{1}{4^3} \begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix}.$$

Ceci nous fait deviner que

$$M^n = \frac{1}{4^n} \begin{pmatrix} 1 & 0 \\ 2n & 1 \end{pmatrix},$$

ce qui se vérifie facilement par récurrence en utilisant que $M^{n+1} = M^n M$.

Comme les coefficients de M^n tendent tous (très vite) vers 0 lorsque $n \rightarrow +\infty$, on en déduit que les composantes de $\Delta(n) = X(n) - X = M^n \Delta(0)$ tendent vers 0 quand $n \rightarrow +\infty$.

En conclusion, on voit que les effectifs du cycle convergent vers les effectifs stables : 120 étudiants en L_1 et 90 en L_2 , et ceci quels que soient les effectifs initiaux !

10.2 Matrices de transition et évolution de systèmes

L'exemple précédent décrit l'évolution d'une population à deux états (les deux années du cycle) et dont les transitions d'un état à l'autre sont données par des probabilités. On peut modéliser systématiquement ce genre de problème par du calcul matriciel.

Les chaînes de Markov.

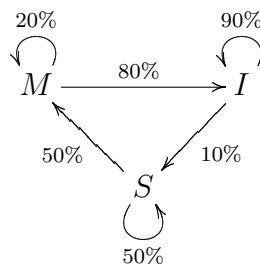
On se donne un système qui peut être dans un certain nombre d'états E_1, E_2, \dots, E_n .
On suppose qu'il évolue d'un instant t à $t + 1$ en suivant des *probabilités de transition*.

On note p_{ij} = probabilité de passer de l'état E_j à l'état E_i entre t et $t + 1$.

Prenons par exemple le problème d'épidémie suivant, à trois états. On peut être malade (M), immunisé (I), ou sain mais non immunisé (S). D'une semaine à l'autre,

- 20% des malades le restent, 80% guérissent et sont immunisés ;
- 90% des immunisés le restent, 10% ne le sont plus mais restent sains (mutation du virus par exemple) ;
- 50% des personnes saines mais non immunisées le restent, 50% tombent malades.

L'ensemble des données peut se représenter dans un graphe, qui s'appelle une *chaîne de Markov*.



C'est joli, mais pour faire des calculs et quand le nombre des états augmente, il est bien plus efficace de représenter les données sous forme matricielle.

Définition 10.1. La *matrice de transition* de la chaîne de Markov est la matrice $M = (p_{ij})$, où p_{ij} représente la probabilité de passer de E_j à E_i entre un instant et le suivant.

Cela donne dans notre exemple avec $E_1 = M$, $E_2 = I$, $E_3 = S$

$$M = \begin{pmatrix} 0.2 & 0 & 0.5 \\ 0.8 & 0.9 & 0 \\ 0 & 0.1 & 0.5 \end{pmatrix}.$$

On remarque que la **somme des coefficients de chaque colonne fait toujours 1**, c'est-à-dire que pour tout j ,

$$\sum_{i=1}^n p_{ij} = 1. \quad (12)$$

C'est une propriété générale des matrices de transition lorsque la liste des états est complète. Ces sommes représentent la probabilité totale de passer d'un état donné à n'importe quel autre.

La notion d'état mélangé. Puisque l'évolution au cours du temps de l'état de chaque individu est aléatoire, on s'intéresse plutôt à l'évolution des différentes proportions x_i de la population qui sont dans les états E_i à chaque instant. On fabrique ainsi un vecteur

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Par exemple dans le cas précédent on part de $X(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ avant l'arrivée de l'épidémie,

et si $X = \begin{pmatrix} 0.5 \\ 0.2 \\ 0.3 \end{pmatrix}$, alors 50% de la population est malade, 20% est immunisée et 30% saine mais non immunisée.³

Définition 10.2. Un *état mélangé* X du système est un vecteur de composantes $x_i \geq 0$ et satisfaisant

$$\sum_{i=1}^n x_i = 1. \quad (13)$$

Dans ce cas, les x_i s'interprètent comme les probabilités pour le système d'être dans un état E_i , ou encore si on étudie une population, la proportion des individus dans l'état E_i .

3. Attention pour écrire les nombres en Python : le séparateur décimal est le point.

Évolution du système. L'évolution de ces probabilités d'un état mélangé est très simple à calculer avec la matrice de transition M .

Théorème 10.3. On a

$$X(t+1) = MX(t).$$

Démonstration. La composante $x_i(t+1)$ de $X(t+1)$ est la probabilité d'être dans l'état E_i à l'instant $t+1$. On a donc en décomposant le calcul suivant les états possibles à l'instant t :

$$\begin{aligned} x_i(t+1) &= \sum_{j=1}^n P(E_i(t+1) \cap E_j(t)) \\ &= \sum_{j=1}^n P(E_i(t+1) \text{ sachant } E_j(t)) \times P(E_j(t)) \\ &= \sum_{j=1}^n p_{ij} x_j(t) = (MX(t))_i \end{aligned}$$

□

On en déduit immédiatement par récurrence que :

Corollaire 10.4. On a $X(n) = M^n X(0)$.

Ces calculs s'effectuent très vite avec Python par exemple, et permettent donc d'observer l'évolution de l'état du système. Dans le cas du modèle d'épidémie par exemple, on observe un phénomène de convergence vers une *état mélangé stationnaire* X lorsque $t \rightarrow +\infty$.⁴ De plus cet état mélangé limite est indépendant de la situation de départ. C'est un phénomène assez général comme on va le voir.

10.3 Existence des états mélangés stationnaires

État stationnaire.

Définition 10.5. On dit qu'un état mélangé est *stationnaire* si $X(n) = X$ est constant au cours du temps. Comme on a toujours $X(n+1) = MX(n)$, d'après le théorème 10.3, on a donc

$$MX = X$$

pour X stationnaire.

4. On verra peut-être en TP Python.

Attention, au sens de cette définition. Par exemple dans le cas d'une épidémie, le fait que X reste stable ne signifie pas que *chaque individu* reste dans le même état (sain, malade, immunisé). C'est juste la *proportion* de la population dans chaque état qui reste stable.

La première question que l'on peut se poser est de savoir si une chaîne de Markov possède toujours un état mélangé stationnaire. C'est bien le cas.

Théorème 10.6 (Existence d'état stationnaire). *Soit M une matrice de transition. Alors M possède au moins un état mélangé stationnaire.*

Démonstration. C'est un résultat intéressant et accessible, mais assez délicat, avec plusieurs étapes. On commence par montrer qu'il existe un *vecteur non nul* X tel que $MX = X$.

On utilise la *matrice ligne* $L = (1, 1, \dots, 1)$. Elle correspond à une application linéaire de \mathbb{R}^n dans \mathbb{R} définie par $LX = \sum_{i=1}^n x_i$ qui intervient dans l'équation (13) $LX = 1$ des états mélangés.

On commence par remarquer que $LM = L$. En effet

$$\begin{aligned} LM &= (1, 1, \dots, 1) \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \\ &= \left(\sum_{i=1}^n p_{i1}, \sum_{i=1}^n p_{i2}, \dots, \sum_{i=1}^n p_{in} \right) = (1, 1, \dots, 1) = L \end{aligned}$$

d'après (12).

Remarque 10.7. La propriété $LM = L$ implique en particulier que M transforme un état mélangé en un autre, comme il découle déjà du théorème 10.3. En effet, si X est un état mélangé, alors X est un vecteur à coordonnées positives et tel que $LX = 1$ par (13), alors MX est aussi à coordonnées positives (car $p_{ij} \geq 0$) et $L(MX) = LX = 1$.

On poursuit la preuve. Comme $LM = L$, alors $L(M - I_n) = 0$. Il découle que l'image de $M - I_n$ est inclus dans $\ker L$. En effet, pour tout $V \in \mathbb{R}^n$, $L((M - I_n)V) = 0$. Comme $L : \mathbb{R}^n \rightarrow \mathbb{R}$ est surjective (puisque $L \neq 0$), on a $\dim \ker L = n - 1$ par le thm du rang. Par conséquent,

$$\text{rang}(M - I_n) = \dim \text{Im}(M - I_n) \leq \dim \ker L = n - 1.$$

Mais alors, de nouveau par le thm du rang appliqué à $M - I_n$, on a

$$\dim \ker(M - I_n) = n - \text{rang}(M - I_n) \geq 1,$$

et il existe un vecteur non nul X tel que $(M - I_n)X = 0 \Leftrightarrow MX = X$.

Ce n'est pas fini car le vecteur $X = (x_1, x_2, \dots, x_n)$ n'est pas nécessairement un état mélangé! Il faudrait en plus que $x_i \geq 0$ et $\sum_{i=1}^n x_i = 1$. On considère à la place le vecteur

noté $|X| = (|x_1|, |x_2|, \dots, |x_n|)$. De $MX = X$, on déduit que pour chaque i

$$(M|X|)_i = \sum_{j=1}^n p_{ij}|x_j| \geq \left| \sum_{j=1}^n p_{ij}x_j \right| \geq |(MX)_i| = |x_i| = |X|_i. \quad (14)$$

On a aussi

$$\sum_{i=1}^n (M|X|)_i = L(M|X|) = L|X| = \sum_{i=1}^n |X|_i,$$

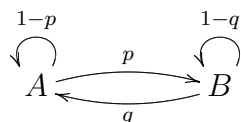
puisque $LM = M$. La somme des inégalités (14) est donc une égalité. Il s'ensuit que chaque inégalité (14) est une égalité. D'où $(M|X|)_i = |X|_i$ pour tout i et donc $M|X| = |X|$. Alors le vecteur multiple $Y = |X| / \sum_{i=1}^n |x_i|$ est un état mélangé stationnaire. \square

Concrètement, pour calculer un état stationnaire, il faut résoudre le système linéaire $MX = X \Leftrightarrow (M - I_n)X = 0$ et chercher une solution avec $x_i \geq 0$. On renormalise ensuite le vecteur obtenu pour avoir aussi $\sum_{i=1}^n x_i = 1$.

10.4 Le théorème de Perron-Frobenius.

La question suivante est de savoir si la suite des états mélangés $X(n)$ converge (se « stabilise ») vers un état stationnaire? Celui-ci est-il indépendant de la condition initiale? Ceci n'est pas vrai en général sans hypothèse supplémentaire sur M .

Exemple 10.8. Considérons un système à deux états A et B . On note p la probabilité de passer de A à B entre t et $t + 1$, et q celle de passer B à A .



La matrice de transition associée (avec la convention de ce cours) est

$$M = \begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix}.$$

- Si $0 < p < 1$ et $0 < q < 1$, on trouve un seul état stationnaire, et on peut montrer que $X(n)$ converge vers celui-ci (à voir en TD).

- Si $p = 0$ et $q = 0$, par contre tout état initial est stationnaire puisque $M = I_2$. Il y en a donc plusieurs possibles. Ici les sous-systèmes A et B sont déjà stables.

- Si $p = 1$ et $q = 1$ par exemple, le système oscille indéfiniment entre A et B . On a $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. On vérifie qu'il y a un unique état mélangé stationnaire : $X = (1/2, 1/2)$. Par contre, comme $M^2 = I_2$, la suite de vecteur $X(n) = M^n X(0)$ est périodique de période 2 et ne converge donc pas du tout vers X (sauf si on part de X)!

On introduit une condition qui permet d'éviter ces cas de figure.

Définition 10.9. On dit que la matrice de transition est *irréductible* si il existe une puissance k telle que **tous** les coefficients de M^k sont > 0 .

Par exemple, dans le cas précédent, M est irréductible ssi $0 < p < 1$ et $0 < q < 1$.

Dans le modèle d'épidémie du paragraphe 10.2, M est irréductible puisque

$$M^2 = \begin{pmatrix} 0.04 & 0.05 & 0.35 \\ 0.88 & 0.81 & 0.4 \\ 0.08 & 0.14 & 0.25 \end{pmatrix}.$$

L'hypothèse d'irréductibilité assure que M *mélange* les différents états. On peut passer de n'importe quel état j à un état k en k étapes. L'évolution du système ne possède pas de « cycle » ni de « cul-de-sac ». Il ne se décompose pas non plus en sous-systèmes qui évolueraient sans interaction.

On a alors le résultat général suivant.

Théorème 10.10 (Théorème de Perron-Frobenius). *Si M est irréductible, alors il existe un unique état mélangé stationnaire X , et de plus, quel que soit l'état mélangé initial du système, on a*

$$X(n) \rightarrow X \text{ lorsque } n \rightarrow +\infty.$$

On a vu la preuve de l'existence d'un état mélangé stationnaire en général. Mais son unicité et le résultat de convergence est admis car la démonstration est du niveau L3 math. Cela ne doit pas empêcher de connaître ni d'utiliser ce résultat pratique !

10.5 Sur le PageRank de Google

Les notions précédentes ont de nombreuses applications, en biologie, statistique, etc. Une application assez récente et spectaculaire est liée au PageRank de Google.

Lorsque l'on envoie une requête à un moteur de recherche, celui-ci doit d'abord rechercher tous les sites référencés liés à votre liste de mots, et ensuite, *classer* les résultats pour les présenter. Il faut donc en quelque sorte attribuer une « note » aux documents, sans pour autant qu'aucun humain ne les ait tous lus pour juger de leur intérêt relatif ! Cette note est le *PageRank*, dont l'efficacité fut une raison du succès de Google.

L'idée est la suivante. Le réseau internet est modélisé par un graphe gigantesque. Ces sommets sont les pages référencées ($N \simeq 25 \times 10^9$ avec les chiffres de 2007). Des liens

pointent de ces pages vers d'autres pages.

On imagine maintenant que l'utilisateur clique *au hasard* sur ces liens et on essaye de déterminer l'importance d'une page donnée en fonction de la fréquence ou probabilité de tomber dessus en un grand nombre de clics. Cela ressemble aux problèmes précédents de chaîne de Markov, avec des états numérotés par les différentes pages, mais cela ne marche pas tel quel. En effet, certaines pages sont des « culs-de-sac » car elles ne contiennent pas de liens. Pour fabriquer une matrice de transition avec le réseau, on rajoute des liens virtuels un peu partout, mais de très faible poids, à l'aide d'un paramètre $0 < q < 1$ supplémentaire.⁵

Supposons qu'il y ait N pages P_j référencées, et que chaque page P_j contienne k_j liens. Les fondateurs de Google, Brin et Page, ont défini la « matrice du réseau » G , de taille $N \times N$ par

$$G_{ij} = \begin{cases} \frac{1-q}{k_j} + \frac{q}{N} & \text{si } P_j \text{ pointe vers } P_i, \\ \frac{q}{N} & \text{si } P_j \text{ ne pointe pas vers } P_i \text{ et } k_j \neq 0, \\ \frac{1}{N} & \text{si } k_j = 0. \end{cases}$$

Proposition 10.11. *La matrice G est une matrice de transition.*

Démonstration. On a pour tout j ,

$$\sum_{i=1}^N g_{ij} = \begin{cases} k_j \left(\frac{1-q}{k_j} + \frac{q}{N} \right) + (N - k_j) \frac{q}{N} = 1 & \text{si } k_j \neq 0, \\ N \times \frac{1}{N} = 1 & \text{si } k_j = 0. \end{cases}$$

□

Comme de plus, on a $g_{ij} > 0$ pour tout i, j , la matrice G est irréductible et le théorème 10.10 s'applique. On a donc qu'il existe un état stationnaire unique R tel que quelque soit l'état X d'origine

$$G^n X \rightarrow R.$$

La composante R_i de R représente la fréquence de visite de la page P_i , *en naviguant au hasard* sur internet. C'est une mesure de « l'importance » de cette page, et est un ingrédient important du PageRank de la page P_i . Bien sûr, ce n'est pas le seul, le business ayant son mot à dire !

Les étudiants que ces notions intéressent peuvent consulter les pages Wikipedia des chaînes de Markov ou du PageRank.

5. Il semble que $q \simeq 0,15$ soit utilisé en pratique.

11 Diagonalisation en dimension 2

On a vu dans le chapitre précédent des problèmes d'évolution de systèmes mettant en jeu une suite de vecteurs $X(n)$ satisfaisant une relation récurrence du type $X(n+1) = MX(n)$ avec M matrice de transition.

La solution abstraite est $X(n) = M^n X(0)$ mais le problème concret reste d'évaluer M^n en général, ou au moins de donner son comportement lorsque n est grand. Le théorème de Perron-Frobenius répond à la question pour des chaînes de Markov irréductibles mais reste un peu mystérieux ! Existe-t-il une méthode générale pour attaquer le problème pour d'autres situations mettant en jeu des puissances de matrices ?

11.1 Motivations. Exemples de problèmes linéaires en dimension 2

Voici d'autres exemples de problèmes d'évolution qui peuvent se modéliser en algèbre linéaire. Mais attention, ces modèles ne sont que des modèles de phénomènes naturels simplifiés sous certaines hypothèses, pas la réalité !

La suite de Fibonacci.

Voici le problème posé par Leonardo Fibonacci en 1202. Il s'agit d'une modélisation de la croissance d'une population de lapins.

« Quelqu'un a déposé un couple de lapins dans un certain lieu, clos de toutes parts, pour savoir combien de couples seraient issus de cette paire en une année, car il est dans leur nature de générer un autre couple en un seul mois, et qu'ils enfantent dans le second mois après leur naissance. »

Ce n'est évidemment qu'un modèle ! Ici les lapins n'ont pas de prédateurs, pas de problème de place ni de nourriture, et ils sont immortels ! Cela peut malgré tout marcher un peu sur des laps de temps assez court par rapport à leur durée de vie et dans un environnement protégé.

On note u_n le nombre de couple de lapins l'année n . On a ici $u_1 = u_2 = 1$, puis $u_3 = u_2 + u_1 = 2$, $u_4 = u_3 + u_2 = 3$, etc... Cette suite de nombres entiers s'appelle la *suite de Fibonacci* (voir la page Wikipedia : Suite de Fibonacci).

C'est la suite définie par la donnée de la condition initiale $u_0 = 0$ et $u_1 = 1$ (ou $u_1 = u_2 = 1$) et la formule de récurrence

$$u_{n+2} = u_{n+1} + u_n.$$

On aimerait avoir une formule pour u_n , et connaître en particulier la vitesse d'expansion de la population. Par exemple, est-ce que u_{n+1}/u_n converge ?

Puisqu'en fait la suite de récurrence dépend de deux termes consécutifs, une idée est de suivre l'évolution d'un vecteur de \mathbb{R}^2 . Si on pose

$$X(n) = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix},$$

alors on a

$$X(n+1) = \begin{pmatrix} u_{n+1} \\ u_{n+2} \end{pmatrix} = \begin{pmatrix} u_{n+1} \\ u_{n+1} + u_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = AX(n)$$

avec $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. La solution matricielle est donc $X(n) = A^n X(0)$. La solution du problème revient donc à trouver une méthode pour calculer explicitement A^n !

De manière plus générale, toutes les suites récurrentes de la forme

$$u_{n+2} = au_{n+1} + bu_n$$

se modélisent de la même manière $X(n+1) = AX(n)$ avec une matrice associée

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

Bib Bip et les coyotes.

Voici un exemple de système proies-prédateurs tiré du livre « Linear Algebra with Applications » d'Otto Bretscher.

On trouve dans le désert du Mexique deux espèces : les coyotes et les coucous-coueurs ou Grand Géocoucou)⁶. Si si, Bip Bip existe vraiment et ces oiseaux sont tellement rapides qu'il peuvent attaquer et tuer un crotale, voir la vidéo Géocoucou vs crotale ! D'une année à l'autre la population $c(t)$ des coyotes et $r(t)$ des coucous (roadrunners en anglais) évolue suivant les formules

$$\begin{cases} c(t+1) = 0.86c(t) + 0.08r(t) \\ r(t+1) = -0.12c(t) + 1.14r(t) \end{cases}$$

tant que ces quantités restent positives. Remarquer les signes des coefficients, les coucous se reproduisent d'autant plus qu'il y a de coucous (à manger!) ($+0.08r(t)$), alors que c'est l'inverse pour les coyotes ($-0.12c(t)$).

Ici le vecteur $X(t) = (c(t), r(t))$ suit la loi d'évolution

$$X(t+1) = AX(t) \quad \text{avec } A = \begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix},$$

et il faut encore calculer A^n pour prédire l'évolution du système.

Un exemple en biologie : le couple glucose-insuline.

Encore un exemple tiré du livre « Linear Algebra with Applications » d'Otto Bretscher.

L'insuline est une hormone qui régule le taux de glucose dans le sang. On mesure leur concentration ($g(t), h(t)$) heure après heure après un repas copieux, relativement à leur taux à jeun. On peut modéliser que

$$\begin{cases} g(t+1) = 0.9g(t) - 0.4h(t) \\ h(t+1) = 0.1g(t) + 0.9r(t) \end{cases},$$

6. Leurs aventures ont été immortalisées dans les dessins animés « Bib bip et Vil Coyote ».

ce qui est encore un système d'évolution du type

$$X(t+1) = AX(t) \quad \text{avec} \quad A = \begin{pmatrix} 0.9 & -0.4 \\ 0.1 & 0.9 \end{pmatrix}.$$

Comment les concentrations vont-elles évoluer vers leur valeur à jeun ((0, 0) ici) ?

Deux oscillateurs couplés.

Nous n'avons décrit pour l'instant que des systèmes d'évolution avec des temps t 'discrets'. On avait $t = n$ avec n entier. Il existe aussi des systèmes à temps continu, c'est-à-dire du type

$$X'(t) = AX(t) \quad \text{ou} \quad X''(t) = AX(t)$$

avec $X(t)$ vecteur et $X'(t)$, $X''(t)$ ses vecteurs dérivées (vecteurs vitesse et accélération en physique).

Voici un exemple qui vient de la mécanique. On considère deux masses M_1 et M_2 liées à un support fixe par deux ressorts R_1 et R_2 (une force élastique). M_1 et M_2 sont de plus liées entre-elles par un troisième ressort R_3 . Il s'agit de deux *oscillateurs couplés* car les masses interagissent entre-elles. Pour simplifier les calculs on suppose que les masses de M_1 et M_2 sont égales à m . On suppose aussi que les ressorts R_1 et R_2 ont même raideur k . On note k_3 la raideur de R_3 .

On mesure la position des masses $x_1(t)$ et $x_2(t)$ *par rapport* à leur position d'équilibre, c'est-à-dire que $x_1 = x_2 = 0$ est une position d'équilibre du système et on mesure les oscillations par rapport à ses positions. Quand on déplace M_1 et M_2 , il s'exerce un surcroît de force élastique sur elles. On trouve :

- sur M_1 : $f_1 = -kx_1(t) + k_3(x_2(t) - x_1(t)) = -(k + k_3)x_1(t) + k_3x_2(t)$,
- et sur M_2 : $f_2 = k_3(x_1(t) - x_2(t)) - kx_2(t) = -(k + k_3)x_2(t) + k_3x_1(t)$.

D'après la principe de la dynamique (loi de Newton) on a

$$(S) \begin{cases} mx_1''(t) = f_1 = -(k + k_3)x_1(t) + k_3x_2(t) \\ mx_2''(t) = f_2 = -(k + k_3)x_2(t) + k_3x_1(t) \end{cases}.$$

On note, comme pour des oscillateurs simples,

$$\omega_0 = \sqrt{\frac{k}{m}} \quad \text{et} \quad \omega = \sqrt{\frac{k_3}{m}}.$$

Ce sont des quantités homogènes à une fréquence (en $1/t$).

En notant $X(t) = (x_1(t), x_2(t))$, le système (S) s'écrit

$$X''(t) = AX(t) \quad \text{avec} \quad A = \begin{pmatrix} -\omega_0^2 - \omega^2 & \omega^2 \\ \omega^2 & -\omega_0^2 - \omega^2 \end{pmatrix} \quad (15)$$

On aimerait savoir comment le système oscille ? Possède-t-il des mouvements périodiques ? À quelle(s) fréquence(s) ?

On remarque que si $k_3 = 0$, c'est-à-dire qu'il n'y a pas de ressort entre les deux masses, alors $\omega = 0$ et la matrice A est diagonale. La solution est simple dans ce cas. En effet, le système associé devient

$$\begin{cases} x_1''(t) = -\omega_0^2 x_1(t) \\ x_2''(t) = -\omega_0^2 x_2(t) \end{cases}.$$

On a deux masses oscillant indépendamment (découplées). Les oscillations de chaque masse sont périodiques de période $2\pi/\omega_0$ dans ce cas avec

$$x_1(t) = C_1 \cos(\omega_0 t + \varphi_1) \quad \text{et} \quad x_2(t) = C_2 \cos(\omega_0 t + \varphi_2).$$

Remarque 11.1. À-propos de tous ces calculs avec cos, sin etc, la notation complexe est bien plus efficace pour les faire ! Les solutions complexes sont du type $x(t) = Ce^{i\omega_0 t}$ avec C complexe, et les solutions réelles sont les parties réelles de ces fonctions. On a bien

$$x'(t) = Ci\omega_0 e^{i\omega_0 t} \quad \text{et} \quad x''(t) = Ci^2\omega_0^2 e^{i\omega_0 t} = -\omega_0^2 x(t).^7$$

11.2 Matrices diagonales et matrices diagonalisables

Nous allons voir comment *diagonaliser* les matrices 2×2 . Cette technique permettra de résoudre les problèmes que nous avons abordés. La plupart des idées exposées ici s'appliquent aussi en dimension quelconque, mais pour simplifier la présentation et certaines preuves nous allons nous restreindre à \mathbb{R}^2 . Vous reverrez et approfondirez sûrement cela dans la suite de votre cursus.

Les matrices diagonales.

Définition 11.2. On dit qu'une matrice D de $M_2(\mathbb{R})$ est *diagonale* si elle s'écrit sous la forme

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

avec λ_1 et λ_2 des nombres réels.

Une matrice diagonale respecte les axes de coordonnées engendrés par la base canonique \vec{e}_1 et \vec{e}_2 et agit par homothétie sur chacune de ces droites car $D(\vec{e}_1) = \lambda_1 \vec{e}_1$ et $D(\vec{e}_2) = \lambda_2 \vec{e}_2$.

7. Après cette introduction fleuve, vous allez voir que les maths vous paraîtront faciles !

Puissances de D et orbites des vecteurs.

Le calcul des puissance D^n d'une matrice diagonale est immédiat. On a

$$D^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

En particulier la solution de la récurrence linéaire $X(n+1) = DX(n)$ avec $X(0) = (x_0, y_0)$ est

$$X(n) = (\lambda_1^n x_0, \lambda_2^n y_0).$$

Le comportement de chaque composante de $X(n)$ pour n grand dépend de λ_i . On rappelle que l'on a

$$|\lambda^n| \begin{cases} \rightarrow +\infty & \text{si } |\lambda| > 1 \\ \rightarrow 0 & \text{si } |\lambda| < 1 \\ = 1 & \text{si } \lambda = \pm 1. \end{cases}$$

On peut ainsi avoir idée de l'allure des *orbites* des vecteurs, c'est-à-dire de la suite des $X(n)$ pour différentes conditions initiales $X(0)$.

Exercice 11.3. Dessiner quelques orbites pour

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -1/2 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix}.$$

Système différentiel $X' = DX$.

Les solutions des systèmes différentiels de la forme $X'(t) = DX(t)$ se calculent aussi facilement. On obtient pour chaque composante que $x'(t) = \lambda_1 x(t)$ et $y'(t) = \lambda_2 y(t)$ d'où

$$X(t) = (e^{t\lambda_1} x_0, e^{t\lambda_2} y_0).$$

Ici, le comportement pour t grand dépend du signe des λ_i . On rappelle que

- si $\lambda_i > 0$, $e^{t\lambda_i}$ tend vers $+\infty$ si t tend vers $+\infty$,
- si $\lambda_i < 0$, $e^{t\lambda_i}$ tend vers 0 si t tend vers $+\infty$,

Exercice 11.4. Dessiner quelques orbites (trajectoires) $X(t)$ pour

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

Les endomorphismes et matrices diagonalisables.

Aucune des matrices rencontrées dans les exemples du premier paragraphe n'était diagonale. On a besoin d'une notion plus souple pour résoudre ces problèmes. C'est celle d'endomorphisme (de matrice) *diagonalisable*.

Définition 11.5. On dit qu'un endomorphisme $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est *diagonalisable* s'il existe une base B de \mathbb{R}^2 dans laquelle la matrice de f est diagonale.

De manière équivalent, on dit qu'une matrice carrée $A \in M_2(\mathbb{R})$ est diagonalisable s'il existe une matrice inversible P telle que $D = P^{-1}AP$ soit une matrice diagonale.

Ces deux définitions sont équivalentes. En effet si A est la matrice de f dans la base canonique B_{can} et P est la matrice de passage de B_{can} à la nouvelle base B , alors on sait que

$$\text{Mat}_B(f) = P^{-1} \text{Mat}_{B_{can}}(f)P = P^{-1}AP$$

et donc f est diagonalisable ssi 'sa' matrice A l'est.

On a déjà étudié des endomorphismes diagonalisables dans ce cours. Par exemple les projections et symétries sont diagonalisables.

En effet si on a une somme directe $\mathbb{R}^2 = D_1 \oplus D_2$, on peut prendre une base B de \mathbb{R}^2 de la forme $B = (\vec{u}_1, \vec{u}_2)$ avec $D_1 = \text{Vect}(\vec{u}_1)$ et $D_2 = \text{Vect}(\vec{u}_2)$. Si on note p la projection sur D_1 le long de D_2 et s la symétrie par rapport à D_1 suivant D_2 , alors on a par définition

$$\begin{cases} p(\vec{u}_1) = \vec{u}_1 & \text{et} & p(\vec{u}_2) = \vec{0} \\ s(\vec{u}_1) = \vec{u}_1 & \text{et} & s(\vec{u}_2) = -\vec{u}_2. \end{cases}$$

Les matrices de p et s dans la base B sont donc de la forme

$$P = \text{Mat}_B(p) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad S = \text{Mat}_B(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

Notez que si D_1 et D_2 ne sont pas les axes de coordonnées, les matrices de p et s ne sont pas diagonales dans la base canonique (voir chapitre 8 et 9 du cours).

11.3 Valeurs propres et vecteurs propres

On voudrait savoir comment reconnaître les endomorphismes diagonalisables et trouver des bases de diagonalisation. Cela passe par les notions vectorielles⁸ de vecteur propre et valeur propre.

8. C'est-à-dire indépendantes du repère.

Définition 11.6. Soit f un endomorphisme de \mathbb{R}^2 . On dit qu'un vecteur \vec{v} non nul est un vecteur propre de f s'il existe un nombre réel λ tel que

$$f(\vec{v}) = \lambda \vec{v}.$$

Le nombre λ associé à un vecteur propre \vec{v} s'appelle une *valeur propre*.

Un endomorphisme f de \mathbb{R}^2 est diagonalisable s'il existe une base $B = (\vec{u}_1, \vec{u}_2)$ dans laquelle la matrice de f soit diagonale. Il existe donc des réels λ_1 et λ_2 tels que

$$f(\vec{u}_1) = \lambda_1 \vec{u}_1 \quad \text{et} \quad f(\vec{u}_2) = \lambda_2 \vec{u}_2.$$

Les vecteurs \vec{u}_i de la base de diagonalisation sont donc des vecteurs propres de f et λ_1, λ_2 sur la diagonale des valeurs propres de f . On a donc par définition.

Proposition 11.7. Un endomorphisme de \mathbb{R}^2 est diagonalisable si et seulement si il possède une base de vecteurs propres.

Les endomorphismes de \mathbb{R}^2 ne sont pas tous diagonalisables. Voici deux exemples typiques.

Exemples d'endomorphismes de \mathbb{R}^2 non diagonalisables.

- Il existe des endomorphismes de \mathbb{R}^2 qui ne possèdent aucun vecteur propre.

Prenons par exemple une rotation d'angle θ non nul modulo π . On a

$$R_\theta(\vec{u}) = \cos \theta \vec{u} + \sin \theta \vec{v}$$

où $\vec{v} = R_{\pi/2} \vec{u}$ (faire un dessin). Si $\vec{u} \neq 0$, on a donc $R_\theta(\vec{u}) = \lambda \vec{u}$ ssi $\sin \theta = 0$.

- Il existe aussi des endomorphismes de \mathbb{R}^2 qui possèdent une valeur propre mais ne sont pas diagonalisables. Soit par exemple $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. On a

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow y - \lambda x = 0 \text{ et } \lambda y = 0.$$

Si $\lambda \neq 0$ on trouve $y = x = 0$ et donc λ n'est pas une valeur propre. Par contre $\lambda = 0$ est valeur propre et les vecteurs propres associés sont proportionnels à $\vec{e}_1 = (1, 0)$. On ne peut donc obtenir une base de vecteurs propres pour A .

Comment reconnaître une valeur propre ? La notion d'espace propre.

L'équation $f(\vec{v}) = \lambda \vec{v}$ se réécrit $f(\vec{v}) - \lambda \vec{v} = \vec{0}$ c'est-à-dire $(f - \lambda \text{Id})\vec{v} = \vec{0}$. On a donc

$$f(\vec{v}) = \lambda \vec{v} \Leftrightarrow \vec{v} \in \ker(f - \lambda \text{Id}).$$

Proposition 11.8. Le nombre réel λ est une valeur propre de f ssi l'espace

$$E_\lambda = \ker(f - \lambda \text{Id}) \neq \{\vec{0}\},$$

c'est-à-dire ssi $f - \lambda \text{Id}$ n'est **pas** inversible.

Lorsque λ est une valeur propre, le sous-espace vectoriel $E_\lambda = \ker(f - \lambda \text{Id})$ s'appelle l'*espace propre* associé à λ . C'est l'ensemble des vecteurs satisfaisant $f(\vec{v}) = \lambda \vec{v}$. En dimension 2, lorsque λ est une valeur propre, E_λ est soit une droite, soit tout \mathbb{R}^2 .

Exercice 11.9. Décrire f lorsque $E_\lambda = \mathbb{R}^2$.

Exercice 11.10. Trouver les valeurs propres et vecteurs propres d'une matrice diagonale.

11.4 Le polynôme caractéristique

On va voir l'outil 'magique' qui permet de calculer les valeurs propres d'une matrice 2×2 .

On a vu dans le chapitre 5.4 qu'une matrice $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est inversible ssi son déterminant $\det A = ad - bc \neq 0$. On a donc inversement que A n'est pas inversible ssi $\det A = 0$.

Remarque 11.11. Si $\det A = 0$, on voit directement que les colonnes de A sont proportionnelles. On peut aussi remarquer que les vecteurs $(d, -b)$ et $(-c, a)$ sont dans $\ker A$.

D'après la proposition 11.8, λ est donc une valeur propre de A ssi

$$\det(A - \lambda I_2) = 0.$$

Définition 11.12. Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. La fonction $\lambda \mapsto P_A(\lambda) = \det(A - \lambda I_2)$ est un polynôme de degré 2. Il s'appelle le *polynôme caractéristique* de A . On a

$$\begin{aligned} P_A(\lambda) &= \begin{vmatrix} a - \lambda & c \\ b & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - \lambda \text{Tr } A + \det A \end{aligned}$$

avec $\text{Tr } A = a + d$ et $\det A = ad - bc$.

9. Quand elles existent, et même sinon!!!

On a donc finalement le résultat suivant.

Théorème 11.13. Les valeurs propres (réelles) de A sont les (éventuelles) racines réelles du polynôme caractéristique de A .

Le polynôme caractéristique $P_A(\lambda) = \lambda^2 - \lambda \operatorname{Tr} A + \det A$ de degré 2 a au plus deux racines réelles, suivant la valeur du discriminant

$$\Delta = (\operatorname{Tr} A)^2 - 4 \det A.$$

Cela donne au plus 2 valeurs propres (réelles). En fait P_A possède *toujours* des racines

$$\lambda_1 = \frac{\operatorname{Tr} A - \sqrt{\Delta}}{2} \quad \text{et} \quad \lambda_2 = \frac{\operatorname{Tr} A + \sqrt{\Delta}}{2} \quad (16)$$

mais elles sont *complexes* lorsque $\Delta < 0$ car alors $\pm\sqrt{\Delta} = \pm i\sqrt{|\Delta|} \in \mathbb{C}$. Lorsque $\Delta = 0$, P possède une seule racine (double) $\lambda = \lambda_1 = \lambda_2 = \frac{1}{2} \operatorname{Tr} A$.

On a aussi toujours finalement $P_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$ et les relations

$$\lambda_1 + \lambda_2 = \operatorname{Tr} A \quad \text{tandis que} \quad \lambda_1 \lambda_2 = \det A.$$

Lorsque λ est une racine *complexe* (non réelle) de $P_A(\lambda)$, on dit aussi que λ est une valeur propre de A (mais complexe) cette fois. Cela correspond encore à une matrice $A - \lambda I_2$ non inversible mais agissant sur les vecteurs à coordonnées complexes. Il existe donc un vecteur $\vec{v} = (z_1, z_2) \in \mathbb{C}^2$ avec $A\vec{v} = \lambda\vec{v}$. Attention, ce vecteur n'est pas réel et ne peut être utilisé pour diagonaliser A dans \mathbb{R}^2 . On verra malgré tout quoi en faire plus loin.

Exercice 11.14. Calculer les valeurs propres et vecteur propres complexes de $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

A est-elle diagonalisable dans \mathbb{R}^2 ?

11.5 Synthèse

On peut maintenant faire la synthèse des résultats et énoncer le théorème principal de diagonalisation dans \mathbb{R}^2 .

Théorème 11.15. Soit f un endomorphisme de \mathbb{R}^2 de matrice A (dans une base quelconque). Soit

$$P_A(\lambda) = \det(A - \lambda I_2) = \lambda^2 - \operatorname{Tr} A \lambda + \det A$$

le polynôme caractéristique de A et $\Delta = (\operatorname{Tr} A)^2 - 4 \det A$ son discriminant. Alors :

1. Si $\Delta > 0$, f possède deux valeurs propres distinctes λ_1 et λ_2 données par (16) et f est diagonalisable dans une base $B = (\vec{u}_1, \vec{u}_2)$ de vecteurs propres avec $f(\vec{u}_1) = \lambda_1 \vec{u}_1$ et $f(\vec{u}_2) = \lambda_2 \vec{u}_2$.
2. si $\Delta = 0$, alors $\lambda = \frac{1}{2} \text{Tr } A$ est la seule valeur propre de f et f n'est diagonalisable que si $f = \lambda \text{Id}$.
3. Si $\Delta < 0$, f n'a pas de valeur propre réelle et n'est pas diagonalisable dans \mathbb{R}^2 .

Démonstration. 1. Il faut juste vérifier que 2 vecteurs propres (non nuls) \vec{v}_1 et \vec{v}_2 associés à des valeurs propres différentes λ_1 et λ_2 forment bien une base de \mathbb{R}^2 . Il suffit de vérifier qu'ils ne peuvent pas être colinéaires. En effet, si on avait $\vec{v}_2 = c\vec{v}_1$ alors

$$\lambda_2 \vec{v}_2 = f(\vec{v}_2) = cf(\vec{v}_1) = c\lambda_1 \vec{v}_1 = \lambda_1 \vec{v}_2 \Leftrightarrow (\lambda_2 - \lambda_1) \vec{v}_2 = \vec{0},$$

ce qui n'est pas possible avec $\lambda_1 \neq \lambda_2$ et $\vec{v}_2 \neq \vec{0}$.

2. Si f ne possède qu'une valeur propre λ est diagonalisable, alors il faut trouver une base de vecteur propre $B = (\vec{v}_1, \vec{v}_2)$ telle que $f(\vec{v}_1) = \lambda \vec{v}_1$ et $f(\vec{v}_2) = \lambda \vec{v}_2$. Mais alors $f(\vec{v}) = \lambda \vec{v}$ pour tout \vec{v} et $f = \lambda \text{Id}$ est une homothétie, avec $\text{Mat}(f) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ diagonale dans toute base!

3. Déjà vu. Le polynôme caractéristique $P_A(\lambda)$ n'a aucune racine réelle. □

Que se passe-t-il si $\Delta < 0$?

On peut compléter la discussion quand il n'y a pas de valeurs propres réelles. (C'est à la limite du programme mais utile pour traiter certains exemples du premier paragraphe.) Si $\Delta < 0$, le polynôme caractéristique possède 2 racines complexes conjuguées

$$\lambda = \frac{\text{Tr } A + i\sqrt{|\Delta|}}{2} \quad \text{et} \quad \bar{\lambda} = \frac{\text{Tr } A - i\sqrt{|\Delta|}}{2}.$$

Pour simplifier les notations, on va noter $\lambda = \alpha + i\beta$. On va montrer le résultat suivant qui permet de calculer quand même les puissances de A (de f).

Théorème 11.16. Avec les notations et hypothèses ci-dessus. Il existe une base $B = (\vec{v}_1, \vec{v}_2)$ de \mathbb{R}^2 telle que

$$\text{Mat}_B(f) = P^{-1}AP = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

Autrement dit, f agit dans la base B comme une similitude $h_\lambda : z \rightarrow \lambda z$ avec $\lambda = \alpha + i\beta$.

Si on écrit $\lambda = \alpha + i\beta = |\lambda|e^{i\theta}$ alors on a pour $n \in \mathbb{N}$,

$$\text{Mat}_B(f^n) = P^{-1}A^nP = |\lambda|^n \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{pmatrix}.$$

Démonstration. La preuve va donner la méthode de calcul de la base B .

On sait que $A - \lambda \text{Id}$ n'est pas inversible puisque $P_A(\lambda) = \det(A - \lambda I_2) = 0$. Il existe donc un vecteur à coordonnées complexes $\vec{v} = (z_1, z_2)$ tel que $A\vec{v} = \lambda\vec{v}$. On a $z_1 = x_1 + iy_1$ et $z_2 = x_2 + iy_2$ avec x_i, y_i réels. On peut décomposer le vecteur complexe \vec{v} sous la forme

$$\vec{v} = (x_1, x_2) - i(y_1, y_2) = \vec{v}_1 - i\vec{v}_2$$

avec $\vec{v}_1 = (x_1, y_1)$ et $\vec{v}_2 = (x_2, y_2)$ vecteurs réels.

On a alors

$$\begin{aligned} A\vec{v} &= A\vec{v}_1 - iA\vec{v}_2 = \lambda\vec{v} = (\alpha + i\beta)(\vec{v}_1 - i\vec{v}_2) \\ &= (\alpha\vec{v}_1 + \beta\vec{v}_2) - i(-\beta\vec{v}_1 + \alpha\vec{v}_2). \end{aligned}$$

On a donc, en identifiant les parties réelles et complexes,

$$A\vec{v}_1 = \alpha\vec{v}_1 + \beta\vec{v}_2 \quad \text{et} \quad A\vec{v}_2 = -\beta\vec{v}_1 + \alpha\vec{v}_2.$$

Cela donne le résultat à condition que $B = (\vec{v}_1, \vec{v}_2)$ soit une base de \mathbb{R}^2 . Sinon, \vec{v}_2 et \vec{v}_1 seraient colinéaires dans \mathbb{R}^2 et donc \vec{v} serait aussi multiple de \vec{v}_1 ou \vec{v}_2 , ce qui contredirait le fait que A ne possède pas de vecteur propre réel.

Comme la matrice de f dans B est celle de h_λ , celle de f^n est aussi celle de $(h_\lambda)^n = h_{\lambda^n}$ avec $\lambda^n = |\lambda|^n e^{in\theta}$. On peut aussi dire que si on note

$$S = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = |\lambda| \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

alors $S = P^{-1}AP$ implique que

$$S^n = (P^{-1}AP)(P^{-1}AP) \cdots (P^{-1}AP) = P^{-1}A^n P.$$

□

Le résultat n'est pas évident, mais concrètement les orbites $A^n X(0)$ d'un vecteur $X(0)$ vont décrire des spirales divergentes si $|\lambda| > 1$, convergentes vers $\vec{0}$ si $|\lambda| < 1$, et des ellipses si $|\lambda| = 1$. La rotation est due à θ et la dilatation à $|\lambda|$. On verra des exemples en Python et plus loin.

11.6 Applications aux problèmes d'évolution

Méthode générale.

Dans les applications, on a besoin de calculer des puissances de matrices A^n , ou de résoudre des systèmes différentiels $X'(t) = AX(t)$ ou $X''(t) = AX(t)$, avec des matrices A non diagonales.

On peut résoudre le problème si on sait diagonaliser A . En effet, si on note P la matrice de passage de la base canonique à la base $B = (\vec{u}_1, \vec{u}_2)$ de diagonalisation, on a

$$D = P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \Leftrightarrow A = PDP^{-1},$$

d'où

$$\begin{aligned} A^n &= (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^nP^{-1} \\ &= P \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} P^{-1}. \end{aligned} \quad (17)$$

Le principe est que les calculs deviennent diagonaux lorsque l'on travaille dans la base de diagonalisation. Concrètement, si on utilise $Y(n) = P^{-1}X(n)$ qui donne les coordonnées de $X(n)$ dans la nouvelle base, on a bien

$$Y(n+1) = P^{-1}APY(n) = DY(n).$$

Proposition 11.17. Finalement, on peut retenir que les solutions sont de la forme

$$X(n) = C_1 \lambda_1^n \vec{u}_1 + C_2 \lambda_2^n \vec{u}_2 \quad (18)$$

où C_1 et C_2 sont obtenus par la condition initiale $X(0) = C_1 \vec{u}_1 + C_2 \vec{u}_2$.

Démonstration. C'est exactement ce que fait la formule (17)! On a $Y(0) = P^{-1}X(0) = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ et

$$\begin{aligned} A^n X(0) &= PD^nP^{-1}X(0) = P \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = P \begin{pmatrix} C_1 \lambda_1^n \\ C_2 \lambda_2^n \end{pmatrix} \\ &= C_1 \lambda_1^n \vec{u}_1 + C_2 \lambda_2^n \vec{u}_2, \end{aligned}$$

puisque par définition la matrice de passage P a pour colonnes \vec{u}_1 et \vec{u}_2 . □

Systèmes différentiels. De la même façon, pour résoudre les systèmes différentiels

$$X'(t) = AX(t) \quad \text{ou} \quad X''(t) = AX(t),$$

on utilise

$$Y(t) = P^{-1}X(t).$$

$Y(t)$ donne les coordonnées de $X(t)$ dans la nouvelle base $B = (\vec{u}_1, \vec{u}_2)$. On a

$$\begin{aligned} Y'(t) &= P^{-1}X'(t) = P^{-1}AX(t) = P^{-1}APY(t) \\ &= DY(t). \end{aligned}$$

Le système différentiel devient donc diagonal dans la base B . On a

$$Y(t) = (e^{t\lambda_1}y_1, e^{t\lambda_2}y_2)$$

dans la base B d'après le paragraphe 11.2.

La solution générale de $X'(t) = AX(t)$ dans la base canonique est de la forme

$$X(t) = PY(t) = y_1 e^{t\lambda_1} \vec{u}_1 + y_2 e^{t\lambda_2} \vec{u}_2, \quad (19)$$

où y_1 et y_2 sont définis par la condition initiale $X(0) = y_1 \vec{u}_1 + y_2 \vec{u}_2$.

- Dans le cas de $X''(t) = AX(t)$ on obtient $Y''(t) = DY(t)$, c'est-à-dire

$$y_1''(t) = \lambda_1 y_1(t) \quad \text{et} \quad y_2''(t) = \lambda_2 y_2(t),$$

pour $Y(t) = (y_1(t), y_2(t))$. On obtient alors

$$y_1(t) = a_1 e^{t\sqrt{\lambda_1}} + b_1 e^{-t\sqrt{\lambda_1}} \quad \text{et} \quad y_2(t) = a_2 e^{t\sqrt{\lambda_2}} + b_2 e^{-t\sqrt{\lambda_2}}, \quad (20)$$

puis $X(t) = PY(t)$. Les quatre inconnues a_1, b_1, a_2, b_2 sont déterminées ici par les données de $X(0)$ et $X'(0)$ (position et vitesse).

Attention ! Si λ_1 (ou λ_2) est < 0 , les racines $\pm\sqrt{\lambda_1}$ sont des nombres imaginaires $\pm i\sqrt{|\lambda_1|}$. Les solutions complexes sont de la forme

$$y_1(t) = a_1 e^{i\omega_1 t} + a_2 e^{-i\omega_1 t} \quad \text{avec} \quad \omega_1 = \sqrt{|\lambda_1|}.$$

Les solutions réelles sont les parties réelles de ces solutions. Elles sont oscillantes de la forme

$$y_1(t) = a \cos(\omega_1 t) + b \sin(\omega_1 t) = r \cos(\omega_1 t + \varphi). \quad (21)$$

C'est le cas que l'on rencontre dans les problèmes mécaniques avec des forces de rappel élastiques.

Retour sur la suite de Fibonacci.

On applique la technique à la suite de Fibonacci $u_{n+2} = u_{n+1} + u_n$ avec $u_0 = 0$ et $u_1 = 1$ du début de chapitre 11.1. On a ici

$$X(n) = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} \quad \text{et} \quad X(n+1) = AX(n) \quad \text{avec} \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de A est

$$P_A(\lambda) = \det(A - \lambda I_2) = \begin{vmatrix} -\lambda & 1 \\ 1 & 1 - \lambda \end{vmatrix} = \lambda^2 - \lambda - 1.$$

Le discriminant est $\Delta = 5 > 0$ donc la matrice est diagonalisable par le théorème 11.15, avec des valeurs propres

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{et} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Les vecteur propres de A satisfont

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} y = \lambda x \\ x + y = \lambda y \end{cases} \Leftrightarrow \begin{cases} y - \lambda x = 0 \\ (1 + \lambda - \lambda^2)x = 0 \end{cases} \quad (22)$$

La seconde équation est bien satisfaite lorsque λ est valeur propre et on a une base de vecteurs propres en prenant $\vec{u}_1 = (1, \lambda_1)$ et $\vec{u}_2 = (1, \lambda_2)$. D'après la proposition 11.17, on a

$$X(n) = C_1 \lambda_1^n \vec{u}_1 + C_2 \lambda_2^n \vec{u}_2$$

avec

$$X(0) = (0, 1) = C_1 \vec{u}_1 + C_2 \vec{u}_2 \Leftrightarrow C_1 = \frac{1}{\sqrt{5}} = -C_2.$$

Et finalement, en prenant la première composante de $X(n)$, on a calculé la suite de Fibonacci

$$u_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

Ce n'était pas évident ! C'est bien une suite d'entiers (malgré les apparences). Comme $\lambda_2 = \frac{1 - \sqrt{5}}{2} \approx -0,61$, on a que λ_2^n tend très vite vers 0 lorsque n tend vers $+\infty$. Finalement, à un très petit nombre près, on a

$$u_n \approx \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n.$$

Et notre population de lapins s'étend exponentiellement avec une croissance

$$\frac{u_{n+1}}{u_n} \rightarrow \lambda_1 = \frac{1 + \sqrt{5}}{2}.$$

Remarque 11.18. Le nombre réel $\lambda_1 = \frac{1 + \sqrt{5}}{2}$ est célèbre ! Il s'appelle le *nombre d'or*. Il est présent dans la nature, et cette « divine proportion » a été utilisée en peinture et architecture : voir par exemple Wikipedia nombre d'or.

Remarque sur les suites récurrentes générales.

La méthode présentée s'applique à (presque) toutes les suites récurrentes de la forme

$$u_{n+2} = au_{n+1} + bu_n$$

avec a et b réels (ou même complexes). En effet, en posant $X(n) = (u_n, u_{n+1})$ elles se modélisent de la même façon $X(n+1) = AX(n)$ avec

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

Le polynôme caractéristique est $P_A(\lambda) = \lambda^2 - a\lambda - b$, et l'équation des valeurs propres

$$P_A(\lambda) = \lambda^2 - a\lambda - b = 0$$

s'appelle *l'équation caractéristique* de la suite. Si le discriminant $\Delta = a^2 - 4b \neq 0$, on trouve deux valeurs propres λ_1 et λ_2 , et des vecteurs propres $\vec{u}_1 = (1, \lambda_1)$ et $\vec{u}_2 = (1, \lambda_2)$ comme dans (22). En conclusion, les solutions sont aussi de la forme

$$u_n = C_1 \lambda_1^n + C_2 \lambda_2^n,$$

avec C_1 et C_2 calculés par les valeurs initiales u_0 et u_1 .

À vous de jouer !

Je laisse la résolution des problèmes de Bip bip et les coyotes et du couple glucose-insuline en exercices, ou à illustrer en Python. Il s'agit de calculer les valeurs propres des matrices pour savoir dans quelle situation on est en fonction des valeurs obtenues. A-t-on à faire à des systèmes stables, instables, oscillants ? Les coucous-coueurs et les coyottes peuvent-ils coexister ? Cela dépend-t-il des conditions initiales ?

Les oscillateurs couplés.

On va revenir pour finir sur l'exemple des oscillateurs couplés. On doit étudier d'après l'équation (15) le système différentiel

$$X''(t) = AX(t) \quad \text{avec} \quad A = \begin{pmatrix} -\omega_0^2 - \omega^2 & \omega^2 \\ \omega^2 & -\omega_0^2 - \omega^2 \end{pmatrix}.$$

Le polynôme caractéristique de A est

$$\begin{aligned} P_A(\lambda) &= \lambda^2 - (\text{Tr } A)\lambda + \det A \\ &= \lambda^2 - 2(\omega_0^2 + \omega^2)\lambda + (\omega_0^2 + \omega^2)^2 - \omega^4. \end{aligned}$$

Son discriminant est

$$\Delta = 4(\omega_0^2 + \omega^2)^2 - 4((\omega_0^2 + \omega^2)^2 - \omega^4) = 4\omega^4.$$

- Le discriminant est nul pour $\omega = 0$, ce qui correspond à un système de deux masses découplées oscillant à la même fréquence $\omega_0/2\pi$, voir paragraphe 11.1.

- Si $\omega \neq 0$, la matrice A possède deux valeurs propres distinctes

$$\lambda_1 = -\omega_0^2 \quad \text{et} \quad \lambda_2 = -\omega_0^2 - 2\omega^2.$$

Comme ces valeurs propres sont négatives, les solutions seront oscillantes d'après l'équation (21) et de la forme

$$X(t) = r_1 \cos(\omega_0 t + \varphi_1) \vec{u}_1 + r_2 \cos(\omega_2 t + \varphi_2) \vec{u}_2$$

avec $\omega_2 = \sqrt{|\lambda_2|} = \sqrt{\omega_0^2 + 2\omega^2}$, et \vec{u}_1, \vec{u}_2 les vecteurs propres correspondants de A .

On trouve facilement que $\vec{u}_1 = (1, 1)$, ce qui correspond à une oscillation simultanée des masses, avec $x_1(t) = x_2(t)$, tandis que $\vec{u}_2 = (1, -1)$ correspond à des oscillations en opposition de phase des masses : $x_1(t) = -x_2(t)$.

Un mouvement général (non périodique !) est une superposition de ces deux mouvements.

Les deux fréquences $\omega_0/2\pi$ et $\omega_2/2\pi$ s'appellent les fréquences propres ou *fréquences de résonance* du système. Vous verrez en physique pourquoi ! Les mouvements correspondants s'appellent des modes propres.