Algèbre linéaire propédeutique
— version abrégée —

Ioan Manolescu

November 6, 2017

Avant-Propos

Ces notes accompagnent le cours d'algèbre linéaire propédeutique donné au semestre d'automne des années 2017-2018 à l'Université de Fribourg.

Le but du cours est double. Il vise principalement de familiariser les étudiants avec l'algèbre linaire, en accentuant les aspects pratiques comme le calcul matriciel, la résolution des systèmes linéaires et les notions de valeur et de vecteur propres. Un deuxième but est de faire découvrir aux étudiants le fonctionnement des mathématiques, à savoir la construction d'une théorie à partir d'axiomes. Ce processus vise à créer un cadre abstrait qui permet de décrire des problèmes issus de la réalité. Trouver un cadre abstrait, général, plutôt que d'en créer un pour chaque problème rencontré, cela nous permet d'apercevoir des connexions entre différents problèmes et d'en offrir des solutions plus robustes.

Enfin, dans le développement de toute science, arrive un moment où, pour dépasser les approches ad-hoc, il est nécessaire de construire un cadre théorique qui permet une étude systématique. Les sciences plus "mathématiques" comme la physique ou l'informatique travaillent déjà dans un tel cadre. Dans les sciences traditionnellement plus empiriques (biologie, médecine, économie etc.) ce cadre est en train de se former par les travaux de modélisation de plus en plus fréquentes. Ainsi, il est essentiel pour les scientifiques en formation de se confronter à une construction abstraite, comme celle présentée dans ce cours.

Cette version du polycopié résume les définitions et les résultats essentiels du cours; elle correspond à ce qui est présenté en classe. Ainsi, ce sont uniquement les preuves simples et illustratrices qui y sont présentées. Ceux qui sont intéressés par les preuves plus complexes peuvent consulter la version complète du polycopié. Les exercices sont destinés à offrir une compréhension plus profonde du cours et sont facultatifs. Ils sont différents des exercices des séries qui consistent surtout en exemples concrets.

Ce polycopié n'est pas censé remplacer le cours donnée en classe; il se veut plutôt un complément qui permet aux étudiants de revoir certains points.

Contents

Chapter 1

Espaces vectoriels

Alors que la notion de nombre nous semble très naturelle, celle de vecteur peut nous paraitre artificielle. Pourtant, la nécessité d'additionner des objets plus complexes que des simples nombres apparait naturellement quand on essaye de modéliser mêmes des phénomènes simples. Les vecteurs (et la notion d'espace vectoriel) offre le cadre pour ce type d'opérations.

En physique, on travail avec des forces (ou encore avec des vitesses, accélérations, etc.) qu'on peut additionner pour obtenir une force totale. Une telle force n'est pas un simple nombre, elles est formé d'un nombre (le module) accompagné d'une direction spatiale; celle-ci ayant une importance cruciale pour comprendre l'effet de la force. En effet, la somme de deux forces de module 1 dépend fortement de leur directions. Suivant l'angle entre les directions, la force totale peut avoir un module compris entre 0 et 2.

Un autre point de vu est le suivant: imaginons une situations ou plusieurs ressources non-interchangeables sont disponibles. Dans cette situation, un avoir n'est pas un simple nombre, mais une collection de nombres (un pour chaque ressource). Les avoirs de plusieurs individus, comme les forces en physique, peuvent être additionnés pour obtenir un avoir total, lui aussi représenté par une collection de nombres.

Dans les deux exemples, on rencontre des quantités représentées par plus qu'un simple nombre, mais qu'on peut additionner avec des règles similaires à celles qu'on utilise pour additionner des nombres. De plus, dans les deux cas, on peut aussi multiplier les quantités en questions par des nombres. Ces deux opérations (et les règles qui les accompagnes) sont les spécificités des espaces vectoriels.

En plus de leur intérêt comme outil de modélisation, les espaces vectoriels (avec la théorie afférente de l'algèbre linéaire) offre un cadre pour donner de solutions simples et élégantes à différents problèmes pratiques (voir les exemples à venir).

1.1 Définitions et exemples

On va étudier ici seulement les espaces vectoriels sur le corps des nombres réels $\mathbb{R}$. On va souvent appeler les nombres réels des *scalaires*.

Définition 1.1. *Soient* E *un ensemble non-vide et* $+ : E^2 \to E$ *et* $\cdot : \mathbb{R} \times E \to E$ *des* *opérations (c.-à-d. des fonctions). On dit que E est un* espace vectoriel[i] *(e.v.) si les propriétés suivantes sont satisfaites pour tous éléments $x, y, z \in E$ et $\lambda, \mu \in \mathbb{R}$:*

(i) $x + y = y + x$ *(commutativité de +);*

(ii) $(x + y) + z = x + (y + z)$ *(associativité de +);*

(iii) il existe un élément $0_E \in E$ tel que $u + 0_E = u$ pour tout $u \in E$;

(iv) pour chaque $u \in E$, il existe un élément $-u \in E$ tel que $u + (-u) = 0_E$;

(v) $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$ *(associativité de ·);*

(vi) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$ *(distributivité pour +);*

(vii) $\lambda(x + y) = \lambda x + \lambda y$ *(distributivité pour +);*

(viii) $1 \cdot x = x$.

Les éléments de E sont appelés vecteurs[ii]*.*

Les propriétés qui sont énumérées dans la définition d'un espace vectoriel ne sont rien d'autre que des règles naturelles de calcul. Il s'avère que ce sont les bonne règles à imposer pour avoir une structure d'espace vectoriel qui correspond à notre intuition. De plus, les fait suivants (qui sont tout aussi naturels) en découlent:

(a) L'élément 0_E est unique et on l'appelle *l'élément neutre*[iii]pour l'addition.
(b) $0 \cdot x = 0_E$, pour tout $x \in E$;
(c) $\lambda \cdot 0_E = 0_E$, pour tout $\lambda \in \mathbb{R}$;
(d) si $\lambda \cdot x = 0_E$, alors $\lambda = 0$ ou $x = 0_E$.
(e) Pour chaque $x \in E$, l'élément $-x$ qui satisfait la propriété (iv) est unique.
(f) $(-1) \cdot x = -x$, pour tout $x \in E$;

Remarque 1.2. Alors que les conditions de la définition 1.1, ainsi que leur conséquences peuvent sembler trop nombreuses et compliqués à retenir, ce dont on doit se souvenir est:

Dans un espace vectoriel E on peut faire des sommes de vecteurs et on peut multiplier les vecteurs par des nombres (qu'on appelle des scalaires). Les règles usuelles d'addition et multiplication s'y appliquent.

Attention: Ce qu'on ne peut pas faire (pour $\lambda \in \mathbb{R}$ et $x, y \in E$) est

- faire la somme entre un scalaire et un vecteur (ne jamais écrire $\lambda + x$);
- écrire un produit avec le scalaire à droite (ne pas écrire $x\lambda$);
- faire un produit de deux vecteurs. Pour l'instant $x \cdot y$ n'as aucun sens.

Désormais, quand aucune ambiguïté est possible, on supprime l'indice E de 0_E.

[i] Vektorraum
[ii] Vektoren

1.2 Sous-espaces vectoriels

Définition 1.3. *Soit $(E, +, \cdot)$ un espace vectoriel. Un ensemble $F \subset E$ est appelé* sous-espace vectoriel[iv] *de E si $(F, +, \cdot)$ est un espace vectoriel.*

Attention! L'addition et la multiplication par une constante utilisées pour F *sont celles de l'espace vectoriel $(E, +, \cdot)$.*

Proposition 1.4. *Soit $(E, +, \cdot)$ un espace vectoriel et $F \subset E$ un sous-ensemble non-vide de E. Alors F est un sous-espace vectoriel de E si et seulement si les deux conditions suivantes sont vérifiées:*

(a) pour tout $x, y \in F$, $x + y \in F$,

(b) pour tout $x \in F$ et $\lambda \in \mathbb{R}$, $\lambda x \in F$.

De plus, les deux conditions peuvent s'écrire de façon plus compacte comme:

$$\text{pour tout } x, y \in F \text{ et } \lambda, \mu \in \mathbb{R}, \qquad \lambda x + \mu y \in F. \tag{1.1}$$

Preuve: On ne donne pas une preuve complète de ce résultat. On mentionne juste que les propriétés (i - viii) pour F suivent des proriétés correspondantes pour E. Ce qui est nécessaire (et non trivial) pour que F soit un espace vectoriel est que les opérations $+$ et $\cdot$ soient en effet définies sur F. Cela correspond aux points (a) et (b) de la proposition 1.4, respectivement.

La condition (1.1) est simplement une écriture plus compacte des conditions (a) et (b). □

Remarque: si F est un sous-espace vectoriel d'un espace vectoriel E, alors $0_E \in F$ (voir la preuve de la proposition pour l'explication de ce fait).

Exemples On donne ici quelques exemples d'espaces et sous-espaces vectoriels qu'on va reprendre à plusieurs reprises plus tard dans le cours.

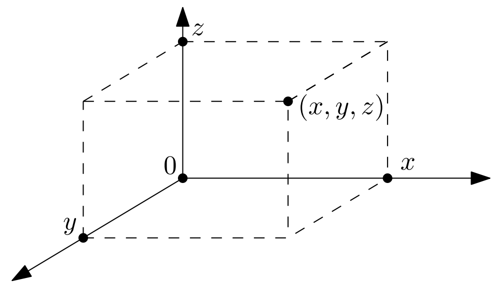
1. L'exemple le plus simple d'espace vectoriel qui vient à l'esprit est $\mathbb{R}^d$ pour $d \geq 1$. L'ensemble $\mathbb{R}^d$ est l'ensemble des d-uplets de nombre réels. Plus précisément, un élément générique de $\mathbb{R}^d$ s'écrit $(x_1, \dots, x_d)$ avec $x_1, \dots, x_d \in \mathbb{R}$.

   L'addition et la multiplication par un scalaire sont définies comme suit. Pour $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ et $\lambda \in \mathbb{R}$,

   $$x + y := (x_1 + y_1, \dots, x_d + y_d) \quad \text{et} \quad \lambda x := (\lambda x_1, \dots, \lambda x_d).$$

   L'espace $\mathbb{R}^3$ s'identifie à l'espace de dimension 3 avec un système de coordonnées. En effet, un point $(x, y, z) \in \mathbb{R}^3$ correspond au point de coordonnées x, y et z. Voir image:

[iv] Untervektorraum



2. L'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels, munit de l'addition standard et de la multiplication par une constante standard est un $\mathbb{R}$-espace vectoriel.
3. Soit X un ensemble quelconque et $\mathcal{F}(X, \mathbb{R})$ l'ensemble des fonctions de X dans $\mathbb{R}$. L'addition dans $\mathcal{F}(X, \mathbb{R})$ est définie composante par composante. La multiplication par un scalaire aussi. Ainsi, pour $f, g \in \mathcal{F}(X, \mathbb{R})$ et $\lambda \in \mathbb{R}$, les fonctions $f + g : X \to \mathbb{R}$ et $\lambda \cdot f : X \to \mathbb{R}$ sont définies par
$$(f+g)(x) = f(x) + g(x) \quad \text{et} \quad (\lambda \cdot f)(x) = \lambda f(x), \quad \text{pour tout } x \in X.$$
Muni de ces deux opérations, $\mathcal{F}(X, \mathbb{R})$ est un espace vectoriel.
4. Un cas particulier du point précèdent est l'ensemble des suites $\ell(\mathbb{R}) = \mathbb{R}^{\mathbb{N}} := \mathcal{F}(\mathbb{N}, \mathbb{R})$. Une suite réelle, notée $(a_n)_{n \in \mathbb{N}}$ est une famille de nombre réels indexée par les nombres naturels. L'addition et la multiplication par un scalaire se font composante par composante, comme pour les fonctions.
5. Les suites définies par une relation de récurrence forment un sous-espace vectoriel de $\ell(\mathbb{R})$ (voir section 1.4).
6. L'ensemble des solutions d'un système linéaire homogène forment un espace vectoriel (voir la partie 2.4). Considérons par exemple le système suivant
$$\begin{cases} x + y + 2z &= 0 \\ x - 2y - z &= 0, \end{cases} \tag{1.2}$$
à deux équations et trois inconnues $x, y, z \in \mathbb{R}$. L'ensemble $\mathcal{S}$ des solutions (x, y, z) est un sous-espace vectoriel de $\mathbb{R}^3$. En effet, si (x, y, z) et (x', y', z') sont des solutions de (1.2), alors $(x, y, z) + (x', y', z') := (x + x', y + y', z + z')$ en est une aussi. De plus, pour $\lambda \in \mathbb{R}$, le triplet $\lambda(x, y, z) := (\lambda x, \lambda y, \lambda z)$ est aussi solution de (1.2).

   On peut même calculer $\mathcal{S}$ et donner la forme générale de ses éléments. En effet, en soustrayant la première ligne à la deuxième, le système (1.2) devient
$$\begin{cases} x + y + 2z &= 0 \\ -3y + 3z &= 0, \end{cases}$$
   On divise la seconde équation par -3 pour obtenir:
$$\begin{cases} x + y + 2z &= 0 \\ y + z &= 0, \end{cases}$$

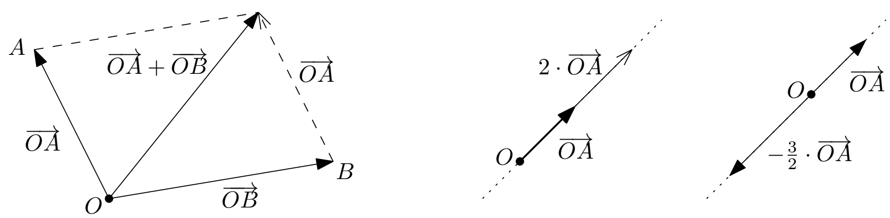


Figure 1.1: L'addition des vecteurs peut se faire par la règle du parallélogramme. Dans l'image de gauche, les cotes opposées du parallélogramme sont *le même vecteur*. La multiplication par une constante se fait en gardant la même direction, mais en multipliant la longueur du vecteur. Si la constant est négative, la flèche change de sens.

En fin, on soustrait deux fois la seconde équation à la première et on obtient

$$\begin{cases} x+z &= 0 \\ y+z &= 0, \end{cases}$$

Les opérations qu'on vient de faire ne change pas l'ensemble de solutions du système. Ainsi, les solutions sont $\mathcal{S} = \{(z, z, -z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$.

Géométriquement, $\mathcal{S}$ est une droite de $\mathbb{R}^3$, plus précisément c'est la droite qui passe par les points $(0,0,0)$ et $(1,1,-1)$. On verra par la suite que les sous-espaces vectoriels (non-triviaux) de $\mathbb{R}^3$ sont les droites et les plans passant par l'origine $(0,0,0)$. Ce premier exemple témoigne de la nature géométrique des espaces vectoriels.

7. En physique les forces, vitesses, accélérations etc. sont représentées par des flèches dans l'espace usuel à trois dimensions. On on appelle ces flèches des vecteurs. Une flèche entre deux points O et A est notée $\overrightarrow{OA}$. On peut additionner les flèches par la règle du parallélogramme. De plus, pour un scalaire $\lambda \in \mathbb{R}$, on peut multiplier la flèche $\overrightarrow{OA}$ par le scalaire $\lambda \in \mathbb{R}$. En effet $\lambda\overrightarrow{OA}$ est une flèche de longueur $|\lambda| \cdot |\overrightarrow{OA}|$, de même direction que $\overrightarrow{OA}$ et de même sens si $\lambda \geq 0$ ou de sens opposé si $\lambda < 0$.

   Fixons le point O et notons G_O^3 l'ensemble des flèches d'origine O.

   Le nom de vecteur qu'on donne à ces flèches n'est pas accidentel. En effet, G_O^3 est un espace vectoriel (les conditions (i-viii) sont facilement vérifiées).

   Une flèche $\overrightarrow{AB}$ dont le point d'origine n'est pas O est identifiée à la flèche parallèle, de même longueur et même sens, mais dont l'origine est bien O (en autre mots on considére que $\overrightarrow{AB}$ est *la même* flèche que sont translaté). Ainsi, en pratique on peut se limiter à l'espace vectoriel G_O^3.

Exercice 1.1.
Prenons $E = \mathbb{R}^3$.

(a) Soit $F = \{(x, y, z) \in \mathbb{R}^3 : z = 0\} \subset E$. Montrer que F est un s.e.v. de E.
(b) Soit $G = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \subset E$. Montrer que G est un s.e.v. de E.
(c) Soit $H = \{(x, y, z) \in \mathbb{R}^3 : x+y+z = 1\} \subset E$. Montrer que H n'est pas un s.e.v. de E.

Exercice 1.2.
Prenons $E = \mathbb{R}[X]$.

(a) Soit $F = \{P \in \mathbb{R}[X] : P(0) = 0\} \subset E$. Montrer que F est un s.e.v. de E.
(b) Soit $G = \{P \in \mathbb{R}[X] : P(0) = 1\} \subset E$. Montrer que G n'est pas un s.e.v. de E.
(c) Soit $H = \{P \in \mathbb{R}[X] : \deg P \leq 4\} \subset E$. Montrer que H est un s.e.v. de E.
(d) Soit $I = \{P \in \mathbb{R}[X] : \deg P = 4\} \subset E$. Montrer que I n'est pas un s.e.v. de E.

1.3 Base, dimension finie, dimension

1.3.1 Familles libres, génératrices

Fixons pour le reste du chapitre un espace vectoriel E. Dans cette partie on va considérer à plusieurs reprises des familles finies de vecteurs $x_1, \dots, x_n \in E$. Pour une telle famille, on appelle *combinaison linéaire*(v) de $x_1, \dots, x_n$ tout vecteur de la forme

$$\lambda_1 x_1 + \cdots + \lambda_n x_n,$$

ou $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ sont des scalaires quelconques.

Définition 1.5. *Soit $x_1, \dots, x_n \in E$ une famille finie de vecteurs. Le sous-espace vectoriel de E engendré par $(x_1, \dots, x_n)$*(vi) *est*

$$\mathrm{Vect}(x_1, \dots, x_n) := \{\lambda_1 x_1 + \cdots + \lambda_n x_n : \lambda_1, \dots, \lambda_n \in \mathbb{R}\}.$$

On dit que la famille $(x_1, \dots, x_n)$ est génératrice *(pour E) si* $\mathrm{Vect}(x_1, \dots, x_n) = E$.

En autre mots, $\mathrm{Vect}(x_1, \dots, x_n)$ est l'ensemble formé de toutes les combinaisons linéaires de $x_1, \dots, x_n$. La famille $(x_1, \dots, x_n)$ est dite génératrice si tout vecteur de E peut s'écrire comme combinaison linéaire de $x_1, \dots, x_n$.

L'ordre des vecteurs $x_1, \dots, x_n$ n'affecte pas l'espace qu'ils engendre, donc ni le fait que la famille soit génératrice ou pas.

Remarque 1.6. La définition de $\mathrm{Vect}(x_1, \dots, x_n)$ ci-dessus nécessite une preuve. En effet $\mathrm{Vect}(x_1, \dots, x_n)$ est défini comme un *sous-ensemble* de E, non-pas comme un sous-espace vectoriel. Il faut ainsi montrer (en utilisant la proposition 1.4 par exemple) que $\mathrm{Vect}(x_1, \dots, x_n)$ est un effet un sous-espace vectoriel de E.

Preuve du fait que $\mathrm{Vect}(x_1, \dots, x_n)$ **est un s.e.v.:** Prenons $x_1, \dots, x_n \in E$ une famille finie de vecteurs. Alors, $0_E = 0 \cdot x_1 + \cdots + 0 \cdot x_n \in \mathrm{Vect}(x_1, \dots, x_n)$. En particulier $\mathrm{Vect}(x_1, \dots, x_n)$ n'est pas vide.

Soit $u, v \in \mathrm{Vect}(x_1, \dots, x_n)$. Alors on peut écrire

$$u = \lambda_1 x_1 + \cdots + \lambda_n x_n \quad \text{et} \quad v = \mu_1 x_1 + \cdots + \mu_n x_n,$$

(v) Linearkombination
(vi) lineare Hülle von $(x_1, \dots, x_n)$, aussi notée $\mathrm{span}(x_1, \dots, x_n)$

pour des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ et $\mu_1, \dots, \mu_n \in \mathbb{R}$. Ainsi, pour tout $\nu \in \mathbb{R}$:

$$\begin{aligned} u + v &= (\lambda_1 + \mu_1) \cdot x_1 + \cdots + (\lambda_n + \mu_n) \cdot x_n \in \mathrm{Vect}(x_1, \dots, x_n) \quad \text{et} \\ \nu u &= \nu\lambda_1 \cdot x_1 + \cdots + \nu\lambda_n \cdot x_n \in \mathrm{Vect}(x_1, \dots, x_n). \end{aligned}$$

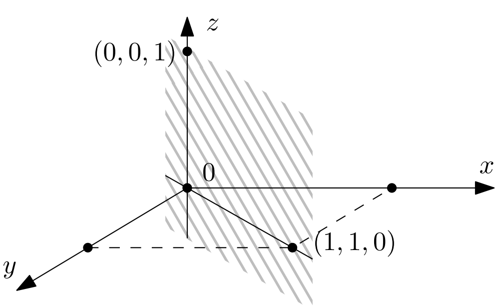
Par la proposition 1.4, on conclut que $\mathrm{Vect}(x_1, \dots, x_n)$ est un sous-espace vectoriel de E. □

Exemple: Dans l'espace vectoriel $\mathbb{R}^2$ la famille $((0,1);(1,0))$ est génératrice. En effet, tout vecteur $(a,b) \in \mathbb{R}^2$ s'écrit $(a,b) = b \cdot (0,1) + a \cdot (1,0)$.

Dans l'espace vectoriel $\mathbb{R}^3$, la famille $((1,1,0);(0,0,1))$ n'est pas génératrice. En effet, le vecteur $(0,1,0)$ ne peut pas s'écrire comme $\alpha \cdot (1,0,0) + \beta \cdot (0,1,1)$ avec $\alpha, \beta \in \mathbb{R}$.
On peut voir que

$$\mathrm{Vect}((1,1,0);(0,0,1)) = \{(a,b,c) \in \mathbb{R}^3 : \; a = b\}.$$

On peut aussi facilement se convaincre que $\mathrm{Vect}((1,1,0);(0,0,1))$ est un s.e.v. de $\mathbb{R}^3$. Il s'agit du plan vertical qui passe par la diagonale $x = y$ du plan horizontal. Voir l'image.



Définition 1.7. *Soit $x_1, \dots, x_n \in E$ une famille finie de vecteurs. On dit que la famille $(x_1, \dots, x_n)$ est* liée *(ou linéairement dépendante)*[vii] *s'il existe des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ non tous nuls, tels que*

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0.$$

Une famille qui n'est pas liée est dite libre *(ou linéairement indépendante)*[viii].

En autre mots, une famille est liée s'il existe une combinaison linéaire non-triviale de ses vecteurs qui vaut 0. Bien évidement, le fait qu'une famille de vecteurs est libre ou liée ne dépend pas de l'ordre des vecteur.

Si on désire montrer qu'une famille $(x_1, \dots, x_n)$ de vecteurs est libre, on peut considérer des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$ et déduire que $\lambda_1 = \cdots = \lambda_n = 0$. Inversement, si on veut montrer que la famille est liée, le plus simple est d'exhiber une combinaison linéaire non-triviale de $x_1, \dots, x_n$ qui vaut 0.

(vii) linear abhängig
(viii) linear unabhängig

Exemple: Dans l'espace vectoriel $\mathbb{R}^2$ la famille $((0,1);(1,0))$ est libre. En effet, si $\lambda, \mu \in \mathbb{R}$ sont tels que $\lambda \cdot (0,1) + \mu \cdot (1,0) = (\lambda, \mu) = (0,0)$, alors $\lambda = \nu = 0$.

Par contre, la famille $((0,1);(1,0);(1,1))$ est liée. En effet, on a

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

Remarque 1.8. Le fait qu'une famille $(x_1, \ldots, x_n)$ de vecteurs d'un espace vectoriel E est libre est quelque chose qui dépend uniquement de $x_1, \ldots, x_n$ pas de l'espace ambiant E. Plus précisément, si F est un s.e.v. de E et $x_1, \ldots, x_n \in F$, alors le fait que famille $(x_1, \ldots, x_n)$ est libre ne dépend pas de si on la considère comme famille de E ou de F. A l'opposé, le fait qu'une famille est génératrice dépend essentiellement de l'espace global E. En effet, même si $(x_1, \ldots, x_n)$ n'est pas génératrice pour E, elle est toujours génératrice pour $\text{Vect}(x_1, \ldots, x_n)$.

Ainsi, dit tout simplement qu'une famille est libre ou liée, mais insiste souvent qu'une famille est génératrice *pour un espace E.*

Le lemme suivant offre un critère pratique pour montrer qu'une famille de vecteurs est liée.

Lemme 1.9. *Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E. Alors elle est liée si et seulement si il existe un vecteur x_i qui s'écrit comme combinaison linéaire des autres. Plus précisément si et seulement si il existe $i \in [1, \ldots, n]$ et des scalaires $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n$ tels que*

$$x_i = \lambda_1 x_1 + \cdots + \lambda_{i-1} x_{i-1} + \lambda_{i+1} x_{i-1} \cdots + \lambda_n x_n = \sum_{j \neq i} \lambda_j x_j.$$

Exemple: En général le fait qu'une famille est libre ou liée n'implique pas le fait qu'elle est génératrice ou pas. Donnons des exemples pour les quatre situations possibles; on va se placer dans l'espace vectoriel $\mathbb{R}^2$:
La famille $(1,0);(0,1)$ est génératrice et libre.
La famille $(1,0);(1,1);(0,1)$ est génératrice et liée (car $(1,1) = (1,0) + (0,1)$).
La famille formée uniquement du vecteur $(1,0)$ est libre mais pas génératrice.
La famille $(1,0);(0,0);(2,0)$ n'est pas génératrice et est liée.

Exercice 1.3.
Démontrer le lemme 1.9

Exercice 1.4.
Montrer que dans un espace vectoriel E:

(a) Toute famille de vecteurs contentant le vecteur 0_E est liée.

(b) Toute famille de vecteurs contentant deux fois le même vecteur est liée.

(c) Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E, $i \in [1, n]$ et $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Alors les affirmations suivantes sont équivalentes

   (i) $(x_1, \ldots, x_n)$ est libre,

(ii) $(x_1, \dots, x_n + x_i)$ est libre,
(iii) $(x_1, \dots, \lambda x_n)$ est libre.

Exercice 1.5.
Montrer que pour une famille finie de vecteurs $x_1, \dots, x_n \in E$, $\mathrm{Vect}(x_1, \dots, x_n)$ est le plus petit sous-espace vectoriel de E contenant $x_1, \dots, x_n$.
En autres mots, si F est un sous-espace vectoriel de E contenant $x_1, \dots, x_n$, montrer que $\mathrm{Vect}(x_1, \dots, x_n) \subset F$.

Exercice 1.6.
Soit $(x_1, \dots, x_n)$ une famille de vecteurs de E. Montrer que

(i) Si $(x_1, \dots, x_n)$ est libre, alors $(x_1, \dots, x_{n-1})$ est aussi libre;
(ii) Si $(x_1, \dots, x_{n-1})$ est génératrice, alors $(x_1, \dots, x_n)$ est aussi génératrice.
Plus généralement, montrer que $\mathrm{Vect}(x_1, \dots, x_{n-1}) \subset \mathrm{Vect}(x_1, \dots, x_n)$.

Exercice 1.7.
Soit $(x_1, \dots, x_n)$ une famille génératrice de E. Montrer que, si $x_n \in \mathrm{Vect}(x_1, \dots, x_{n-1})$, alors $(x_1, \dots, x_{n-1})$ est aussi génératrice.

1.3.2 Dimension finie; théorèmes essentiels

Définition 1.10. *Un espace vectoriel E est dit de* dimension finie[ix]*, s'il existe une famille finie de vecteurs de E qui soit génératrice pour E.*

Définition 1.11. *Soit E un espace vectoriel. On dit qu'une famille $x_1, \dots, x_n \in E$ est une* base[x] *de E si et seulement si $(x_1, \dots, x_n)$ est libre et génératrice.*

Exemple: Considérons l'espace vectoriel $\mathbb{R}^d$ avec $d \geq 1$ et la famille de vecteurs $\mathcal{E} = (e_1, \dots, e_d)$ ou les composantes de e_i sont toutes 0 à part la i^{eme} qui vaut 1:

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i^{\text{eme}} \text{ ligne} \\ \\ \\ \\ \end{array}$$

[ix] endlich dimensional
[x] Basis

Alors, $\mathcal{E}$ est une base de $\mathbb{R}^d$. En effet
$\mathcal{E}$ **est génératrice** car $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, s'écrit $x = x_1 e_1 + \cdots + x_d e_d$.
$\mathcal{E}$ **est libre** car, si $\lambda_1 x_1 + \cdots + \lambda_d x_d = 0$ avec $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, alors

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 e_1 + \cdots + \lambda_d e_d = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_d \end{pmatrix},$$

donc $\lambda_1 = \cdots = \lambda_d = 0$.

Cette base est souvent appelée la *base canonique* de $\mathbb{R}^d$. Toutefois, cette base n'est pas l'unique base de $\mathbb{R}^d$; pour tout $d \geq 1$, il existe une infinité de bases des de $\mathbb{R}^d$.

Le théorème suivant nous permet de définir la dimension d'un espace vectoriel. Il s'agit d'un théorème difficile, qu'on ne va pas prouvé dans ce cours.

Théorème 1.12. *Soit E un espace vectoriel de dimension finie. Alors:*

(i) *toute famille génératrice de E contient une base:*
si $x_1, \ldots, x_m \in E$ est une famille génératrice, alors il existent $n \leq m$ et des indices $1 \leq i_1 < \cdots < i_n \leq m$ tels que $(x_{i_1}, \ldots, x_{i_n})$ soit une base de E.

(ii) *toute famille libre de E peut être complétée en une base:*
si $x_1, \ldots, x_m \in E$ est une famille libre, alors il existent $n \geq m$ et des vecteurs $x_{m+1}, \ldots, x_n$ tels que $(x_1, \ldots, x_n)$ soit une base de E.

(iii) *toutes les bases de E ont le même nombre de vecteur.*
Ce nombre est appelé la *dimension* de E et est noté $\dim(E)$.

Le théorème précédent est très important. En effet, quand on pense à $\mathbb{R}^d$ on dit naturellement qu'il s'agit d'un espace de dimension d. Cela vient du fait qu'on pense à la base canonique $(e_1, \ldots, e_d)$ qui contient d vecteurs. Il est en effet naturel de définir la dimension comme le nombre de vecteurs d'une base. Pourtant, si on regarde $\mathbb{R}^d$ comme espace vectoriel, il n'y a pas de raison de distinguer la base $(e_1, \ldots, e_d)$ des autres bases. Ainsi, on a besoin de savoir que le nombre de vecteurs des différentes bases d'un espace vectoriel ne dépend pas de la base.

Proposition 1.13. *Soient E un espace vectoriel de dimensions $d < \infty$ et $\mathcal{E} = (e_1, \ldots, e_n)$ une famille de vecteurs de E.*

(i) *Si $\mathcal{E}$ est libre, alors $n \leq d$. De plus, si $n = d$, alors $\mathcal{E}$ est une base de E.*

(ii) *Si $\mathcal{E}$ est génératrice, alors $n \geq d$. De plus, si $n = d$, alors $\mathcal{E}$ est une base de E.*

Preuve: Commençons par le point (i). Supposons que $\mathcal{E} = (e_1, \dots, e_n)$ est libre. Par le théorème 1.12, on peut compléter $(e_1, \dots, e_n)$ en une base $(e_1, \dots, e_{n+m})$ pour un certain $m \geq 0$. Mais le théorème 1.12 indique alors que $m + n$ est la dimension de E, à savoir d. Ainsi $n = d - m \leq d$.

De plus, si $n = d$, alors $m = 0$, donc la famille $(e_1, \dots, e_n)$ est une base.

Passons au point (ii). Supposons que $\mathcal{E} = (e_1, \dots, e_n)$ est génératrice de E. Par le théorème 1.12, $(e_1, \dots, e_n)$ contient une base de E. De plus, par le théorème 1.12, cette base contient exactement d vecteurs. Ainsi $d \leq n$.

De plus, si $d = n$, la base contenue dans $\mathcal{E}$ contient n vecteurs, donc est égale à $\mathcal{E}$. □

Exercice 1.8.
On se place dans l'espace vectoriel $\mathbb{R}^3$. Quelles des familles suivantes de vecteurs sont des bases? Motivez vos réponses

(a) $\big((1,0,0);(0,1,0);(0,0,3)\big)$
(b) $\big((1,1,1);(0,2,0);(1,0,1)\big)$
(c) $\big((1,0,0);(1,1,0);(0,1,0);(0,0,1)\big)$
(d) $\big((1,1,0);(0,0,1)\big)$
(e) $\big((1,1,1);(x,y,z);(x^2,y^2,z^2)\big)$ pour $x, y, z \in \mathbb{R}$ (la réponse peut dépendre de x, y et z).

1.3.3 Sous-espaces vectoriels de dimension finie; rang

Soit F un sous-espace vectoriel d'un espace vectoriel E, ce dernier n'étant pas forcement de dimension finie. Par définition, F est aussi un espace vectoriel, et les définitions de dimension finie et dimension données au-dessus s'y applique également. Il est généralement possible que F soit de dimension finie, alors que E ne l'est pas. L'inverse n'est pas possible, comme l'affirme la proposition suivante.

Proposition 1.14. *Soit E un espace vectoriel de dimension finie et F un sous espace vectoriel de E. Alors F est aussi de dimension finie et* $\dim(F) \leq \dim(E)$.
De plus, si $\dim(F) = \dim(E)$, *alors* $F = E$.

Preuve: On va admettre ce résultat. □

Ainsi, l'espace vectoriel $\mathbb{R}^3$ admet des sous-espaces de dimension $0, 1, 2$ et 3. Les uniques espaces de dimension 0 et 3 sont $\{(0,0,0)\}$ et $\mathbb{R}^3$, respectivement. Les sous-espaces de dimension 1 sont les droites passant par 0; ceux de dimension 2 sont les plans passant par 0.

La terminologie s'étend aux espaces de dimension plus grande: si E est un espace vectoriel de dimension finie d, ses sous-espaces vectoriels de dimension 1 sont appelés des droites et ceux de dimension $d-1$ sont appelés des hyperplans.

Définition 1.15. *Soit $x_1, \dots, x_n \in E$ une famille de vecteurs. Le* rang(xi) *de* $(x_1, \dots, x_n)$

est noté $\operatorname{rang}(x_1,\ldots,x_n)$ *et est défini par* $\operatorname{rang}(x_1,\ldots,x_n) = \dim(\operatorname{Vect}(x_1,\ldots,x_n))$.

Corollaire 1.16. *Soit* $x_1,\ldots,x_n \in E$ *une famille de vecteurs. Alors*

(i) $\operatorname{rang}(x_1,\ldots,x_n) \leq n$. *De plus,* $\operatorname{rang}(x_1,\ldots,x_n) = n$ *si et seulement si la famille est libre.*

(ii) $\operatorname{rang}(x_1,\ldots,x_n) \leq \dim(E)$. *De plus,* $\operatorname{rang}(x_1,\ldots,x_n) = \dim(E)$ *si et seulement si la famille est génératrice pour* E.

Preuve: Soit $x_1,\ldots,x_n \in E$ une famille de vecteurs. Alors, $x_1,\ldots,x_n$ est génératrice pour $\operatorname{Vect}(x_1,\ldots,x_n)$, donc par proposition 1.13, $n \geq \dim\operatorname{Vect}(x_1,\ldots,x_n) = \operatorname{rang}(x_1,\ldots,x_n)$.

En outre, si $n = \operatorname{rang}(x_1,\ldots,x_n)$, alors le proposition 1.13 indique que $(x_1,\ldots,x_n)$ est une base de $\operatorname{Vect}(x_1,\ldots,x_n)$, donc qu'elle est libre.

Inversement, si $(x_1,\ldots,x_n)$ est libre, alors elle est une base de $\operatorname{Vect}(x_1,\ldots,x_n)$ (on a déjà mentionné qu'elle est génératrice pour $\operatorname{Vect}(x_1,\ldots,x_n)$). Ainsi $\operatorname{rang}(x_1,\ldots,x_n) = \dim\operatorname{Vect}(x_1,\ldots,x_n) = n$ $\square$

1.3.4 Représentation d'un vecteur dans une base

Une base d'un espace vectoriel E sert essentiellement à *décomposer* les vecteurs de E dans cette base. Le théorème suivant donne un sens précis à la décomposition d'un vecteur dans une base.

Théorème 1.17. *Soit* E *un espace vectoriel de dimension finie* n *et* $\mathcal{B} = (e_1,\ldots,e_n)$ *une base de* E. *Alors, pour tout* $x \in E$, *il existe une unique famille de scalaires* $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ *telle que*

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n. \tag{1.3}$$

Preuve: Soit $x \in E$. Comme $\mathcal{B}$ est une base de E, $x \in \operatorname{Vect}(\mathcal{B}) = E$. Ainsi, il existe $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ tels que $x = \lambda_1 e_1 + \cdots + \lambda_n e_n$.

Maintenant qu'on a prouvé l'existence de la famille de scalaires désirée, montrons aussi son unicité. Soit $\lambda_1,\ldots,\lambda_n,\mu_1,\ldots,\mu_n \in \mathbb{R}$ tels que

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n = \mu_1 e_1 + \cdots + \mu_n e_n.$$

Alors, $(\lambda_1 - \mu_1)e_1 + \cdots + (\lambda_n - \mu_n)e_n = 0_E$. La famille $e_1,\ldots,e_n$ étant libre, cela implique $\lambda_1 - \mu_1 = \cdots = \lambda_n - \mu_n = 0$. En autre mots, les familles $(\lambda_1,\ldots,\lambda_n)$ et $(\mu_1,\ldots,\mu_n)$ sont égales, ce qui prouve l'unicité. $\square$

(xi) Rang

A chaque $x \in E$ on associe le vecteur $V_{\mathcal{B}}(x) := (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$, ou $\lambda_1, \dots, \lambda_n$ sont les scalaires tels que (1.3) soit vérifié. Pour des raisons qu'on va voir plus tard, on va désormais écrire $V_{\mathcal{B}}(x)$ verticalement, a savoir:

$$V_{\mathcal{B}}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

On appellera le vecteur $V_{\mathcal{B}}(x)$ *l'écriture de x dans la base $\mathcal{B}$*.

Proposition 1.18. *Soit E un espace vectoriel de dimension n et $\mathcal{B}$ une base. Pour $x, y \in E$ et $\lambda \in \mathbb{R}$, notons*

$$V_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{et} \qquad V_{\mathcal{B}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Alors:

$$V_{\mathcal{B}}(x+y) = V_{\mathcal{B}}(x) + V_{\mathcal{B}}(y) := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad \textit{et} \qquad V_{\mathcal{B}}(\lambda x) = \lambda V_{\mathcal{B}}(x) := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

De plus, la fonction $V_{\mathcal{B}} : E \to \mathbb{R}^n$ est une bijection.

Rappel: le fait que $V_{\mathcal{B}}$ est bijective revient à dire que, pour tout vecteur $X \in \mathbb{R}^n$, il existe un unique vecteur $x \in E$ tel que $V_{\mathcal{B}}(x) = X$.

Ainsi, les vecteurs de E et ceux de $\mathbb{R}^n$ sont en correspondance exacte par la fonction $V_{\mathcal{B}}$. De plus cette fonction est cohérente avec les opérations d'espace vectoriel: somme et multiplication par scalaires. On s'aperçoit donc que E et $\mathbb{R}^n$ ont exactement la même structure.

Toutefois, la correspondance entre E et $\mathbb{R}^d$ dépend de la base $\mathcal{B}$ de E qu'on a choisi.

Preuve: On va admettre ce résultat. $\square$

Exemple: Evidement, on peut décomposer un même vecteur dans deux bases différentes. Pour $\mathbb{R}^2$, on dispose de la base canonique $\mathcal{E} = (e_1, e_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$. Mais on peut également considérer la base $\mathcal{F} = (f_1, f_2)$ ou

$$f_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{et} \qquad f_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

(On vérifie facilement que $\mathcal{F}$ est en effet une base de $\mathbb{R}^2$.) Le vecteur $x = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ se décompose alors dans les bases $\mathcal{E}$ et $\mathcal{F}$ comme suit

$$\begin{pmatrix} 3 \\ 7 \end{pmatrix} = 3e_1 + 7e_2 = 5f_1 - 2f_2.$$

Ainsi

$$V_{\mathcal{E}}(x) = \begin{pmatrix} 3 \\ 7 \end{pmatrix} \quad \text{et} \quad V_{\mathcal{F}}(x) = \begin{pmatrix} 5 \\ -2 \end{pmatrix}.$$

1.4 Application: suites définies par récurrence

Le modèle de Fibonacci (1202) Fibonacci a proposé un modèle pour décrire le développement d'une population de lapins. Il admet les hypothèses suivantes:

- Les lapins sont immortels et vivent fidèlement en couples qui sont formés à la naissance. Les partenaires d'un couple ont donc le même âge et nous parlerons de l'âge du couple.
- Dès l'âge de 2 mois chaque couple donne naissance, chaque mois, à un nouveau couple.
- le temps t est discret, l'unité étant le mois. A $t = 0$, la population consiste d'un unique couple de nouveau-nés.

Le problème est de trouver une formule simple pour le nombre de couples après t mois.

Notons u_t le nombre de couples à temps t. Alors on a

$$u_{t+2} = u_{t+1} + u_t, \quad \text{pour tout } t \geq 0 \text{ et} \tag{1.4}$$

$$u_0 = 1,\ u_1 = 1. \tag{1.5}$$

En effet, (1.4) s'explique comme suit. Juste avant le moment $t + 2$, le nombre de couples de lapins en vie est u_{t+1}. Pour obtenir u_{t+2}, on doit y rajouter les couples nées au moment $t + 2$. Chaque couple de lapins âgé d'au moins deux mois au moment $t + 2$ en crée un. Les couples âgés d'au moins deux mois sont exactement les couples en vie au moment t; il y en a donc u_t tels couples. Ainsi, le nombre total de couples de lapins au moment $t + 2$ est égal à $u_{t+1} + u_t$, ce qui donne (1.4).

Au moment initial il y a un seule couple, qui pendant le premier mois est trop jeune pour procréer. Ainsi $u_0 = u_1 = 1$, d'ou (1.5).

Comment faire pour déduire de (1.4) et (1.5) une forme générale pour u_t?

On rappelle que $\mathbb{R}^{\mathbb{N}}$, l'ensemble des suites à valeurs réelles, a une structure de $\mathbb{R}$-espace vectoriel (voir la section 1.1). Posons:

$$\mathcal{S} := \big\{(v_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} :\ (v_n) \text{ satisfait } (1.4)\big\}.$$

Une première observation est que $\mathcal{S}$ est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$ de dimension 2. Montrons ce fait.

Par la proposition 1.4, pour montrer que $\mathcal{S}$ est un sous-espace vectoriel, il suffit de montrer que pour $(v_n)_{n\in\mathbb{N}} \in \mathcal{S}$, $(w_n)_{n\in\mathbb{N}} \in \mathcal{S}$ et $\lambda \in \mathbb{R}$, $(v_n + w_n)_{n\in\mathbb{N}} \in \mathcal{S}$ et $(\lambda v_n)_{n\in\mathbb{N}} \in \mathcal{S}$. Il s'agit d'une simple vérification qu'on laisse en exercice. Il faut en outre montrer que $\mathcal{S}$ est non-vide. Pour cela on observe que la suite nulle $(0, 0, \dots)$ est bien un élément de $\mathcal{S}$.

Montrons maintenant que $\dim(\mathcal{S}) = 2$. Posons $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ les suites définies par

$$a_{n+2} = a_{n+1} + a_n, \quad \text{et} \quad b_{n+2} = b_{n+1} + b_n \quad \text{pour tout } t \geq 0 \text{ et} \tag{1.6}$$

$$a_0 = 1, a_1 = 0, \qquad b_0 = 0, b_1 = 1. \tag{1.7}$$

On prétend que la famille formée des suites (a_n) et (b_n) est une base de $\mathcal{S}$ (donc qu'elle est libre et génératrice).

Montrons que $\big((a_n),(b_n)\big)$ est libre. Soit $\lambda,\eta \in \mathbb{R}$ tels que $\lambda\cdot(a_n)+\mu\cdot(b_n)=0$. Ce dernier fait signifie que

$$\lambda a_n + \mu b_n = 0, \qquad \text{pour tout } n \in \mathbb{N}.$$

En particulier

$$\lambda a_0 + \mu b_0 = \lambda = 0 \quad \text{ et } \lambda a_1 + \mu b_1 = \mu = 0.$$

Ainsi $\lambda = \mu = 0$ donc la famille est bien libre.

Montrons maintenant que $\big((a_n),(b_n)\big)$ est génératrice pour $\mathcal{S}$. Soit $(c_n) \in \mathcal{S}$. On va montrer que la suite (c_n) est égale à la suite $c_0\cdot(a_n)+c_1\cdot(b_n)$, à savoir que

$$c_n = c_0 a_n + c_1 b_n, \qquad \text{pour tout } n \in \mathbb{N}. \tag{1.8}$$

Montrons (1.8) par récurrence sur n.
Pour $n = 0, 1$ on a bien

$$c_0 a_0 + c_1 b_0 = c_0 \quad \text{ et } \quad c_0 a_1 + c_1 b_1 = c_1.$$

Soit $n \geq 1$ et supposons que $c_k = c_0 a_k + c_1 b_k$ pour tout $k \leq n$. Alors

$$\begin{aligned} c_{n+1} &= c_n + c_{n-1} && \text{car } (c_n) \text{ vérifie (1.4)}\\ &= c_0 a_n + c_1 b_n + c_0 a_{n-1} + c_1 b_{n-1} && \text{par l'hypothèse de récurrence}\\ &= c_0 a_{n+1} + c_1 b_{n+1} && \text{car } (a_n) \text{ et } (b_n) \text{ vérifient (1.4).} \end{aligned}$$

Ainsi, par récurrence (1.8) est vrai, donc $(c_n) \in \mathrm{Vect}\big((a_n),(b_n)\big)$. On a donc montré que $\big((a_n),(b_n)\big)$ est une base de $\mathcal{S}$. En particulier $\dim(\mathcal{S}) = 2$.

Malheureusement $\big((a_n),(b_n)\big)$ n'est pas une base de $\mathcal{S}$ qui s'écrit de façon confortable. Il est plus intéressant de chercher des suites géométriques qui forment une base de $\mathcal{S}$.

Soit α et β les deux solutions de l'équation

$$x^2 = x + 1. \tag{1.9}$$

Plus précisément, $\alpha = \frac{1+\sqrt{5}}{2}$ et $\beta = \frac{1-\sqrt{5}}{2}$. On va montrer que les suites $(\alpha^n)_{n\in\mathbb{N}}$ et $(\beta^n)_{n\in\mathbb{N}}$ appartiennent à $\mathcal{S}$ et qu'elles en forment une base.

Montons que $(\alpha^n)_{n\in\mathbb{N}}$ et $(\beta^n)_{n\in\mathbb{N}}$ vérifient (1.4). Soit $n \in \mathbb{N}$. On a

$$\begin{aligned} &\alpha^{n+2} - \alpha^{n+1} - \alpha^n = (\alpha^2 - \alpha - 1)\alpha^n = 0 \quad \text{ et}\\ &\beta^{n+2} - \beta^{n+1} - \beta^n = (\beta^2 - \beta - 1)\beta^n = 0, \end{aligned}$$

ce qui monte bien (1.4) pour $(\alpha^n)_{n\in\mathbb{N}}$ et $(\beta^n)_{n\in\mathbb{N}}$.

Vu que $\mathcal{S}$ est de dimension 2, pour montrer que la famille $\big((\alpha^n)_{n\in\mathbb{N}},(\beta^n)_{n\in\mathbb{N}}\big)$ est une base, il suffit de montrer qu'elle est libre. Soient $\lambda,\mu \in \mathbb{R}$ tels que

$$\lambda\alpha^n + \mu\beta^n = 0 \quad \text{ pour tout } n \in \mathbb{N}.$$

En particulier on a

$$\lambda\alpha^0 + \mu\beta^0 = \alpha + \beta = 0 \quad \text{et} \quad \lambda\alpha^1 + \mu\beta^1 = \frac{\lambda + \mu + \sqrt{5}(\lambda - \mu)}{2} = 0.$$

On en déduit

$$\lambda + \mu = 0 \quad \text{et} \quad \lambda - \mu = 0,$$

d'ou $\lambda = 0$ et $\mu = 0$. Donc $\big((\alpha^n)_{n\in\mathbb{N}}, (\beta^n)_{n\in\mathbb{N}}\big)$ est libre, donc une base.

En fin, vu que la suite (u_n) qui nous intéresse depuis le début fait partie de $\mathcal{S}$, on peut écrire

$$(u_n)_{n\in\mathbb{N}} = A(\alpha^n)_{n\in\mathbb{N}} + B(\beta^n)_{n\in\mathbb{N}}, \tag{1.10}$$

pour une certaine (unique) parie de scalaires $A, B \in \mathbb{R}$.

Trouvons A et B en utilisant (1.5). Si on écrit (1.10) pour $n = 0, 1$ on trouve

$$u_0 = 1 = A + B \quad \text{et} \quad u_1 = 1 = \frac{A + B + \sqrt{5}(A - B)}{2}.$$

Ce système se résout facilement et on trouve $A = \frac{1+\sqrt{5}}{2\sqrt{5}} = \frac{\alpha}{\sqrt{5}}$ et $B = \frac{\sqrt{5}-1}{2\sqrt{5}} = -\frac{\beta}{\sqrt{5}}$. Ainsi,

$$u_n = \frac{\alpha^{n+1}}{\sqrt{5}} - \frac{\beta^{n+1}}{\sqrt{5}}, \qquad \text{for all } n \in \mathbb{N}.$$

Suites récurrentes générales

Exercice 1.9.
Soit $a, b \in \mathbb{R}$ et considérons la suite définie par

$$u_{n+1} = au_n + bu_{n-1}, \quad \text{pour tout } n \geq 1 \tag{1.11}$$

et avec conditions initiales u_0, u_1 données (par exemple $u_0 = 0, u_1 = 1$). Le but c'est de déterminer un algorithme général pour évaluer u_n.

(i) Considérons l'équation quadratique

$$x^2 - ax - b = 0. \tag{1.12}$$

Donner la condition pour que cette équation ai deux solutions réelles $\alpha_1 \neq \alpha_2$.

(ii) Supposons que l'équation (1.12) admet deux solutions réelles $\alpha_1 \neq \alpha_2$. En s'inspirant de l'exemple précédent, donner une forme générale des suites qui satisfont (1.11). Expliquer comment exprimer $(u_n)_{n\geq 0}$ en termes des puissances de α_1 et α_2.

(iii) Supposons que l'équation (1.12) admet une solution double réelle α. En s'inspirant de l'exemple précédent, donner une suite qui satisfait (1.11). Vérifier que la suite $(n\alpha^n)_{n\geq 0} = (0, \alpha, 2\alpha^2, 3\alpha^3, \dots)$ satisfait également (1.11). Conclure.

(iv) Que faire si (1.12) n'admet pas de solutions réelles? Remarquer que l'équation admet quand même deux solutions complexes. Pouvez vous conclure?

(v) Donner une stratégie générale pour trouver la forme des suites définie par récurrence de tout ordre, c.-à-d. les suites de la forme

$$v_n = a_1 v_{n-1} + a_2 v_{n-2} + \dots + a_k v_{n-k}, \quad \text{pour tout } n \geq k,$$

ou $a_1, \dots, a_k$ sont des nombres fixés.

À retenir

- Un *espace vectoriel* (e.v.) E est un ensemble pour lequel on peut additionner les éléments entre eux et multiplier les éléments par des nombres. Ces opérations suivent des règles intuitives.

- Les éléments d'un e.v. sont appelés des *vecteurs*; les nombres sont appelés des *scalaires*.

- Un *sous-espace vectoriel* (s.e.v.) F de E est un sous ensemble de E stable par les opérations de E (addition et multiplication par scalaire).
  Un sous-espace vectoriel contient nécessairement le vecteur nul 0_E.

- Pour une famille $(x_1, \dots, x_n)$ de vecteurs de E, $\mathrm{Vect}(x_1, \dots, x_n)$ est le *sous-espace vectoriel* de E *engendrée* par $x_1, \dots, x_n$. Il est composé de toutes les combinaisons linéaires de $x_1, \dots, x_n$.
  Si $\mathrm{Vect}(x_1, \dots, x_n) = E$ on dit que $x_1, \dots, x_n$ est *génératrice* pour E.

- Une famille $(x_1, \dots, x_n)$ de vecteurs de E est dite *liée* (ou linéairement dépendante) si un des vecteur s'écrit comme combinaison linéaire des autres.
  Si elle n'est pas liée, la famille est dite *libre* (ou linéairement indépendante).

- Une *base* d'un e.v. E est une famille finie de vecteurs qui est libre et génératrice.

- Un e.v. E est dit de *dimension finie* s'il admet une famille génératrice finie.
  Si c'est le cas, alors E admet des bases et tout ses bases ont le même nombre de vecteurs. On appelle ce nombre la *dimension* de E et on l'écrit $\dim(E)$.

- Dans un espace vectoriel de dimension finie d:
  (i) une famille libre a au plus d vecteurs;
  (ii) une famille génératrice a au moins d vecteurs.

- Pour une famille $x_1, \dots, x_n$ on défini le *rang* de la famille par
  $$\mathrm{rang}(x_1, \dots, x_n) = \dim(\mathrm{Vect}(x_1, \dots, x_n)).$$
  On a $\mathrm{rang}(x_1, \dots, x_n) \leq n$.

- Si $\mathcal{B} = (e_1, \dots, e_d)$ est une base de E, alors tout vecteur $x \in E$ s'écrit de façon unique comme combinaison linéaire de $e_1, \dots, e_n$:
  $$x = \lambda_1 e_1 + \cdots + \lambda_d e_d, \qquad \text{avec } \lambda_1, \dots, \lambda_d \in \mathbb{R}.$$
  On écrit alors
  $$V_{\mathcal{B}}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_d \end{pmatrix}.$$

Zu wissen

- Ein *Vektorraum* E ist eine Menge, deren Elementen können addiert und mit Zahlen multipliziert werden können. Diese Operationen folgen elementaren Regeln.

- Die Elemente eines Vektorraums heissen *Vektoren*; eine Zahl heisst ein Skalar.

- Ein *Untervektorraum* F von E ist eine Untermenge von E, die bezüglich der Vektoraddition und der Skalarmultiplikation abgeschlossen ist.
  Ein Untervektorraum enthält notwendigerweise den Nullvektor 0_E.

- Für eine Familie $(x_1, \dots, x_n)$ von Vektoren aus E, definiert man $\mathrm{Vect}(x_1, \dots, x_n)$ (oder $\mathrm{Span}(x_1, \dots, x_n)$) als die lineare Hülle von $(x_1, \dots, x_n)$. Es besteht in allen Linearkombinationen von $x_1, \dots, x_n$; es ist der *Untervektorraum* von E den von $x_1, \dots, x_n$ erzeugt ist.
  Wenn $\mathrm{Vect}(x_1, \dots, x_n) = E$, sagt man, dass $x_1, \dots, x_n$ E *erzeugt*.

- Eine Familie $(x_1, \dots, x_n)$ von Vektoren im E sind *linear abhängig*, wenn sich irgendein Vektor aus $(x_1, \dots, x_n)$ als Linearkombination der anderen Vektoren schreiben lässt. Sonst ist die Familie *linear unabhängig*.

- Eine *Basis* eines Vektorraumes E ist eine endliche Familie von Vektoren die linear unabhängigen ist und die erzeugt E.

- Ein Vektorraum E heisst *endlich-dimensional*, falls er von einer endlichen Familie erzeugt wird.
  In diesem Fall, gibt es Basen von E. Alle Basen enthalten gleichviel Vektoren. Diese Zahl ist die *Dimension* von E und wird durch $\dim(E)$ beschrieben.

- In einem endlich dimensionalen Vektorraum von Dimension d:
  (i) eine lineare unabhängige Familie enthält höchstens d Vektoren;
  (ii) eine erzeugende Familie enthält mindestens d Vektoren.

- Der *Rang* einer Familie von Vektoren $x_1, \dots, x_n$ ist definiert durch

$$\mathrm{rang}(x_1, \dots, x_n) = \dim(\mathrm{Vect}(x_1, \dots, x_n)).$$

  Es gilt $\mathrm{rang}(x_1, \dots, x_n) \leq n$.

- Falls $\mathcal{B} = (e_1, \dots, e_d)$ eine Basis von E ist, dann lassen sich alle Vektoren $x \in E$ als eindeutige Linearkombinationen von $e_1, \dots, e_n$:

$$x = \lambda_1 e_1 + \cdots + \lambda_d e_d, \qquad \text{mit } \lambda_1, \dots, \lambda_d \in \mathbb{R}.$$

  Man schreibt dann

$$V_{\mathcal{B}}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_d \end{pmatrix}.$$

Chapter 2

Applications linéaires et matrices

2.1 Applications linéaires: généralités

Définition 2.1. *Soient E et F deux espaces vectoriels et $u : E \to F$ une fonction. On dit que u est une* application linéaire[(i)] *de E dans F si les conditions suivantes sont satisfaites*

(i) *pour tout $x, y \in E$, $u(x + y) = u(x) + u(y)$.*

(ii) *pour tout $x \in E$ et $\lambda \in \mathbb{R}$, $u(\lambda x) = \lambda u(x)$.*

L'ensemble des applications linéaires de E dans F est noté $\mathcal{L}(E, F)$.

Tout comme pour vérifier qu'un sous-ensemble est un sous-espace vectoriel, on dispose d'une formule abrégée pour vérifier qu'un application est linéaire.

Lemme 2.2. *Soient E et F deux espaces vectoriels et $u : E \to F$ une fonction. Alors u est une* application linéaire *si et seulement si*

$$u(\lambda\, x + \mu\, y) = \lambda\, u(x) + \mu\, u(y) \qquad \textit{pour tout } x, y \in E \textit{ et } \lambda, \mu \in \mathbb{R}.$$

Exercice 2.1.
Soient E et F deux espaces vectoriels. Montrer que pour $u \in \mathcal{L}(E, F)$

(a) $u(0_E) = 0_F$,

(b) pour tout $x \in E$, $u(-x) = -u(x)$.

Dans la suite les applications linéaires d'un espace vectoriel E dans lui-même vont jouer un rôle plus important. On les appelle des *automorphisme* de E et on note l'ensemble des automorphisme de E par $\mathcal{L}(E) := \mathcal{L}(E, E)$.

[(i)] Lineare Abbildung oder Vektorraumhomomorphismus

Exemple: Pour $d \geq 1$, les applications suivantes de $\mathbb{R}^d$ dans $\mathbb{R}^d$ sont linéaires, c.-à-d. appartiennent à $\mathcal{L}(\mathbb{R}^d)$. Pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$\begin{aligned} f(x) &= (-x_1, x_2 \dots, x_n), \\ g(x) &= (x_1, 2x_2 \dots, nx_n), \\ h(x) &= (x_1 + x_2, x_2 + x_3, \dots, x_{n-1} + x_n, x_n + x_1). \end{aligned}$$

Par contre, l'application $i : \mathbb{R}^n \to \mathbb{R}^n$ définie par $i(x) = (x_1 + 1, x_2, \dots, x_n)$ n'est pas linéaire.

Les applications suivantes sont linéaires, de $\mathbb{R}^n$ dans $\mathbb{R}$:

$$\begin{aligned} j(x) &= x_1 + \cdots + x_n \\ k(x) &= x_1 - x_2. \end{aligned}$$

Par contre $\ell, m : \mathbb{R}^d \to \mathbb{R}$ définies par $\ell(x) = \max\{x_1, x_2, \dots, x_n\}$ et $m(x) = |x_1|$ ne sont pas linéaires.

Prenons maintenant un espace vectoriel de dimension infinie. Pour $\mathbf{a} = (a_n)_{n\in\mathbb{N}} \in \ell(\mathbb{R})$ posons

$$\begin{aligned} u(\mathbf{a}) &= a_1 \in \mathbb{R}, \\ v(\mathbf{a}) &= (a_1 - a_0, a_2 - a_1, \dots) \in \ell(\mathbb{R}), \\ \sigma_g(\mathbf{a}) &= (a_1, a_2, \dots) \in \ell(\mathbb{R}), \\ \sigma_d(\mathbf{a}) &= (0, a_0, a_1, \dots) \in \ell(\mathbb{R}). \end{aligned} \tag{2.1}$$

Alors $u \in \mathcal{L}(\ell(\mathbb{R}), \mathbb{R})$ et $v, \sigma_g, \sigma_d \in \mathcal{L}(\ell(\mathbb{R}))$.

Exercice 2.2.

Soit E un espace vectoriel, $x \in E$ un vecteur et $\lambda \in \mathbb{R}$ un scalaire. Quelles des deux applications suivantes sont linéaires?

$$\begin{aligned} u :& E \to E \\ & y \mapsto y + x \end{aligned} \qquad\qquad \begin{aligned} v :& E \to E \\ & y \mapsto \lambda y. \end{aligned}$$

Exercice 2.3.

Soient E, F deux espaces vectoriels et $u \in \mathcal{L}(E, F)$. Montrer que, si G est un sous-espace vectoriel de E, alors $u(G) = \{u(x) : x \in G\}$ est un sous-espace vectoriel de F.

Exercice 2.4.

Quelles parmi les fonctions suivantes sont linéaires? Motivez vos réponses.

$$\begin{aligned} f :& \mathbb{R}^2 \to \mathbb{R}^2 \\ & \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} b \\ a \end{pmatrix} \end{aligned} \qquad \begin{aligned} g :& \mathbb{R}^2 \to \mathbb{R}^2 \\ & \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a+b \\ a-b \end{pmatrix} \end{aligned} \qquad \begin{aligned} h :& \mathbb{R}^2 \to \mathbb{R}^2 \\ & \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a+1 \\ b-1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} u :& \mathbb{R}[X] \to \mathbb{R} \\ & P \mapsto P(1) \end{aligned} \qquad \begin{aligned} v :& \mathbb{R}[X] \to \mathbb{R} \\ & P \mapsto P(0) + P(1) \end{aligned} \qquad \begin{aligned} w :& \mathbb{R}[X] \to \mathbb{R} \\ & P \mapsto \max_{t\in[0,1]} P(t) \end{aligned}$$

L'ensemble $\mathcal{L}$ comme espace vectoriel et algèbre Fixons pour la suite de cette partie deux espaces vectoriels E, F. L'ensemble $\mathcal{L}(E,F)$ est lui-même munit d'une structure d'espace vectoriel, héritée de F. La multitude d'espaces vectoriels peut créer des confusions dans une première phase; ainsi on va re-introduire brièvement l'indice pour le 0 qui marque l'espace vectoriel au quel il appartient.

Les opérations d'addition et multiplication par une constante sont naturellement définies sur $\mathcal{L}(E,F)$ comme suit. Soient $u, v \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{R}$. On définit les fonctions $u+v$ et λu de E dans F par

$$(u+v)(x) = u(x) + v(x) \quad \text{et} \quad (\lambda u)(x) = \lambda u(x), \quad \text{pour tout } x \in E.$$

Il est facile de vérifier que les applications ainsi définies sont en effet des applications linéaires, donc des membres de $\mathcal{L}(E,F)$.

Avec ces opérations, $\mathcal{L}(E,F)$ est un espace vectoriel. La preuve est simple et on la laisse en exercice. On mentionne uniquement que l'élément neutre est la fonction nulle $0_{\mathcal{L}}$ définie par

$$0_{\mathcal{L}}(x) = 0_F \quad \text{pour tout } x \in E.$$

Exercice 2.5.
Montrer que, pour $u, v \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{R}$, on a $u+v \in \mathcal{L}(E,F)$ et $\lambda u \in \mathcal{L}(E,F)$. Montrer que $(\mathcal{L}(E,F), +, \cdot)$ est bien un espace vectoriel.

Si on se restreint de plus aux automorphisme d'espaces vectoriels (à savoir à $\mathcal{L}(E)$), on voit apparaitre une structure encore plus riche.

Pour $u, v \in \mathcal{L}(E)$ on peut définir la *composition* des deux fonctions, qu'on note $u \circ v$. On rappelle qu'il s'agit de la fonction de E dans E définie par

$$(u \circ v)(x) = u(v(x)) \quad \text{pour tout } x \in E.$$

On peut facilement vérifier que $u \circ v$ est une application linéaire de E dans E.

Ainsi $\mathcal{L}(E)$ est muni de trois opérations: l'addition $+$, la multiplication par un scalaire $\cdot$ et la composition $\circ$. En plus des propriétés d'espace vectoriel, $\mathcal{L}(E)$ a les propriétés suivantes. On dit que $\mathcal{L}(E, +, \cdot, \circ)$ est une *algèbre*.

Proposition 2.3.
(i) *Pour tout $u, v, w \in \mathcal{L}(E)$ on a $u \circ (v+w) = u \circ v + u \circ w$ et $(u+v) \circ w = u \circ w + v \circ w$;*
(ii) *pour tout $u, v \in \mathcal{L}(E)$ et $\lambda \in \mathbb{R}$ on a $\lambda(u \circ v) = (\lambda u) \circ v = u \circ (\lambda v)$;*
(iii) *pour tout $u \in \mathcal{L}(E)$ on a $u \circ 0_{\mathcal{L}} = 0_{\mathcal{L}} \circ u = 0_{\mathcal{L}}$;*
(iv) *il existe un unique élément de $\mathcal{L}(E)$ qu'on note* id *tel que pour tout $u \in \mathcal{L}$ $u \circ \mathrm{id} = \mathrm{id} \circ u = u$.*

Mentionnons que le l'élément neutre pour la composition est l'application $\mathrm{id} \in \mathcal{L}(E)$ définie par

$$\mathrm{id}(x) = x, \qquad \text{pour tout } x \in E. \tag{2.2}$$

On va admettre cette proposition, dont la preuve consiste de vérifications immédiates.

2.1.1 Image et noyau

Définition 2.4. *Soit $u \in \mathcal{L}(E, F)$.*
(i) *On appelle* l'image[ii] *de u le sous-ensemble de F*

$$\operatorname{Im}(u) := \{u(x) : x \in E\}.$$

(ii) *On appelle* le noyau[iii] *de u le sous-ensemble de E*

$$\operatorname{Ker}(u) := \{x \in E : u(x) = 0_F\}.$$

Proposition 2.5. *Soit $u \in \mathcal{L}(E, F)$. L'image et le noyau de u sont des sous-espaces vectoriels de F et de E, respectivement.*

Preuve: Commençons par l'image. Comme $u(0_E) = 0_F$ (voir l'exercice 2.1), l'image de u n'est pas vide. Soit $y_1, y_2 \in \operatorname{Im}(u)$ et $\lambda \in \mathbb{R}$. Par la définition de l'image, il existe $x_1, x_2 \in E$ tels que $u(x_1) = y_1$ et $u(x_2) = y_2$. Alors, par linéarité de u,

$$u(x_1 + x_2) = u(x_1) + u(x_2) = y_1 + y_2 \quad \text{et} \quad u(\lambda x_1) = \lambda u(x_1) = \lambda y_1.$$

On en déduit que $y_1 + y_2 \in \operatorname{Im}(u)$ et $\lambda y_1 \in \operatorname{Im}(u)$. Par la proposition 1.4, $\operatorname{Im}(u)$ est un sous-espace vectoriel de F.

Passons à $\operatorname{Ker}(u)$. On a $u(0_E) = 0_F$, donc $0_E \in \operatorname{Ker}(u)$, et en particulier $\operatorname{Ker}(u)$ n'est pas vide. Soit $x_1, x_2 \in \operatorname{Ker}(u)$ et $\lambda \in \mathbb{R}$. On a

$$u(x_1 + x_2) = u(x_1) + u(x_2) = 0_F \quad \text{et} \quad u(\lambda x_1) = \lambda u(x_1) = 0_F,$$

donc $x_1 + x_2 \in \operatorname{Ker}(u)$ et $\lambda x_1 \in \operatorname{Ker}(u)$. Par la proposition 1.4 encore, $\operatorname{Ker}(u)$ est un sous-espace vectoriel de E. □

Bien évidement, $u \in \mathcal{L}(E, F)$ est surjective si et seulement si $\operatorname{Im}(u) = F$. Un critère similaire pour l'injectivité est donné par la proposition suivante.

Proposition 2.6. *Soit $u \in \mathcal{L}(E, F)$. Alors u est injective si et seulement si* $\operatorname{Ker}(u) = \{0_E\}$.

[ii] Bild
[iii] Kern

Preuve: Supposons pour commencer que u est injective. Il est toujours vrai que $u(0_E) = 0_F$, et comme u est injective, il n'y a pas d'autre élément $x \in E$ tel que $u(x) = 0_F$. Ainsi $\mathrm{Ker}(u) = \{0_E\}$.

Inversement, supposons que $\mathrm{Ker}(u) = \{0_E\}$. Soient $x, y \in E$ tels que $u(x) = u(y)$. Afin de montrer que u est injective, il faut montrer que $x = y$. Par linéarité de u,

$$u(x - y) = u(x) - u(y) = 0_F,$$

donc $x - y \in \mathrm{Ker}(u)$. Ainsi $x - y = 0_E$, d'ou $x = y$. □

2.1.2 Inverses des applications linéaires

Soit $u \in \mathcal{L}(E, F)$. Rappelons-nous que u est avant tout une fonction, il existe donc une notion d'inverse de u en tant que fonction. On va utiliser cette même notion d'inverse dans le cadre des applications linéaires.

Comme toute fonction, u admet un inverse si et seulement si elle est bijective. On dira souvent que u est *inversible* pour dire simplement que u est bijective. Si u est inversible, l'application inverse de u, notée $u^{-1} : F \to E$, est définie par:

$$\text{pour tout } y \in F,\ u^{-1}(y) \text{ est le seul vecteur de } E \text{ tel que } u(u^{-1}(y)) = x.$$

Proposition 2.7. *Si $u \in \mathcal{L}(E, F)$ est inversible, alors $u^{-1} \in \mathcal{L}(F, E)$.*

Preuve: On admet ce résultat. □

Exercice 2.6.
Soit $u, v \in \mathcal{L}(E)$ et supposons que u est bijective. Monter que
(a) $\mathrm{Ker}(u \circ v) = \mathrm{Ker}(v)$,
(b) $\mathrm{Im}(v \circ u) = \mathrm{Im}(v)$,

2.1.3 Applications linéaires en dimension finie

Fixons deux espaces vectoriels E, F. On va supposer que E est de dimension finie. Le théorème suivant est un outil très puissant pour l'étude des applications linéaires.

Théorème 2.8 (Théorème du rang)**.** *Soit $u \in \mathcal{L}(E, F)$. Alors*

$$\dim(\mathrm{Ker}(u)) + \dim(\mathrm{Im}(u)) = \dim(E).$$

Preuve: Ce résultat est nécessite une preuve assez complexe, on va donc l'admettre. □

Un corollaire immédiat du théorème du rang est le suivant.

Corollaire 2.9. *L'image de u est de dimension finie et* $\dim(\mathrm{Im}(u)) \leq \dim(E)$.

La dimension de $\mathrm{Im}(u)$ est souvent appelée le *rang* de u et est notée

$$\mathrm{rang}(u) := \dim(\mathrm{Im}(u)).$$

Deux conséquences qui témoignent de la puissance du théorème 2.8 sont les suivantes

Corollaire 2.10. *S'il existe une application linéaire inversible $u \in \mathcal{L}(E, F)$, alors* $\dim(E) = \dim(F)$.

Preuve: Soit $u \in \mathcal{L}(E, F)$ inversible. Alors $\dim(E) = \dim(\mathrm{Ker}(u)) + \dim(\mathrm{Im}(u)) = \dim(F)$, car $\mathrm{Ker}(u) = \{0\}$, donc $\dim(\mathrm{Ker}(u)) = 0$ et $\mathrm{Im}(u) = F$, donc $\dim(\mathrm{Im}(u)) = \dim(F)$. □

Corollaire 2.11. *Soit $u \in \mathcal{L}(E)$. Alors on a équivalence de:*

(i) *u est injective;*
(ii) *u est surjective;*
(iii) *u est inversible.*

Preuve: Evidement $(iii) \Rightarrow (i)$ et $(iii) \Rightarrow (ii)$. On va montrer les deux implications inverses.

Soit $u \in \mathcal{L}(E)$ injective. Alors $\dim(\mathrm{Ker}(u)) = 0$, donc par le théorème 2.8, $\dim(E) = \dim(\mathrm{Im}(u))$. Mais $\mathrm{Im}(u)$ est un s.e.v. de E; la proposition 1.14 implique alors que $\mathrm{Im}(u) = E$, donc que u est surjective. Ainsi u est surjective et injective, donc bijective.

Supposons maintenant que $u \in \mathcal{L}(E)$ est surjective. Alors le théorème 2.8 dit que $\dim(E) = \dim(\mathrm{Ker}(u)) + \dim(\mathrm{Im}(u)) = \dim(\mathrm{Ker}(u)) + \dim(E)$, donc que $\dim(\mathrm{Ker}(u)) = 0$. En autres mots $\mathrm{Ker}(u) = \{0_E\}$, ce qui veut dire que u est injective, donc bijective. □

Exercice 2.7.
Soit $E = \mathbb{R}^3$ et soient $u, v : \mathbb{R}^3 \to \mathbb{R}^3$ définies par

$$u\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3a + 2c \\ a + b + c \\ c - a \end{pmatrix} \quad \text{et} \quad v\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3a + 2c \\ a + b + c \\ b + 2c \end{pmatrix},$$

pour $a, b, c \in \mathbb{R}$.

(a) Montrer que $u, v \in \mathcal{L}(\mathbb{R}^3)$
(b) Est-ce que u ou v est inversible?
Indication: Chercher à voir si le noyaux est réduit à $\{0\}$.

2.2 Introduction aux matrices

Les matrices sont des objets omniprésentes dans les différentes domaines des mathématiques. De plus elles sont très souvent utilisées dans les applications des mathématiques. Une de leur utilisations essentielles est la description des applications linéaires.

On commence par donner une descriptions des matrices et des opérations s'y appliquant, ensuite on va aborder le lien avec les applications linéaires.

Définition 2.12. *Soient $m, n \in \mathbb{N}$. Une matrice A de taille $m \times n$ à coefficients dans $\mathbb{R}$ est une famille $(a_{i,j})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ d'éléments de $\mathbb{R}$. Elle est représenté par un tableau rectangulaire*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

L'indice i est l'indice des lignes et l'indice j est celui des colonnes.
L'ensemble de matrices de taille $m \times n$ à coefficients dans $\mathbb{R}$ est noté $\mathcal{M}_{m,n}(\mathbb{R})$.

2.2.1 Opérations sur les matrices

Les matrices peuvent être additionnées, multipliées par des constantes et dans certains cas multipliées entre elles.

Soient $m, n \in \mathbb{N}$ et deux matrices $A = (a_{i,j})$ et $B = (b_{i,j})$ de $\mathcal{M}_{m,n}(\mathbb{R})$. De plus, soit $\lambda \in \mathbb{R}$. On défini les matrices $A + B \in \mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda \cdot A \in \mathcal{M}_{m,n}(\mathbb{R})$ par

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{pmatrix} \quad \text{et} \quad \lambda \cdot A = \begin{pmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{pmatrix}.$$

Proposition 2.13. *L'ensemble $\mathcal{M}_{m,n}(\mathbb{R})$ avec les opérations définies au-dessus est un espace vectoriel. L'élément neutre est la matrice*

$$0_{m,n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}).$$

Exercice 2.8.
Démontrer la proposition 2.13. (Il s'agit simplement de vérifier que les conditions de la définition 1.1.)

Soient $m, n, p \in \mathbb{N}$. Pour $A = (a_{i,k}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B = (b_{k,j}) \in \mathcal{M}_{n,p}(\mathbb{R})$ on définit le *produit des matrices* A et B comme la matrice $AB = (c_{i,j}) \in \mathcal{M}_{m,p}(\mathbb{R})$ (aussi notée $A \times B$) dont les entrées sont données par

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \qquad \text{pour } 1 \leq i \leq m \text{ et } 1 \leq j \leq p.$$

Attention! pour des matrices A, B dont les tailles ne sont pas compatibles (comme dans la définition au-dessus) le produit de A et B n'est pas définit.

La multiplication matriciélle suit des règles intuitives, similaires aux règles de la multiplication des nombres. Elles sont décrites dans la proposition suivante.

Proposition 2.14. *Soient $m, n, p, q \in \mathbb{N}$.*
(i) *Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B, C \in \mathcal{M}_{n,p}(\mathbb{R})$ on a $A(B + C) = AB + AC$.*
(ii) *Pour tout $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ et $C \in \mathcal{M}_{n,p}(\mathbb{R})$ on a $(A + B)C = AC + BC$.*
(iii) *Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et $C \in \mathcal{M}_{p,q}(\mathbb{R})$ on a $(AB)C = A(BC)$.*
(iv) *Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B, C \in \mathcal{M}_{n,p}(\mathbb{R})$ et $\lambda \in \mathbb{R}$ on a $\lambda(AB) = (\lambda A)B = A(\lambda B)$.*

On va surtout utiliser la multiplication des matrices pour les matrices carrées, à savoir les matrices de $\mathcal{M}_{n,n}(\mathbb{R})$ pour un $n \in \mathbb{N}$. On écrira $\mathcal{M}_n(\mathbb{R})$ à la place de $\mathcal{M}_{n,n}(\mathbb{R})$ pour raccourcir la notation. La proposition 2.14 est surtout importante dans ce contexte; on la complète par propriété suivante.

Proposition 2.15. *Soit $n \in \mathbb{N}$. Alors la matrice*

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

est l'unique matrice de $\mathcal{M}_n(\mathbb{R})$ telle que, pour tout $A \in \mathcal{M}_n(\mathbb{R})$,

$$AI_n = I_n A = A. \tag{2.3}$$

On appelle I_n la matrice identité *de taille n.*

Preuve: Commençons par montrer que 2.3 est satisfait pour tout $A \in \mathcal{M}_n(\mathbb{R})$. Par la règle de multiplication des matrices, pour une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$, on a

$$A \cdot I_n = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = A.$$

De même $I_n \cdot A = A$.

Montrons maintenant l'unicité de la matrice I_n. Soit $J \in \mathcal{M}_n(\mathbb{R})$ telle que $A \cdot J = J \cdot A = A$ pour tout $A \in \mathcal{M}_n(\mathbb{R})$. En appliquant cela à $A = I_n$, et en utilisant (2.3) on trouve

$$I_n = I_n \cdot J = J.$$

□

Attention! Généralement la multiplication des matrices n'est pas commutative. En effet, pour $n \geq 2$ il existe $A, B \in \mathcal{M}_n(\mathbb{R})$ tels que $AB \neq BA$.

Exercice 2.9.
Trouver pour tout $n \geq 2$ des matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ tels que $AB \neq BA$.

2.2.2 Matrices inversibles

Définition 2.16. *Soit $n \in \mathbb{N}$ et $A \in \mathcal{M}_n(\mathbb{R})$. On dit qu'une matrice $B \in \mathcal{M}_n(\mathbb{R})$ est un inverse de A si $AB = BA = I_n$. Si une telle matrice existe, on dit que A est* inversible[iv].

L'étude des matrices inversibles est particulièrement intéressant. On donne une première proposition les concernant; on va y revenir dans les parties 2.3.3 et 2.3.4.

Proposition 2.17.

(i) *Si A est inversible, alors il existe un unique inverse qu'on note A^{-1}.*

(ii) *Soient $A, B \in \mathcal{M}_n(\mathbb{R})$ deux matrices inversibles. Alors AB est une matrice inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

Exercice 2.10.
Démontrer la proposition 2.17.

Exercice 2.11.
Soient $A, B \in \mathcal{M}_n(\mathbb{R})$ telles que $AB = 0$. Montrer que A et B ne sont pas inversibles.
Exemple: Calculer

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Déduire des exemples de matrices non-nulles mais non-inversibles.

[iv] Reguläre Matrix

2.2.3 La transposition

Une opération spécifique aux matrices est la transposition. Elle consiste à échanger les lignes et les colonnes d'une matrice. On peut aussi la voir comme une réflexion de la matrice suivant sa diagonale.

Définition 2.18. *Soient $m, n \in \mathbb{N}$ et $A = (a_{i,j})_{\substack{1\le i\le m\\1\le j\le n}} \in \mathcal{M}_{m,n}(\mathbb{R})$. La* transposée *de A est la matrice $A^T = (a^T_{i,j})_{\substack{1\le i\le n\\1\le j\le m}} \in \mathcal{M}_{n,m}(\mathbb{R})$ définie par $a^T_{i,j} = a_{j,i}$ pour tout $1 \le i \le n$ et $1 \le j \le m$.*

Donnons quelques exemples pour illustrer la notion.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Proposition 2.19. *Soit $A, B \in \mathcal{M}_n(\mathbb{R})$. Alors $(AB)^T = B^T A^T$.*

Preuve: Montrons l'égalité élément par élément. Notons $A = (a_{i,j})_{1\le i,j\le n}$ et $B = (b_{i,j})_{1\le i,j\le n}$. On va aussi noter $a^T_{i,j} = a_{j,i}$ et $b^T_{i,j} = b_{j,i}$ les entrées de A^T et B^T.

Soient $k, \ell \in [1, n]$. L'entrée de la ligne k et la colonne ℓ de $(AB)^T$ est égale à l'entrée de la ligne ℓ et la colonne k de AB. Il s'agit donc de

$$(AB)^T_{k,\ell} = a_{\ell,1}b_{1,k} + \cdots + a_{\ell,n}b_{n,k}.$$

D'autre part, l'entrée de la ligne k et la colonne ℓ de $B^T A^T$ est

$$(B^T A^T)_{k,\ell} = b^T_{k,1}a^T_{1,\ell} + \cdots + b^T_{k,n}a^T_{n,\ell} = b_{1,k}a_{\ell,1} + \cdots + b_{n,k}a_{\ell,n}.$$

Ainsi $(AB)^T = B^T A^T$. □

Corollaire 2.20. *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors A est inversible si et seulement si A^T l'est. Quand les deux sont inversibles, $(A^T)^{-1} = (A^{-1})^T$.*

Preuve: En utilisant la proposition 2.19 on vérifie que si A est inversible

$$(A^{-1})^T A^T = (AA^{-1})^T = I_n = (A^{-1}A)^T = A^T(A^{-1})^T.$$

Donc A^T est inversible et $(A^T)^{-1} = (A^{-1})^T$. □

Exercice 2.12.
Trouver toutes les matrices $A \in \mathcal{M}_n(\mathbb{R})$ telles que
(a) $A = A^T$.
(b) $A = -A^T$.

2.2.4 Formes particulières

Il est souvent commode d'écrire une matrice comme une collection de lignes ou de colonnes. Soit $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$. On peut alors écrire

$$A = \begin{pmatrix} \\ C_1 \dots C_n \\ \\ \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix},$$

ou $C_1, \dots, C_n$ sont les colonnes de A et $L_1, \dots, L_m$ sont les lignes. Plus précisément, pour $1 \leq i \leq m$ et $1 \leq j \leq n$,

$$L_i = (a_{i,1}, \dots, a_{i,n}) \in \mathcal{M}_{1,n} \quad \text{et} \quad C_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} \in \mathcal{M}_{m,1}.$$

Pour les matrices carrées, certaines formes de matrices sont particulièrement intéressantes.

Définition 2.21. *Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$. On dit que A est une* matrice triangulaire supérieure *si $a_{i,j} = 0$ pour tout $i > j$, et* triangulaire inférieure, *si $a_{i,j} = 0$ pour tout $i < j$.*
On dit que A est une matrice diagonale si $a_{i,j} = 0$ pour tout $i \neq j$.

Les formes générales des matrices triangulaires supérieures, triangulaires inférieures et diagonales, respectivement, sont:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n} \\ 0 & a_{22} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}; \qquad \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}; \qquad \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}.$$

2.2.5 Puissance des matrices

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Comme pour les nombres, on peut multiplier A par elle même plusieurs fois pour obtenir les puissances de A. Ainsi on pose, pour $k \in \mathbb{N}$, $k \geq 1$:

$$A^k = \underbrace{A \dots A}_{k \text{ fois}}.$$

On pose aussi par convention $A^0 = I_n$.

Certaines propriétés des puissances des nombres sont conservées. Pour $k, \ell \in \mathbb{N}$ on a bien

$$A^k A^\ell = A^{k+\ell} \quad \text{et} \quad \left(A^k\right)^\ell = A^{k\ell}.$$

D'autres ne le sont pas...

Exercice 2.13.
Trouver trois solutions distinctes $A \in \mathcal{M}_2(\mathbb{R})$ de l'équation $A^2 = A$.

Puissance des matrices diagonales Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonale:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}.$$

On peut alors facilement calculer les puissances de A:

$$A^k = \begin{pmatrix} a_{11}^k & 0 & \dots & 0 \\ 0 & a_{22}^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n}^n \end{pmatrix}, \quad \text{pour tout } k \in \mathbb{N}. \tag{2.4}$$

Exercice 2.14.
Montrer la formule (2.4) par récurrence sur n.

Le calcul des puissances d'une matrice est un problème très intéressant d'un point de vu pratique. On a vu quelques exemples ou le calcul est possible, mais en général, il s'agit d'un problème compliqué. Illustrons son utilité par un exemple.

2.2.6 Les matrices comme outil de modélisation

On a déjà vu comment l'algèbre linéaire peut être utile pour la résolution d'équations de récurrence comme celle de Fibonacci (voir partie 1.4). En utilisant les matrices on peut décrire (et résoudre, comme on va le voir plus tard) des équations de récurrence plus complexes.

Population de deux types de bactéries Supposons qu'on a une population de bactéries de deux types A et B. A chaque instant, une bactérie de type A donne naissance à une bactérie de type A et quatre de type B, puis meure. Une bactérie de type B donne naissance à une bactérie de type A et une de type B. Quelle est la population après n étapes, sachant qu'on commence par une seule bactérie de type A?

On peut modéliser cela comme suit. Notons a_n et b_n le nombre de bactéries de type A et B, respectivement, à l'étape n. Alors on a la formule suivante:

$$a_{n+1} = a_n + b_n \quad \text{et} \quad b_{n+1} = 4a_n + b_n. \tag{2.5}$$

De plus la condition initiale est $a_0 = 1$ et $b_0 = 0$.

La population totale de bactéries après n étapes s'écrit $p_n = a_n + b_n$. Remarquons qu'on n'a pas une simple équation de récurrence pour p_n comme celles traitées dans la partie 1.4.

Par contre, si on note $X_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ pour chaque $n \in \mathbb{N}$, on observe que

$$X_{n+1} = \begin{pmatrix} a_n + b_n \\ 4a_n + b_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} X_n.$$

Ainsi

$$X_n = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}^n X_0 = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Malheureusement il n'est pas evident de deviner la forme générale de la matrice $\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}^n$. On va voir dans la partie 3 une technique pour le faire. Il se trouve que dans le cas présent, si on pose $c_n = 2a_n - b_n$ et $d_n = 2a_n + b_n$ pour tout n, alors

$$c_{n+1} = -c_n \quad \text{et} \quad d_{n+1} = 3d_n, \quad \forall n \in \mathbb{N}.$$

Ainsi

$$\begin{pmatrix} c_n \\ d_n \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}^n \begin{pmatrix} c_0 \\ d_0 \end{pmatrix} = \begin{pmatrix} (-1)^n c_0 \\ 3^n d_0 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-1)^n \\ 2 \cdot 3^n \end{pmatrix}.$$

Cela suffit de retrouver a_n et b_n car

$$a_n = \frac{1}{4}(c_n + d_n) = \frac{1}{2}(3^n + (-1)^n) \quad \text{et} \quad b_n = \frac{1}{2}(d_n - c_n) = 3^n - (-1)^n.$$

Matrices d'adjacence et matrices stochastiques. Un graphe est un couple d'ensembles $G = (V, E)$ ou les éléments de V sont les sommets de G et $E \subset V \times V$ représente les arêtes de G. Il symbolise un réseau avec des arêtes orientées (ou pas, suivant la définition exacte) entre certains de ses sommets. Pour un graphe fini G (c.à-d. avec V fini), on définit sa matrice d'adjacence comme suit.

On commence par noter $V = \{v_1, \dots, v_n\}$. La matrice d'adjacence $A = (a_{i,j})_{1 \le i,j \le n} \in \mathcal{M}_n(\mathbb{R})$ est définie par

$$a_{i,j} = 1 \text{ si et seulement si } (v_i, v_j) \in E,$$

donc si et seulement si il y a une arête de v_i vers v_j. Sinon, on pose $a_{i,j} = 0$. Ainsi la matrice d'adjacence décrit parfaitement le graphe G.

Si de plus le graphe G a un poids positif w_e associé à chaque arête $e \in E$, on peut rajouter cette information à la matrice A, en posant

$$a_{i,j} = \begin{cases} w_{(v_i, v_j)} & \text{si } (v_i, v_j) \in E, \\ 0 & \text{si } (v_i, v_j) \notin E. \end{cases}$$

Un vecteur colonne $X \in \mathcal{M}_{n,1}(\mathbb{R})$ (avec des entrées positives) peut être interprété comme une attribution de "masse" à chaque sommet: chaque sommet v_i a une masse x_i.

Imaginons la dynamique suivante: à chaque étape, chaque sommet v_i transfère une proportion $a_{i,j}$ de sa masse au sommet v_j. Pour que cela ai un sens, il faut que

$$\sum_{j=1}^{n} a_{i,j} = 1, \qquad \forall 1 \le i \le n. \tag{2.6}$$

Alors, si on commence par une attribution de masse X, après une étape, on aura une distribution de masse Y, ou

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \text{avec} \quad y_j = \sum_{i=1}^{n} a_{i,j} x_i, \ \forall 1 \leq j \leq n.$$

On peut écrire cela de façon plus compacte comme $Y = A^T X$. Plus généralement, après k étapes, on aura une distribution $(A^T)^k X$.

La condition (2.6) garantie qu'il y a conservation de la masse totale dans le graphe. Si cette condition est satisfaite, on dit que A est une *matrice stochastique* à droite et A^T une matrice stochastique à gauche. Ainsi, dans une matrice stochastique à droite, la somme des entrées sur chaque ligne vaut 1; dans une matrice stochastique à gauche, c'est somme des entrées sur chaque colonne qui vaut 1.

Cette dynamique décrit une *marche aléatoire* sur le graphe G (un cas particulier d'une *chaine de Markov*). Imaginons qu'on dispose d'un kilo de sable. On se donne un vecteur colonne $X = (x_i)_{1 \leq i \leq n} \in \mathcal{M}_{n,1}$ avec

$$x_i \geq 0, \text{ pour tout } 1 \leq i \leq n \quad \text{et} \quad \sum_{i=1}^{n} x_i = 1.$$

Pour chaque i, on place x_i kilos de sable au sommet v_i de G. On lance ensuite la dynamique décrite avant: le sable de v_i est distribué entre ses voisins, v_j recevant de v_i une proportion $a_{i,j}$ de sable.

Supposons qu'un grain de sable est coloré en rouge et qu'on le suit pendent cette évolution. Alors, à chaque étape, si le grain se trouve à v_i, il a une probabilité $a_{i,j}$ d'être envoyé au sommet v_j. On va supposer que cette dynamique est markovienne, c.à-d. que l'évolution future du grain de sable dépend uniquement de sa position actuelle, pas de son évolution passée.

Ainsi, la positon du grain rouge après k étapes est décrite par le vecteur

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \left(A^T\right)^k X.$$

En effet, il y a une quantité y_i de sable au sommet v_i, donc une probabilité y_i que le grain rouge s'y trouve.

Exemple: la parade nuptiale des bourdons. Comme discuté avant, les matrices peuvent être utilisées pour décrire certaines évolutions aléatoires appelées des *chaines de Markov*. On va illustrer cela par un exemple concret.

L'accouplement des bourdons suit une procédure à plusieurs étapes. Celles-ci peuvent être classées comme suit:

(App) Approche: un mâle se dirige vers la reine. Il s'approche à courte distance de la reine, et peut continuer la parade, ou se retirer (c'est le plus souvent la cas).

(IF) Inspection de la femelle: le mâle suit la reine avec ses antennes tendues vers elle. Il inspecte souvent la reine au niveau de la tête (région où se trouvent les glandes produisant les phéromones sexuelles), mais parfois au niveau de l'abdomen.

(T) Tentative d'accouplement: le mâle s'approche de la reine, il s'accroche à elle. Il frotte de ses pattes antérieures l'extrémité de l'abdomen de la femelle. Il sort ses génitalias (appareil reproducteur) et tente de pénétrer la reine.

(Acc) Accouplement: lors de l'accouplement, le comportement du mâle se caractérise par des mouvements de battements des pattes sur l'extrémité de l'abdomen de la reine.

Pour observer la séquence, on place 80 bourdons dans un milieu favorable, et on les observe pendant 15 minutes. Les bourdons passent par les différentes phases de la parade nuptiale. On observe chaque bourdon chaque minute pour déterminer les phases par les quelles il passe. On rajoute quelques états dans notre tableau:

(D) depart: la situation de depart.

(AA) Accouplement accompli: le mâle quitte la séance après accouplement.

(AM) Abandon du mâle: lors de la séquence, le bourdon mâle peut adopter un comportement indifférent vis-à-vis de la reine; il sort de la parade nuptiale et n'y revient jamais.

Les observations sont classées dans le tableau suivant:

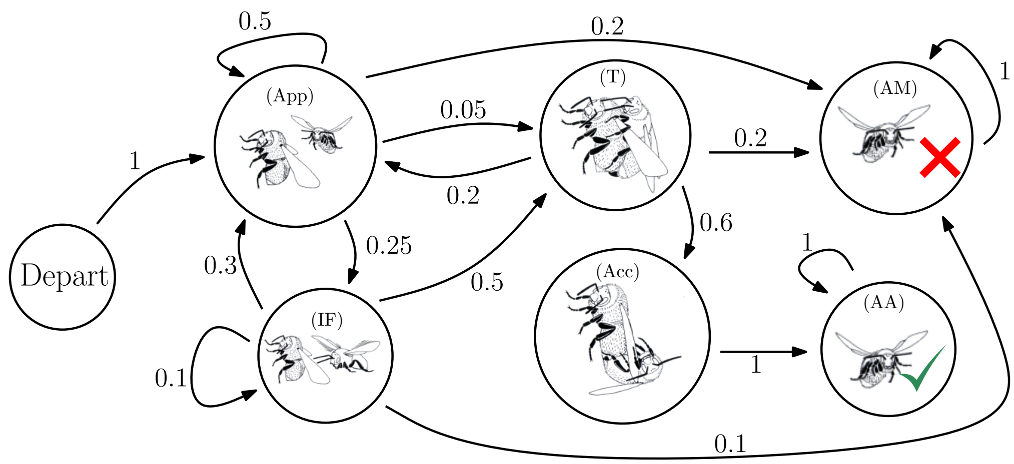
	Vers					
De ↓	App	IF	T	Acc	AM	Total
D	80	0	0	0	0	80
App	102	51	10	0	41	204
IF	16	6	28	0	7	57
T	6	0	0	22	10	38

Si on suppose l'évolution des bourdons markovienne, on peut estimer à partir de ces observations les probabilités de passage d'une phase à une autre. Elles sont obtenues en divisant chaque ligne du tableau précédent par le nombre total d'observations y correspondent. Rajoutons les transition à partir des états Acc, AM et AA:

	App	IF	T	Acc	AM	AA
D	1	0	0	0	0	0
App	$\frac{102}{204} = 0.5$	$\frac{51}{204} = 0.25$	$\frac{10}{204} = 0.05$	0	$\frac{41}{204} = 0.2$	0
IF	$\frac{16}{57} = 0.3$	$\frac{6}{57} = 0.1$	$\frac{28}{57} = 0.5$	0	$\frac{7}{57} = 0.1$	0
T	$\frac{6}{38} = 0.15$	0	0	$\frac{22}{38} = 0.6$	$\frac{10}{38} = 0.25$	0
Acc	0	0	0	0	0	1
AM	0	0	0	0	1	0
AA	0	0	0	0	0	1

(2.7)

Cette évolution peut être décrite par le graphe suivant



La matrice associé à ce graphe (en ignorant l'état D) est

$$A = \begin{pmatrix} 0.5 & 0.25 & 0.05 & 0 & 0.2 & 0 \\ 0.3 & 0.1 & 0.5 & 0 & 0.1 & 0 \\ 0.15 & 0 & 0 & 0.6 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad A^T = \begin{pmatrix} 0.5 & 0.3 & 0.15 & 0 & 0 & 0 \\ 0.25 & 0.1 & 0 & 0 & 0 & 0 \\ 0.05 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0 & 0 \\ 0.2 & 0.1 & 0.25 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Exercice 2.15.
Vérifier que les matrices A et A^T sont stochastiques (à droite et à gauche, respectivement). Si on suppose que les bourdons se comportent de façon markovienne, donner une distribution, minute par minute, des bourdons entre les différents états de la parade nuptiale. Pour cela, prenons X_0 le vecteur colonne avec 80 en première coordonnée et 0 ailleurs, et posons $X_{k+1} = A^T X_k$ pour $k \geq 0$. Alors X_k contient le nombre moyen de bourdons dans chaque état de la parade après k minutes.
Est-ce que les observations sont cohérentes avec les simulations? Pour cela il faut voir si le nombre total (sur toutes les étapes de l'algorithme) de bourdons observés dans chaque état est le même dans notre simulation que dans le tableau (2.7).

Exemple: Page rank algorithm. Comment décider quelles pages internet sont plus importantes que d'autres lors une recherche? Le contenu individuel de chaque page (par exemple le nombre de fois que le mot recherché y apparait) n'est évidement pas un bon indicateur. On peut par contre considérer que plus il y a de liens vers une page, plus cette page est importante. Cette observation est à l'origine des algorithmes utilisés par les moteurs de recherche.

Imaginons l'internet comme un graphe $G = (V, E)$ dirigé, chaque sommet étant une page internet, et chaque arête symbolisant un lien. Posons $n = |V|$ et donnons pour commencer une importance égale à chaque page:

$$X_0 = \begin{pmatrix} 1/n \\ \vdots \\ 1/n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}).$$

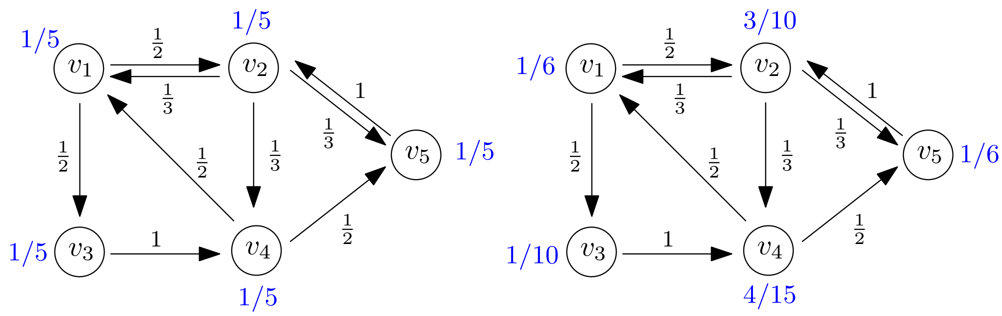


Figure 2.1: Un graphe avec les proportions transmise par chaque arête. A gauche une commence avec chaque sommet ayant une importance égale, à savoir 1/5. Après un pas de redistribution, on obtient X_1, décrit à droite.

Cette attribution d'importance n'est pas réaliste. On va donc considérer en première approximation, que chaque page est aussi importante qu'il y a d'arêtes pointant vers elle. On considère ainsi que chaque page v_i distribue son importance uniformément parmi ses voisins.

Notons $d(v_i)$ le nombre d'arêtes sortantes de v_i et posons

$$A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R}) \text{ avec } a_{i,j} = \begin{cases} \frac{1}{d(v_i)} & \text{si } (v_i, v_j) \in E \\ 0 & \text{sinon.} \end{cases}$$

La matrice A est stochastique à droite. L'importance de chaque sommet v_i est alors donné en cette première approximation par $X_1 = A^T X_0$.

Cette façon de classer les pages n'est pas parfaite non-plus: une page devrait être plus importante si d'autres pages importantes y sont reliées. Ainsi on pose

$$X_2 = A^T X_1,\ X_3 = A^T X_2,\ etc.$$

Prenons l'exemple du graphe de la figure 2.1. Dans ce cas la matrice A s'écrit

$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad A^T = \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Si on applique A^T de façon répétée on trouve

	X	$A^T X$	$(A^T)^2 X$	$(A^T)^3 X$	$(A^T)^4 X$	$(A^T)^5 X$	$(A^T)^6 X$	$(A^T)^7 X$	$(A^T)^8 X$
v_1	20%	17%	23%	18%	20%	21%	19%	20%	20%
v_2	20%	30%	25%	35%	28%	30%	31%	29%	31%
v_3	20%	10%	8%	12%	9%	10%	10%	10%	10%
v_4	20%	27%	20%	17%	23%	18%	20%	21%	19%
v_5	20%	17%	23%	18%	20%	21%	19%	20%	20%

2.3 Applications linéaires et matrices

Fixons pour l'intégralité de cette partie deux espaces vectoriels E, F de dimension finie n et m respectivement. De plus, soient $\mathcal{E} = (e_1, \dots, e_n)$ un base de E et $\mathcal{F} = (f_1, \dots, f_m)$ une base de F.

2.3.1 Représentation des applications linéaires par des matrices

On rappelle que tout vecteur $x \in E$ admet une écriture dans la base $\mathcal{E}$ qu'on a noté $V_{\mathcal{E}}(x)$. (Voir la partie 1.3.4.) De plus, on a fait la convention d'écrire $V_{\mathcal{E}}(x)$ verticalement, ainsi $V_{\mathcal{E}}(x) \in \mathcal{M}_{n,1}(\mathbb{R})$.

Dans la partie précédente on a vu que $\mathcal{M}_{n,1}(\mathbb{R})$ admet une structure d'espace vectoriel. Dans le langage des applications linéaires, la proposition 1.18 nous dit que la fonction $V_{\mathcal{E}} : E \to \mathcal{M}_{n,1}(\mathbb{R})$ est une application linéaire bijective.

Dans cette partie, on va voir que les applications linéaires admettent une représentation similaire par des matrices.

Définition 2.22. *Soit $u \in \mathcal{L}(E, F)$. La matrice $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ définie par*

$$\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} = V_{\mathcal{F}}\big(u(e_j)\big), \qquad \textit{pour tout } j \in [1, n], \tag{2.8}$$

est la matrice de u dans les bases $\mathcal{E}$, $\mathcal{F}$.

Le plus souvent on va traiter le cas des automorphisme de E. Dans ce cadre il est naturel d'utiliser la même base pour l'espace d'arrivée et de départ. Ainsi, pour $u \in \mathcal{L}(E)$, on écrit $\mathrm{Mat}_{\mathcal{E}}(u) = \mathrm{Mat}_{\mathcal{E},\mathcal{E}}(u) \in \mathcal{M}_n(\mathbb{R})$.

Théorème 2.23. *Soit $u \in \mathcal{L}(E, F)$ et $x \in E$. Alors*

$$V_{\mathcal{F}}\big(u(x)\big) = \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) V_{\mathcal{E}}(x). \tag{2.9}$$

De plus $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)$ est l'unique matrice qui a cette propriété.

Preuve: On va commencer par montrer (2.9) pour les vecteurs de la base $\mathcal{E}$. Soit e_j un vecteur de

$\mathcal{E}$. Alors

$$V_{\mathcal{E}}(e_j) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ou l'entrée 1 est sur la j^{eme} ligne. Ainsi,

$$\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) V_{\mathcal{E}}(e_j) = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} = V_{\mathcal{F}}\big(u(e_j)\big). \tag{2.10}$$

La première égalité vient de la règle de multiplication des matrices, la deuxième vient de la definition de $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)$

Montrons maintenant (2.9) pour un vecteur quelconque x. Soit $x \in E$. Comme $\mathcal{E}$ est une base de E, il existe des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $x = \lambda_1 e_1 + \dots + \lambda_n e_n$. Alors, par la linéarité de u, $u(x) = \lambda_1 u(e_1) + \dots + \lambda_n u(e_n)$. Ainsi

$$\begin{aligned}
&\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) V_{\mathcal{E}}(x) \\
&\quad = \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)\big(\lambda_1 V_{\mathcal{E}}(e_1) + \dots + \lambda_n V_{\mathcal{E}}(e_n)\big) && \text{par linéarité de } V_{\mathcal{E}} \\
&\quad = \lambda_1 \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) V_{\mathcal{E}}(e_1) + \dots + \lambda_n \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) V_{\mathcal{E}}(e_n) && \text{par linéarité de multip. matricielle} \\
&\quad = \lambda_1 V_{\mathcal{F}}\big(u(e_1)\big) + \dots + \lambda_n V_{\mathcal{F}}\big(u(e_j)\big) && \text{par (2.10)} \\
&\quad = V_{\mathcal{F}}\big(\lambda_1 u(e_1) + \dots + \lambda_n u(e_n)\big) && \text{par linéarité de } V_{\mathcal{F}} \\
&\quad = V_{\mathcal{F}}\big(u(\lambda_1 e_1 + \dots + \lambda_n e_n)\big) && \text{par linéarité de } u \\
&\quad = V_{\mathcal{F}}\big(u(x)\big).
\end{aligned}$$

En fin, observons que $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)$ est l'unique matrice qui satisfait (2.9) pour tout $x \in E$. En effet si on applique (2.9) au vecteur e_i de la base $\mathcal{E}$, on obtient (2.8). Ainsi (2.8) n'est qu'un cas particulier de (2.9). $\square$

Quand $E = F$ on utilise par défaut la même base pour l'espace de depart et celui d'arrivé. Ainsi, pour $u \in \mathcal{L}(E)$ et $\mathcal{E}$ une base de E, on écrit

$$\mathrm{Mat}_{\mathcal{E}}(u) = \mathrm{Mat}_{\mathcal{E},\mathcal{E}}(u).$$

Exemple: Considérons les transformations suivantes du plan $\mathbb{R}^2$: u est la rotation autour de 0 d'angle θ (pour un certain $\theta \in [0, 2\pi]$) et soit v la réflexion par rapport à la diagonale principale de $\mathbb{R}^2$ (voir fig. 2.2) Les deux sont des applications linéaires. On peut calculer leur matrices dans la base canonique $\mathcal{E}$ de $\mathbb{R}^2$, à savoir la base formée de $e_1 = (1, 0)$ et $e_2 = (0, 1)$. On a

$$u(e_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \cos\theta e_1 + \sin\theta e_2, \qquad u(e_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = -\sin\theta e_1 + \cos\theta e_2,$$
$$\text{et} \quad v(e_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e_2, \qquad v(e_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1.$$

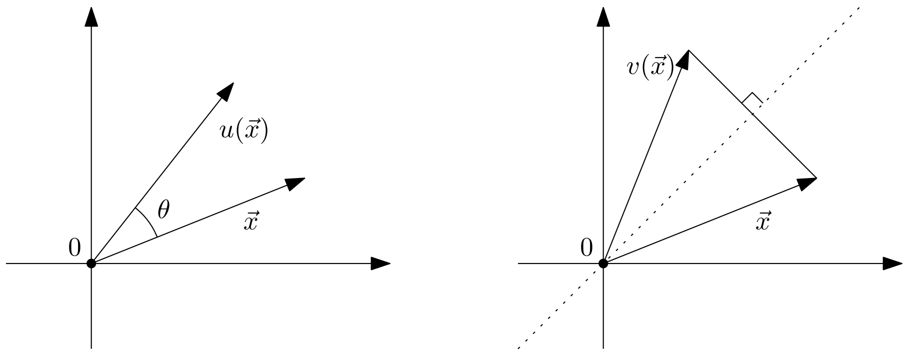


Figure 2.2: L'effet des transformations u et v.

Ainsi

$$\mathrm{Mat}_{\mathcal{E}}(u) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{E}}(v) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

On peut également essayer d'écrire les mêmes transformations dans une autre base. Posons $f_1 = e_1 + e_2$ et $f_2 = -e_1 + e_2$. Alors il est facile de voir que $\mathcal{F} = (f_1, f_2)$ est une base de $\mathbb{R}^2$. Un calcul rapide nous montre que

$$\mathrm{Mat}_{\mathcal{F}}(u) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{F}}(v) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Exercice 2.16.
Soient $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$. Montrer que si $AX = BX$ pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, alors $A = B$.

2.3.2 Addition et multiplication par un scalaire

On voit ici une parallèle se créer entre les application linéaires et les matrices. De plus, on a vu que les applications linéaires de $\mathcal{L}(E, F)$ et les matrices de $\mathcal{M}_{m,n}(\mathbb{R})$ ont toutes les deux une structure d'espace vectoriel. Il n'est pas surprenant que ces deux structures sont équivalentes.

Proposition 2.24. *L'application* $\mathrm{Mat}_{\mathcal{F},\mathcal{E}} : \mathcal{L}(E, F) \to \mathcal{M}_{m,n}(\mathbb{R})$ *est linéaire et bijective. Plus précisément, pour* $u, v \in \mathcal{L}(E, F)$ *et* $\lambda \in \mathbb{R}$*,*

(a) $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u + v) = \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) + \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(v)$,

(b) $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(\lambda u) = \lambda \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)$,

(c) $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) = 0$ *si et seulement si* $u = 0$,

(d) *pour toute matrice* $M \in \mathcal{M}_{m,n}(\mathbb{R})$*, il existe* $w \in \mathcal{L}(E, F)$ *telle que* $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(w) = M$.

Preuve: On va admettre la proposition. Mentionnons simplement quelle est la stratégie de preuve pour le point (a). Sachant que $\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)$ est l'unique matrice M qui satisfait

$$V_{\mathcal{F}}\big((u+v)(x)\big) = M\,V_{\mathcal{E}}(x) \qquad \forall x \in E,$$

il suffit de montrer que

$$V_{\mathcal{F}}\big((u+v)(x)\big) = \big[\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) + \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(v)\big]V_{\mathcal{E}}(x) \qquad \forall x \in E.$$

Cela peut être vérifier simplement. Les autres points sont traités de la même façon. □

Exercice 2.17.
De la proposition 2.24 et du corollaire 2.10 on déduit que $\mathcal{L}(E,F)$ et $\mathcal{M}_{m,n}(\mathbb{R})$ ont la même dimension. Trouver leur dimension et exhibant une base de $\mathcal{M}_{m,n}(\mathbb{R})$.

2.3.3 Composition vs. multiplication; inverse

Introduisons un troisième espace vectoriel G de dimension p et une base $\mathcal{G} = (g_1, \dots, g_p)$ de G.

Dans certaines situations, on peut composer des applications linéaires. De même, si elles ont la bonne taille, on peut multiplier deux matrices. Les deux opérations sont intimement liées.

Proposition 2.25. *Soit $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$. Alors,*

$$\mathrm{Mat}_{\mathcal{G},\mathcal{E}}(v \circ u) = \mathrm{Mat}_{\mathcal{G},\mathcal{F}}(v)\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u).$$

Preuve: Soit $x \in E$. Alors, par (2.9) appliqué trois fois,

$$\begin{aligned}\mathrm{Mat}_{\mathcal{G},\mathcal{F}}(v)\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)V_{\mathcal{E}}(x) &= \mathrm{Mat}_{\mathcal{G},\mathcal{F}}(v)V_{\mathcal{F}}\big(u(x)\big)\\ &= V_{\mathcal{G}}\big(v(u(x))\big)\\ &= \mathrm{Mat}_{\mathcal{G},\mathcal{E}}(v\circ u)V_{\mathcal{E}}(x).\end{aligned}$$

On a déjà mentionné que la matrice satisfaisant (2.9) est unique. Ainsi $\mathrm{Mat}_{\mathcal{G},\mathcal{F}}(v)\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u) = \mathrm{Mat}_{\mathcal{G},\mathcal{E}}(v \circ u)$. □

Dans le cadre de $\mathcal{L}(E)$ on a vu que certaines applications linéaires accepte des inverse. Cette notion est bien sur reliée à celle d'inverse de matrice.

Proposition 2.26. *Soit $u \in \mathcal{L}(E)$. Alors u est inversible si et seulement si $\mathrm{Mat}_{\mathcal{E}}(u)$ est inversible. Dans le cas ou les deux sont inversibles, on a*

$$\mathrm{Mat}_{\mathcal{E}}\big(u^{-1}\big) = \mathrm{Mat}_{\mathcal{E}}(u)^{-1}.$$

Preuve: Supposons pour commencer que $\mathrm{Mat}_{\mathcal{E}}(u)$ est inversible. Comme $\mathrm{Mat}_{\mathcal{E}} : \mathcal{L}(E) \to \mathcal{M}_n(\mathbb{R})$ est bijective, il existe $v \in \mathcal{L}(E)$ telle que

$$\mathrm{Mat}_{\mathcal{E}}(v) = \mathrm{Mat}_{\mathcal{E}}(u)^{-1}.$$

Alors

$$\mathrm{Mat}_{\mathcal{E}}(v \circ u) = \mathrm{Mat}_{\mathcal{E}}(v)\mathrm{Mat}_{\mathcal{E}}(u) = I_m = \mathrm{Mat}_{\mathcal{E}}(u)\mathrm{Mat}_{\mathcal{E}}(v) = \mathrm{Mat}_{\mathcal{E}}(u \circ v).$$

En utilisant encore une fois le fait que $\mathrm{Mat}_{\mathcal{E}}$ est bijective, on déduit $v \circ u = u \circ v = \mathrm{id}$. On conclut donc que u est bien inversible et que $v = u^{-1}$.

Inversement, si on suppose que u est inversible, alors

$$\begin{aligned}\mathrm{Mat}_{\mathcal{E}}\big(u^{-1}\big)\mathrm{Mat}_{\mathcal{E}}\big(u\big) &= \mathrm{Mat}_{\mathcal{E}}\big(u^{-1} \circ u\big)\\ &= \mathrm{Mat}_{\mathcal{E}}(\mathrm{id}) = I_m = \mathrm{Mat}_{\mathcal{E}}\big(u \circ u^{-1}\big)\\ &= \mathrm{Mat}_{\mathcal{E}}\big(u\big)\mathrm{Mat}_{\mathcal{E}}\big(u^{-1}\big).\end{aligned}$$

Donc $\mathrm{Mat}_{\mathcal{E}}(u)$ est inversible et

$$\mathrm{Mat}_{\mathcal{E}}\big(u^{-1}\big) = \mathrm{Mat}_{\mathcal{E}}(u)^{-1}.$$

□

2.3.4 Image, noyaux et rang des matrices

En pratique on travail souvent avec la forme matricielle des applications linéaires. Dans le cas dégénéré ou les applications linéaires ne sont pas inversibles, il est utile d'identifier le noyaux, l'image et le rang de l'application linéaire à partir de sa forme matricielle.

Commençons par définir l'image et le noyaux d'une matrice $M \in \mathcal{M}_{m,n}(\mathbb{R})$.

Définition 2.27. *On pose*

$$\begin{aligned}\mathrm{Im}(M) &= \big\{X \in \mathcal{M}_{m,1}(\mathbb{R}) : \exists Y \in \mathcal{M}_{m,1} \ t.q.\ MY = X\big\} \subset \mathbb{R}^m,\\ \mathrm{Ker}(M) &= \big\{X \in \mathcal{M}_{n,1}(\mathbb{R}) : MX = 0\big\} \subset \mathbb{R}^n,\\ \mathrm{rang}(M) &= \dim(\mathrm{Im}(M)).\end{aligned}$$

Remarque 2.28. On peut facilement vérifier que $\mathrm{Im}(M)$ et $\mathrm{Ker}(M)$ sont des s.e.v. de $\mathbb{R}^m$ et $\mathbb{R}^n$, respectivement.

Les définition de image, noyaux et rang pour les matrices sont très similaires à ceux pour les applications linéaire. Ce n'est pas par hasard, ce sont deux façons de voir le même objet.

Proposition 2.29. *Soit $u \in \mathcal{L}(E, F)$. Alors,*

$$\mathrm{Im}(\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)) = V_{\mathcal{F}}(\mathrm{Im}(u)) \quad \textit{et} \quad \mathrm{Ker}(\mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)) = V_{\mathcal{E}}(\mathrm{Ker}(u)).$$

On laisse la preuve en exercice.
Un corollaire immédiat est la version du théorème du rang pour les matrices.

Corollaire 2.30. *Soit $M \in \mathcal{M}_{m,n}(\mathbb{R})$, alors*

$$\operatorname{rang}(M) + \dim(\operatorname{Ker}(M)) = n.$$

Pour les matrices, le rang peut se calculer de plusieurs manières, comme l'illustre la proposition suivante.

Proposition 2.31.
Soit $M \in \mathcal{M}_{m,n}(\mathbb{R})$, $M = \begin{pmatrix} C_1 \dots C_n \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}$. Alors $\operatorname{Im}(M) = \operatorname{Vect}(C_1, \dots, C_n)$ *et*

$$\operatorname{rang}(M) = \operatorname{rang}(C_1, \dots, C_n) = \operatorname{rang}(L_1, \dots, L_m).$$

Preuve: On commence par montrer $\operatorname{Im}(M) = \operatorname{Vect}(C_1, \dots, C_n)$.
Pour $i \in \{1, \dots, n\}$, soit $Y_i \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur contenant un 1 sur la ligne i et 0 partout ailleurs. Alors $Y_1, \dots, Y_n$ est la base canonique de $\mathbb{R}^n$.

Une simple application de la règle de multiplication des matrices montre que

$$MY_i = C_i, \qquad \text{pour tout } 1 \leq i \leq n.$$

Ainsi $C_1, \dots, C_n \in \operatorname{Im}(M)$. Comme $\operatorname{Im}(M)$ est un espace vectoriel, $\operatorname{Vect}(C_1, \dots, C_n) \subset \operatorname{Im}(M)$.

Inversement, soit $X \in \operatorname{Im}(M)$ et $Z \in \mathcal{M}_{n,1}(\mathbb{R})$ avec $X = MZ$. Si on note $z_1, \dots, z_n$ les coefficients de Z, à savoir $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, alors $Z = z_1 Y_1 + \dots + z_n Y_n$. Ainsi

$$\begin{aligned} X = MZ &= M(z_1 Y_1 + \dots + z_n Y_n) \\ &= z_1 MY_1 + \dots + z_n MY_n = z_1 C_1 + \dots + z_n C_n \in \operatorname{Vect}(C_1, \dots, C_n). \end{aligned}$$

On a donc montré que $\operatorname{Im}(M) = \operatorname{Vect}(C_1, \dots, C_n)$.

Une conséquence directe est

$$\operatorname{rang}(M) = \dim(\operatorname{Im}(M)) = \dim(\operatorname{Vect}(C_1, \dots, C_n)) = \operatorname{rang}(C_1, \dots, C_n).$$

La dernière égalité (à savoir $\operatorname{rang}(C_1, \dots, C_n) = \operatorname{rang}(L_1, \dots, L_m)$) est plus délicate et nécessite une construction qu'on ne traite pas dans ce cours. On va l'admettre. □

Une conséquence immédiate de la proposition 2.29 et le critère suivant pour l'invisibilité des matrices.

Corollaire 2.32. *Soit $M \in \mathcal{M}_n(\mathbb{R})$. Alors on a équivalence de*
- (i) *la famille des colonnes $C_1, \dots, C_n$ est libre,*
- (ii) *la famille des lignes $L_1, \dots, L_n$ est libre,*
- (iii) *la matrice M est inversible.*

Preuve: La matrice M est inversible si et seulement si $\mathrm{rang}(M) = n$. Par la proposition précédente, cela revient à $\mathrm{rang}(C_1, \dots, C_n) = \mathrm{rang}(L_1, \dots, L_n) = n$. Mais cela est équivalent au fait que les familles $(C_1, \dots, C_n)$ et $(L_1, \dots, L_n)$ sont libres. □

Lemme 2.33. *Soit $M \in \mathcal{M}_{m,n}(\mathbb{R})$, $P \in \mathcal{M}_m(\mathbb{R})$ et $Q \in \mathcal{M}_n(\mathbb{R})$. Alors*
- (i) *si P est inversible, alors* $\mathrm{Ker}(M) = \mathrm{Ker}(PM)$ *et* $\mathrm{rang}(M) = \mathrm{rang}(PM)$*;*
- (ii) *si Q est inversible, alors* $\mathrm{Im}(M) = \mathrm{Im}(MQ)$ *et* $\mathrm{rang}(M) = \mathrm{rang}(MQ)$.

Ce qu'il faut retenir de ce lemme est que le rang d'une matrice n'est pas modifié si on la multiplie, à gauche ou à droite, par des matrices inversibles.

On laisse la preuve en exercice.

Exercice 2.18.
Démontrer la proposition 2.29. (La relation (2.9) peux être utile.)

Exercice 2.19.
Démonter le lemme 2.33.

Exercice 2.20.
Soient $A, B \in \mathcal{M}_n(\mathbb{R})$ tels que $AB = I_n$. Montrer que A est inversible et que $A^{-1} = B$.
Trouver un couple de matrices $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}(n,m)(\mathbb{R})$ pour $m < n$, tels que $AB = I_m$. Calculer BA. Que dire des images et noyaux de A et B?

2.3.5 Changement de base

En pratique il est des fois nécessaire de travailler avec plusieurs bases d'un même espace vectoriel (voir par exemple les suites définies par récurrence traités dans la partie 1.4, ou encore la diagonalisation des matrices traitée dans la partie 3). Ainsi, il est important d'avoir un outil qui permet d'obtenir l'écriture d'un vecteur dans une base à partir de son écriture dans une autre base. La solution est donnée par les matrices de changement de base.

Fixons E un espace vectoriel de dimension finie n et $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{E} = (e_1, \dots, e_n)$ deux bases de E.

Définition 2.34. *La* matrice de changement de base *de $\mathcal{E}$ à $\mathcal{B}$ est la matrice notée* $\mathrm{P}_{\mathcal{B},\mathcal{E}} =$

$(p_{ij}) \in \mathcal{M}_n(\mathbb{R})$ *dont les entrées sont données colonne par colonne par*

$$C_i = \begin{pmatrix} p_{1,i} \\ \vdots \\ p_{n,i} \end{pmatrix} = V_{\mathcal{B}}(e_i) \quad \text{pour tout } i \in \{1, \dots, n\}. \tag{2.11}$$

Remarque 2.35. On peut voir la matrice de changement de base comme la matrice de l'application linéaire identité id $\in \mathcal{L}(E)$:

$$\mathrm{P}_{\mathcal{B},\mathcal{E}} = \mathrm{Mat}_{\mathcal{B},\mathcal{E}}(\mathrm{id}). \tag{2.12}$$

Proposition 2.36. *Pour tout $x \in E$, $V_{\mathcal{B}}(x) = \mathrm{P}_{\mathcal{B},\mathcal{E}} V_{\mathcal{E}}(x)$.*

Preuve: La démonstration découle directement de (2.12) et du théorème 2.23. □

On aimerai dire que le changement de base de $\mathcal{B}$ à $\mathcal{E}$ est l'inverse de celui de $\mathcal{E}$ à $\mathcal{B}$. La proposition suivante en donne le sens précis.

Proposition 2.37. *La matrice $\mathrm{P}_{\mathcal{B},\mathcal{E}}$ est inversible et $\mathrm{P}_{\mathcal{B},\mathcal{E}}^{-1} = \mathrm{P}_{\mathcal{E},\mathcal{B}}$.*

Preuve: En appliquant la proposition 2.36 deux fois, on obtient que pour tout $x \in E$,

$$V_{\mathcal{B}}(x) = \mathrm{P}_{\mathcal{B},\mathcal{E}} V_{\mathcal{E}}(x) = \mathrm{P}_{\mathcal{B},\mathcal{E}} \mathrm{P}_{\mathcal{E},\mathcal{B}} V_{\mathcal{B}}(x).$$

Le même calcul s'applique à $\mathrm{P}_{\mathcal{E},\mathcal{B}} \mathrm{P}_{\mathcal{B},\mathcal{E}}$. Comme $V_{\mathcal{B}}$ et$V_{\mathcal{E}}$ sont surjectives, on déduit que

$$\mathrm{P}_{\mathcal{B},\mathcal{E}} \mathrm{P}_{\mathcal{E},\mathcal{B}} X = \mathrm{P}_{\mathcal{E},\mathcal{B}} \mathrm{P}_{\mathcal{B},\mathcal{E}} X = X, \quad \text{pour tout } X \in \mathcal{M}_{n,1}(\mathbb{R}).$$

Il s'en suit (voir par exemple l'exercice 2.16) que $\mathrm{P}_{\mathcal{B},\mathcal{E}} \mathrm{P}_{\mathcal{E},\mathcal{B}} = \mathrm{P}_{\mathcal{E},\mathcal{B}} \mathrm{P}_{\mathcal{B},\mathcal{E}} = I_n$, donc que $\mathrm{P}_{\mathcal{B},\mathcal{E}} = \mathrm{P}_{\mathcal{E},\mathcal{B}}^{-1}$. □

De plus, il est facile de voir que toute matrice inversible est une matrice de changement de base. En effet, si $P \in \mathcal{M}_n(\mathbb{R})$ est inversible et $\mathcal{B}$ est une base de E, il suffit de créer la base $\mathcal{E}$ par la formule (2.11), et on obtient $P = \mathrm{P}_{\mathcal{B},\mathcal{E}}$.

Exercice 2.21.
Soient E un espace vectoriel de dimension n et $\mathcal{E}$ une base de E. Montrer que pour toute matrice inversible $M \in \mathcal{M}_n(\mathbb{R})$ il existe une base $\mathcal{B}$ de E telle que $M = \mathrm{P}_{\mathcal{B},\mathcal{E}}$.

2.3.6 Matrices semblables

Définition 2.38. *Soit $A, B \in \mathcal{M}_n(\mathbb{R})$ deux matrices. On dit qu'elles sont* semblables *s'il existe $P \in \mathcal{M}(\mathbb{R})$ inversible telle que*

$$B = P^{-1}AP. \tag{2.13}$$

La notion de matrices semblable est motivée par la représentation des automorphisme par les matrices carrées. Pour illustrer cela, on commence par une proposition qui donnent la formule essentielle de changement de base pour des automorphisme.

Proposition 2.39. *Soient $\mathcal{E}$ et $\mathcal{B}$ deux bases d'un espace vectoriel de dimension finie E et soit $u \in \mathcal{L}(E)$. Alors,*

$$\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{P}_{\mathcal{E},\mathcal{B}}^{-1}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{P}_{\mathcal{E},\mathcal{B}}. \tag{2.14}$$

Preuve: Notons n la dimension de E et soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Alors, par la surjectivité de $\mathrm{Vect}_{\mathcal{B}}$, il existe $x \in E$, tel que $X = \mathrm{Vect}_{\mathcal{B}}(x)$. Ainsi

$$\begin{aligned}
\mathrm{Mat}_{\mathcal{B}}(u)X &= \mathrm{Mat}_{\mathcal{B}}(u)\mathrm{Vect}_{\mathcal{B}}(x)\\
&= \mathrm{Vect}_{\mathcal{B}}(u(x))\\
&= \mathrm{P}_{\mathcal{B},\mathcal{E}}\mathrm{Vect}_{\mathcal{E}}(u(x))\\
&= \mathrm{P}_{\mathcal{B},\mathcal{E}}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{Vect}_{\mathcal{E}}(x)\\
&= \mathrm{P}_{\mathcal{B},\mathcal{E}}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{P}_{\mathcal{E},\mathcal{B}}\mathrm{Vect}_{\mathcal{B}}(x)\\
&= \mathrm{P}_{\mathcal{E},\mathcal{B}}^{-1}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{P}_{\mathcal{E},\mathcal{B}}X.
\end{aligned}$$

On a donc prouvé que $\mathrm{Mat}_{\mathcal{B}}(u)X = \mathrm{P}_{\mathcal{E},\mathcal{B}}^{-1}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{P}_{\mathcal{E},\mathcal{B}}X$ pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Par l'exercice 2.16, ceci implique l'égalité des deux matrices. □

Soit $A \in \mathcal{M}_n(\mathbb{R})$, E un espace vectoriel de dimension n et $\mathcal{E}$ une base de E. Il existe alors une unique application linéaire $u \in \mathcal{L}(E)$ telle que $\mathrm{Mat}_{\mathcal{E}}(u) = A$ (par bijectivité de $\mathrm{Mat}_{\mathcal{E}}$). Une conséquence immédiate du lemme précédent est la suivante.

Proposition 2.40. *Une matrice $B \in \mathcal{M}_n(\mathbb{R})$ est semblable à A si et seulement si B est la représentation de u dans une base $\mathcal{B}$ de E. De plus, la matrice P de* (2.13) *est alors $\mathrm{P}_{\mathcal{B},\mathcal{E}}$.*

Preuve: Si $B = \mathrm{Mat}_{\mathcal{B}}(u)$ pour une certaine base $\mathcal{B}$ de E, alors, par la proposition 2.39, $B = \mathrm{P}_{\mathcal{B},\mathcal{E}}A\mathrm{P}_{\mathcal{B},\mathcal{E}}^{-1}$.

Inversement, supposons que $B = PAP^{-1}$ pour une certaine matrice inversible P. Alors il existe une base $\mathcal{B}$ de E telle que $P = \mathrm{P}_{\mathcal{B},\mathcal{E}}$ (voir l'exercice 2.21). Ainsi, par la proposition 2.39,

$$B = \mathrm{P}_{\mathcal{B},\mathcal{E}}A\mathrm{P}_{\mathcal{B},\mathcal{E}}^{-1} = \mathrm{Mat}_{\mathcal{B}}(u).$$

□

2.4 Systèmes linéaires

Définition 2.41. *Soient $m, n \in \mathbb{N}$ et $(a_{i,j})_{\substack{1\le i\le m\\ 1\le j\le n}}$, $(b_i)_{1\le i\le m}$ deux familles de scalaires (c.à d. des éléments de $\mathbb{R}$). Le système linéaire de m équations à n inconnues $x_1, \dots, x_n$ avec les coefficients $(a_{i,j})_{i,j}$, $(b_i)_i$ est l'ensemble d'équations*

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ \dots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m. \end{cases} \tag{2.15}$$

Une solution du système est une famille $(x_1, \dots, x_n) \in \mathbb{R}^n$ qui satisfait toutes les équations de (2.15). On appelle $\mathcal{S} \subset \mathbb{R}^n$ l'ensemble des solutions de (2.15).

On peut écrire le système (2.15) sous forme matricielle de la façon suivante. Soient

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}) \quad \text{et} \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathcal{M}_{m,1}(\mathbb{R}).$$

Alors, (2.15) devient

$$AX = B, \tag{2.16}$$

ou $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est le vecteur des inconnues $x_1, \dots, x_n$. On peut ainsi écrire

$$\mathcal{S} = \left\{ X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}) : \ AX = B \right\}.$$

Ainsi, $\mathcal{S} \neq \emptyset$ (c.à d. que le système admet au moins une solution) si et seulement si $B \in \mathrm{Im}(A)$. Supposons que $B \in \mathrm{Im}(A)$, et soit $X_0 \in \mathcal{M}_{n,1}(\mathbb{R})$ une solution. Alors, pour toute solution $X \in \mathcal{S}$,

$$A(X - X_0) = AX - AX_0 = B - B = 0,$$

donc $X - X_0 \in \mathrm{Ker}(A)$. Inversement, si $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est tel que $X - X_0 \in \mathrm{Ker}(A)$, alors le calcul précédent montre que $X \in \mathcal{S}$. On arrive ainsi à la conclusion suivante.

Théorème 2.42. *Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{m,1}(\mathbb{R})$. On note $\mathcal{S}$ l'ensemble des solutions $X \in \mathcal{M}_{n,1}(\mathbb{R})$ de $AX = B$. Alors*

(i) *si $B \notin \mathrm{Im}(A)$, alors $\mathcal{S} = \emptyset$,*

(ii) *si $B \in \mathrm{Im}(A)$, alors $\mathcal{S} \neq \emptyset$. De plus, si X_0 est une solution, on a*

$$\mathcal{S} = X_0 + \mathrm{Ker}(A) = \{X_0 + Y : Y \in \mathrm{Ker}(A)\}.$$

Dans le cas *(ii)* on dit que X_0 est une *solution particulière* du système et que Y est une *solution générale du système homogène.* Le fait que $Y \in \mathrm{Ker}(A)$ s'écrit aussi $AY = 0$, ou encore,

$$\begin{cases} a_{1,1}y_1 + \cdots + a_{1,n}y_n = 0 \\ \cdots \\ a_{m,1}y_1 + \cdots + a_{m,n}y_n = 0. \end{cases}$$

On appelle ce système le *système homogène* associé à (2.15).

Rappelons nous que $\mathrm{Ker}(A)$ est un s.e.v. de $\mathbb{R}^n$. Ainsi, quand le système admet des solutions, $\mathcal{S}$ est le translaté d'un espace vectoriel par un vecteur X_0. On appelle ce type d'espace un *espace affine* et on pose

$$\dim(\mathcal{S}) = \dim(\mathrm{Ker}(A)).$$

Une conséquence du théorème du rang qui peut être utile est que $\dim(\mathrm{Ker}(A)) = n - \mathrm{rang}(A)$.

Exercice 2.22.
Soit $m \leq n$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$. Supposons que la famille des lignes $L_1, \ldots, L_m$ de A est libre. Montrer que, pour tout $B \in \mathcal{M}_{m,1}(\mathbb{R})$, l'équation $AX = B$ admet des solutions $X \in \mathcal{M}_{n,1}(\mathbb{R})$, et que l'ensemble des solutions est un espace affine de dimension $n - m$.

Exercice 2.23.
Soit $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{R})$ et $B \in \mathcal{M}_{n,1}(\mathbb{R})$ tels que l'équation $AX = B$ admet une unique solution $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Montrer que pour tout $B' \in \mathcal{M}_{n,1}(\mathbb{R})$, l'équation $AX = B'$ admet une unique solution $X \in \mathcal{M}_{n,1}(\mathbb{R})$.

2.4.1 Opérations sur les lignes; forme échelonnée

Notons les équations du système comme suit

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 & (L_1) \\ \cdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m. & (L_m) \end{cases} \tag{2.17}$$

Alors pour $i \neq j$ et pour $\lambda \in \mathbb{R}$, on peut rajouter à l'équation (L_i) l'équation (L_j) multipliée par λ, sans changer l'ensemble des solutions $\mathcal{S}$. On dit que le système (2.15) est equivalent au système

$$\begin{cases} \begin{array}{rcllll} a_{1,1}x_1+ & \ldots & +a_{1,n}x_n & = b_1 & & (L_1) \\ \vdots & \vdots & & \vdots & & \vdots \\ (a_{i,1} + \lambda a_{j,1})x_1+ & \ldots & +(a_{i,n} + \lambda a_{j,n})x_n & = b_i + \lambda b_j & & (L_i) + \lambda(L_j) \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m,1}x_1+ & \ldots & +a_{m,n}x_n & = b_m. & & (L_m) \end{array} \end{cases} \tag{2.18}$$

On peut également échanger deux équations du système entre elles et multiplier une équation par un scalaire $\lambda \neq 0$, sans que l'ensemble des solutions change.

Vu l'écriture plus compacte en utilisant les matrices, on va désormais écrire les systèmes linéaires sous forme matricielle. On identifie trois opérations pour les matrices:

- $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$ consiste à rajouter à la ligne i la ligne j multipliée par λ (ici $i \neq j$);
- $\mathrm{Mult}_{\lambda,\mathrm{i}}$ consiste à multiplier la ligne i par λ (ici $\lambda \neq 0$);
- $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$ consiste à échanger les lignes i et j (ici $i \neq j$).

Pour une matrice A, on va écrire $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}\mathrm{A}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}\mathrm{A}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}\mathrm{A}$ pour la matrice obtenue à partir de A par les opérations $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$, respectivement. Le fait que ces opérations ne changent pas les solutions du système s'écrit de la manière suivante.

Proposition 2.43. *Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{m,1}(\mathbb{R})$ et $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Pour $i \neq j$ et $\lambda \in \mathbb{R}$ avec $\lambda \neq 0$ on a équivalence de*

- $AX = B$,
- $(\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}\mathrm{A})\mathrm{X} = \mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}\mathrm{B}$,
- $(\mathrm{Mult}_{\lambda,\mathrm{i}}\mathrm{A})\mathrm{X} = \mathrm{Mult}_{\lambda,\mathrm{i}}\mathrm{B}$,
- $(\mathrm{Ech}_{\mathrm{i};\mathrm{j}}\mathrm{A})\mathrm{X} = \mathrm{Ech}_{\mathrm{i};\mathrm{j}}\mathrm{B}$.

Preuve: Les opérations $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$ consistent à multiplier à droite A par les matrices suivantes de $\mathcal{M}_m(\mathbb{R})$:

$$\begin{array}{cc} \begin{array}{c} i \\ \downarrow \end{array} & \\ \begin{pmatrix} 1 & & & & \\ & 1 & 0 & \lambda & 0 \\ & & \ddots & & 0 \\ & 0 & & 1 & \\ & & & & 1 \end{pmatrix} & \leftarrow j \end{array} \qquad \begin{array}{cc} \begin{array}{c} i \\ \downarrow \end{array} & \\ \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 \\ 0 & & & & \ddots & \\ & & & & & 1 \end{pmatrix} & \leftarrow i \end{array} \qquad \begin{array}{cc} \begin{array}{c} i \quad j \\ \downarrow \quad \downarrow \end{array} & \\ \begin{pmatrix} 1 & & & & & \\ & \ddots & & 0 & & 0 \\ & & 1 & & & \\ & & & 0 & & 1 \\ 0 & & & & 1 & & 0 \\ & & & & & \ddots & \\ & & & & & & 1 \\ & & & 1 & & & & 0 \\ 0 & & 0 & & & & & 1 \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} & \begin{array}{c} \leftarrow i \\ \\ \leftarrow j \end{array} \end{array}$$

On notera ces trois matrices aussi par $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$, respectivement. Ainsi la notation $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}\mathrm{A}$ désigne simplement le produit matriciel. De plus ces trois matrices sont inversibles, leurs inverses étant $\mathrm{Add}_{\mathrm{i};-\lambda,\mathrm{j}}$, $\mathrm{Mult}_{1/\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$, respectivement.

Soit $P \in \mathcal{M}_m(\mathbb{R})$ inversible. Alors si $AX = B$ on a évidement aussi $PAX = PB$. Inversement, si $PAX = PB$, alors, si on multiplie à droite par P^{-1} on obtient $AX = B$. On a donc prouvé que $AX = B$ si et seulement si $PAX = PB$.

En appliquant cette observation aux matrices inversibles $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i};\mathrm{j}}$ on obtient l'équivalence des affirmations. □

Définition 2.44. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$. On dit que A est* échelonnée selon les lignes[(v)] *si A est de la forme*

$$A = \begin{pmatrix}
0 \dots 0\, 1 * \dots * & 0 * \dots * & 0 & \dots * & 0 * \dots * \\
0 \quad \dots \quad \dots & 0\, 1 * \dots * & 0 & \dots * & 0 * \dots * \\
0 \quad \dots \quad \dots & \dots \quad \dots & 0\, 1 & \dots * & 0 * \dots * \\
\vdots & & & \ddots \quad \vdots & \vdots \\
0 \quad \dots \quad \dots & \dots \quad \dots & \dots & \dots & 0\, 1 * \dots * \\
0 \quad \dots \quad \dots & \dots \quad \dots & \dots & \dots & 0 \quad \dots \quad 0 \\
\vdots & & & & \vdots \qquad \vdots \\
0 \quad \dots \quad \dots & \dots \quad \dots & \dots & \dots & 0 \quad \dots \quad 0
\end{pmatrix}$$

Dans la représentation de la matrice échelonnée, les $*$ représentent des nombres quelconques.

Les matrices échelonnées sont particulièrement faciles à étudier. Par exemple, si $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est échelonnée avec k lignes non nulles, alors $\text{rang}(A) = k$ et par conséquence $\dim(\text{Ker}(A)) = n - k$. De plus, la résolution de $AX = B$ est immédiate (voir la partie suivante). Il est donc intéressant de transformer une matrice quelconque A en une matrice échelonnée en utilisant les opérations sur les lignes décrites ci-dessus.

Théorème 2.45. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$. Alors, en lui appliquant une série d'opérations sur les lignes du type $\text{Add}_{\text{i};\lambda,\text{j}}$, $\text{Ech}_{\text{i;j}}$ et $\text{Mult}_{\lambda,\text{i}}$, on peut la transformer en une matrice échelonnée $\tilde{A}$.*

La preuve du théorème nous donne aussi l'algorithme à suivre pour obtenir la forme échelonnée de A. Cet algorithme, appelé le *pivot de Gauss*(vii), est présenté dans la partie suivante dans le cadre des systèmes linéaires.

Soient $P_1, \dots, P_k$ les transformations sur les lignes appliquées à A pour arriver à la matrice échelonnée $\tilde{A}$. La preuve de la proposition 2.43 nous dit que $P_1, \dots, P_k \in \mathcal{M}_m(\mathbb{R})$ sont inversibles et que
$$\tilde{A} = P_k \dots P_1 A.$$
Ainsi, si $B \in \mathcal{M}_{m,1}(\mathbb{R})$, alors
$$AX = B \quad \text{si et seulement si} \quad \tilde{A}X = P_k \dots P_1 B.$$
En outre, comme $P_k \dots P_1$ est inversible, $\text{rang}(A) = \text{rang}(\tilde{A})$.

Attention! L'application des transformations dans l'ordre $P_1, \dots, P_k$ correspond à la multiplication de A à gauche par P_1, puis par P_2 etc. Cela correspond bien à $P_k \dots P_1 A$.

2.4.2 Pivot de Gauss pour la résolution des systèmes linéaires

Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{m,1}(\mathbb{R})$. Plutôt que de résoudre l'équation $AX = B$, on va transformer A en une matrice échelonnée $\tilde{A}$ par des opérations $P_1, \dots, P_k$ sur les lignes, pour ensuite résoudre l'équation équivalente $\tilde{A}X = P_k \dots P_1 B$.

(vi) Reduzierte Stufenform

(vii) Gaußsches Eliminationsverfahren

Vu que les opérations $\mathrm{Add_{i;\lambda,j}}$, $\mathrm{Ech_{i;j}}$ et $\mathrm{Mult_{\lambda,i}}$ doivent être appliquées à A et B simultanément, il va être plus commode d'écrire A et B ensemble sous la forme réduite suivante.

$$(A|B) = \left(\begin{array}{ccc|c} a_{1,1} & \dots & a_{1,n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,n} & b_m \end{array}\right) \tag{2.19}$$

On va procéder de façon itérative, suivant les colonnes. Supposons qu'on a une matrice A dont les $j-1$ premières colonnes forment une matrice échelonnée. Soit $k-1 \leq j-1$ le nombre des lignes non nulles de la matrice formée des les $j-1$ premières colonnes. (Au pas initial, on prend $j=1$ et $k=1$ et on applique la procédure décrite ci-dessous.)

On s'occupe de la colonne j, plus précisément de l'entrée $a_{k,j}$. On distingue plusieurs cas:

(i) Si $a_{k,j} = 1$, alors on soustrait la ligne k multiplié par $a_{i,j}$ à la ligne i pour chaque $i = 1, \dots, k-1, k+1, \dots, n$. Plus précisément on applique $\prod_{i \neq k} \mathrm{Add_{i;-a_{i,j},k}}$ à A. Ainsi on obtient une matrice qui contient un coefficient 1 sur la position k, j et des 0 sur le reste de la colonne j. Les colonnes $1, \dots, j-1$ ne sont pas affectées par ces transformations car la ligne k a ses premières $j-1$ entrées égales à 0. On conclut que les j premières colonnes de la matrice ainsi obtenue sont échelonnées; on peut donc passer à la colonne suivante.

(ii) Si $a_{k,j} \neq 0$ est une valeur quelconque, alors on divise la ligne k par $a_{k,j}$ (c.à-d. on applique $\mathrm{Mult_{k,\frac{1}{a_{k,j}}}}$ à A) et on obtient ainsi une matrice comme celle traitée au point (i). On continue en appliquant le point (i) à $\mathrm{Mult_{k,\frac{1}{a_{k,j}}}A}$.

(iii) Si $a_{k,j} = 0$ mais il existe $\ell > k$ tel que $a_{\ell,j} \neq 0$, alors on échange les lignes k et ℓ (c.à-d. on applique $\mathrm{Ech_{k,\ell}}$ à A) et on se ramène ainsi à une matrice comme celle traité au point (ii). On continue en appliquant le point (ii) à $\mathrm{Ech_{k,\ell}A}$

(iv) Si $a_{k,j} = a_{k+1,j} = \dots = a_{m,j} = 0$, alors la matrice formée des j premières lignes est déjà sous forme échelonnée et on peut passer à la colonne suivante.

L'algorithme fini quand $j = n+1$ ou $k = m+1$. On vérifie facilement que dans les deux cas, la matrice résultante est bien sous forme échelonnée. Un exemple est donné plus bas.

Attention! Pendant tout l'algorithme, les opérations appliquées à A doivent aussi être appliquées à B. Pourtant, on n'essaye pas de mettre B sous forme échelonnée. L'ordre d'application des transformations est important!

Résoudre un système sous forme échelonnée.

Supposons maintenant qu'on veut résoudre un système déjà écrit sous forme échelonnée:

$$\begin{array}{c} \qquad\qquad \overset{j_1}{\downarrow} \qquad\quad \overset{j_1}{\downarrow} \qquad\quad \overset{j_3}{\downarrow} \qquad\quad \overset{j_k}{\downarrow} \qquad\qquad\quad \\ (A|B) = \left(\begin{array}{cccccccccccccccccc|c} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & \dots & * & 0 & * & \dots & * & b_1 \\ 0 & \dots & & \dots & & & 0 & 1 & * & \dots & * & 0 & \dots & * & 0 & * & \dots & * & b_2 \\ 0 & \dots & & \dots & & \dots & & \dots & & & 0 & 1 & \dots & * & 0 & * & \dots & * & b_3 \\ \vdots & & & & & & & & & & & & \ddots & & & \vdots & & \vdots & \vdots \\ 0 & \dots & & \dots & & \dots & & \dots & & \dots & & & \dots & 0 & 1 & * & \dots & * & b_k \\ 0 & \dots & & \dots & & \dots & & \dots & & \dots & & & \dots & & 0 & & \dots & 0 & b_{k+1} \\ \vdots & & & & & & & & & & & & & & \vdots & & & \vdots & \vdots \\ 0 & \dots & & \dots & & \dots & & \dots & & \dots & & & \dots & & 0 & & \dots & 0 & b_m \end{array}\right) \end{array}$$

ou $j_1 < \dots < j_k$ sont les colonnes contentant les premiers 1 de chaque ligne. Alors $X \in \mathcal{M}_{n,1}$ est solution du système si et seulement si

$$AX = \begin{pmatrix} x_{j_1} + \sum_{j>j_1} a_{1,j}x_j \\ \vdots \\ x_{j_k} + \sum_{j>j_k} a_{k,j}x_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_k \\ b_{k+1} \\ \vdots \\ b_m \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \text{ligne } k \\ \\ \\ \\ \end{matrix} \tag{2.20}$$

Ainsi, le système admet des solutions que si $b_{k+1} = \dots = b_m = 0$.

Supposons que $b_{k+1} = \dots = b_m = 0$. Alors les solutions X de $AX = B$ sont obtenues comme suit. On choisi arbitrairement les valeurs de x_j pour $j \notin \{j_1, \dots j_k\}$ et on pose, pour chaque j_ℓ,

$$x_{j_\ell} = b_\ell - \sum_{j>j_k} a_{k,j}x_j.$$

Vu (2.20), il est évident que les vecteurs obtenus comme cela sont biens les solutions de $AX = B$.

En fin, mentionnons que dans le cas ou (2.20) admet des solutions (c.à-d. quand $b_{k+1} = \dots = b_m = 0$), l'espace vectoriel des solutions a dimensions $n - k$.

Exemple: Considérons le système suivant, au inconnues $x_1, \dots, x_5 \in \mathbb{R}$,

$$\begin{cases} 4x_1 - 4x_2 - 4x_3 - 8x_4 - 8x_5 = -8 \\ 3x_1 - 3x_2 - 3x_3 - x_4 + 4x_5 = 9 \\ x_1 + 2x_3 + x_4 + 5x_5 = 5 \\ 2x_1 + 4x_3 + 6x_5 = 4 \\ -2x_1 + x_2 - x_3 + 4x_4 + 3x_5 = 6 \end{cases}$$

On l'écrit sous forme matricielle réduite comme dans (2.19):

$$\left(\begin{array}{rrrrr|r} 4 & -4 & -4 & -8 & -8 & -8 \\ 3 & -3 & -3 & -1 & 4 & 9 \\ 1 & 0 & 2 & 1 & 5 & 5 \\ 2 & 0 & 4 & 0 & 6 & 4 \\ -2 & 1 & -1 & 4 & 3 & 6 \end{array}\right)$$

On applique l'algorithme pour mettre la matrice sous forme échelonnée: on commence par $k = j = 1$. Vu que l'entrée en position $1, 1$ est non-nulle, on applique le point (ii), à savoir on divise la première ligne par 4 pour obtenir un 1 en haut à gauche:

$$\left(\begin{array}{rrrrr|r} 1 & -1 & -1 & -2 & -2 & -2 \\ 3 & -3 & -3 & -1 & 4 & 9 \\ 1 & 0 & 2 & 1 & 5 & 5 \\ 2 & 0 & 4 & 0 & 6 & 4 \\ -2 & 1 & -1 & 4 & 3 & 6 \end{array}\right) \qquad L_1/4$$

Ensuite on applique l'étape (i) de l'algorithme: on additionne la première ligne multipliée par -3, -1, -2, 2 aux lignes $2, 3, 4$ et 5, respectivement. Ainsi on obtient des zéros sur le reste de la première colonne.

$$\left(\begin{array}{ccccc|c} 1 & -1 & -1 & -2 & -2 & -2 \\ 0 & 0 & 0 & 5 & 10 & 15 \\ 0 & 1 & 3 & 3 & 7 & 7 \\ 0 & 2 & 6 & 4 & 10 & 8 \\ 0 & -1 & -3 & 0 & -1 & 2 \end{array}\right) \qquad \begin{array}{l} \\ L_2 - 3L_1 \\ L_3 - L_1 \\ L_4 - 2L_1 \\ L_5 + 2L_1 \end{array}$$

La première colonne est maintenant échelonnée; on passe à la suivante, donc à $j = k = 2$. Comme l'entrée en position $2, 2$ est nulle, mais qu'il y a des valeurs non-nulles sous cette entrée, on applique le point (iii) de l'algorithme: on échange les lignes 2 et 3 pour mettre une valeur non-nulle à la position $2, 2$.

$$\left(\begin{array}{ccccc|c} 1 & -1 & -1 & -2 & -2 & -2 \\ 0 & 1 & 3 & 3 & 7 & 7 \\ 0 & 0 & 0 & 5 & 10 & 15 \\ 0 & 2 & 6 & 4 & 10 & 8 \\ 0 & -1 & -3 & 0 & -1 & 2 \end{array}\right) \qquad \begin{array}{l} \\ L_3 \\ L_2 \\ \\ \\ \end{array}$$

L'entée à la position $2, 2$ est maintenant déjà égale à 1, il n'est donc pas nécessaire d'appliquer le point (ii), on passe directement au point (i). On additionne la ligne 2 multipliée par 1, -2 et 1 aux lignes 1, 3 et 5, respectivement, pour éliminer les autres entrées de la colonne 2:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 5 & 5 \\ 0 & 1 & 3 & 3 & 7 & 7 \\ 0 & 0 & 0 & 5 & 10 & 15 \\ 0 & 0 & 0 & -2 & -4 & -6 \\ 0 & 0 & 0 & 3 & 6 & 9 \end{array}\right) \qquad \begin{array}{l} L_1 + L_2 \\ \\ \\ L_4 - 2L_2 \\ L_5 + L_2 \end{array}$$

On passe à la colonne 3 (à savoir à $j = k = 3$). Toutes les entrées sous le niveau déjà traité sont nulles (on est dans le cas (iv) de l'algorithme), on peut donc passer à la colonne 4, à savoir à $j = 4$ et $k = 3$. On applique le point (ii): on divise la ligne 3 par 5 pour obtenir l'entrée 1 en position $3, 4$:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 5 & 5 \\ 0 & 1 & 3 & 3 & 7 & 7 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -2 & -4 & -6 \\ 0 & 0 & 0 & 3 & 6 & 9 \end{array}\right) \qquad \begin{array}{l} \\ \\ \frac{1}{5}L_3 \\ \\ \\ \end{array}$$

Pour éliminer les autres entrées sur la colonne 4 on applique le point (i): on additionne la ligne 3 multipliée par $-1, -3, 2-3$ aux lignes $1, 2, 4$ et 5, respectivement.

$$\left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & 3 & 2 \\ 0 & 1 & 3 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \begin{array}{l} L_1 - L_3 \\ L_2 - 3L_3 \\ \\ L_4 + 2L_2 \\ L_5 - 3L_2 \end{array}$$

En fin on passe à la colonne 5 ($j = 5$, $k = 4$). On est à nouveau dans le cas (iv). Comme toutes les colonnes ont été analysées, l'algorithme est fini. La matrice ainsi obtenue est échelonnée de rang 3. Le système initial est équivalent à

$$\begin{cases} x_1 + 2x_3 + 3x_5 & = 2 \\ x_2 + 3x_3 + x_5 & = -2 \\ x_4 + 2x_5 & = 3 \end{cases}$$

Les solutions forment un espace affine de dimension 2. Elles sont obtenues en choisissant x_3 et x_5 arbitrairement, puis en posant

$$\begin{cases} x_1 & = 2 - 2x_3 - 3x_5 \\ x_2 & = -2 - 3x_3 - x_5 \\ x_4 & = 3 - 2x_5 \end{cases} \tag{2.21}$$

Ainsi

$$\mathcal{S} = \left\{ \begin{pmatrix} 2 - 2\lambda - 3\mu \\ -2 - 3\lambda - \mu \\ \lambda \\ 3 - 2\mu \\ \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 2 \\ -2 \\ 0 \\ 3 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -3 \\ -1 \\ 0 \\ -2 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

On peut aussi voir les solution comme la somme d'une solution particulière et d'une solution générale du système homogène. Pour obtenir une solution particulière X_0 on pose $x_3 = x_5 = 0$ et on obtient par (2.21)

$$X_0 = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 3 \\ 0 \end{pmatrix}.$$

Alors $\mathcal{S} = \{X_0 + Y : \ Y \text{ solution du système homogène}\}$. Le système homogène s'écrit

$$\begin{cases} x_1 + 2x_3 + 3x_5 & = 0 \\ x_2 + 3x_3 + x_5 & = 0 \\ x_4 + 2x_5 & = 0 \end{cases}$$

Ses solutions forment un espace $\mathcal{S}_{hom}$ de dimension 2 dans $\mathbb{R}^5$. Une base est formée de deux solutions linéairement indépendantes, par exemple celle obtenue en posant $x_3 = 1$ et $x_5 = 0$ et celle obtenue avec $x_3 = 0$ et $x_5 = 1$:

$$Y_1 = \begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad Y_2 = \begin{pmatrix} -3 \\ -1 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

Ainsi $\mathcal{S}_{hom} = \{\lambda Y_1 + \mu Y_2 : \lambda, \mu \in \mathbb{R}\}$ et $\mathcal{S} = \{X_0 + \lambda Y_1 + \mu Y_2 : \lambda, \mu \in \mathbb{R}\}$.

Exercice 2.24.
Représenter dans l'espace $\mathbb{R}^3$ les solutions de

$$\begin{cases} 2x + 2y + z = 4 \\ x + y + z = 0 \end{cases}$$

Exercice 2.25.
Mettre sous forme échelonnée les systèmes suivantes calculer leur solutions:

$$a) \begin{cases} x + 2y + z = 4 \\ 2x + y + z = 3 \\ x + 5y + 2z = 7. \end{cases} \qquad b) \begin{cases} 3x + 2z = 5y - 6 \\ z = x + 2 \\ 2x - y - z = 0. \end{cases}$$

$$c) \begin{cases} x + 2y + z + t = 4 \\ 2x + 3y + 3z - t = 3 \\ x + 5y + 2z + 3t = 7 \\ 2x + 5y + 4z + 3t = 2. \end{cases} \qquad d) \begin{cases} x + y + z + t = 0 \\ x - y + 2z - 2t = 0 \\ x + y + 4z + 4t = 0 \\ x - y + 8z - 8t = 0. \end{cases}$$

2.4.3 Pivot de Gauss pour le calcul de l'inverse

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Il est souvent intéressant de vérifier si A est inversible et de calculer son inverse A^{-1}. Le pivot de Gauss offre une procédure pratique pour faire cela.

Appliquons le pivot de Gauss à A et à la matrice identité I_n simultanément. Plus précisément, chaque transformation appliquée à A dans l'algorithme du pivot de Gauss, est aussi appliquée à I_n. Notons $\tilde{A}$ la forme échelonnée de A qui en résulte, et $\tilde{I}$ le résultat pour I_n.

Proposition 2.46.
Si $\tilde{A} \neq I_n$, alors A n'est pas inversible.
Si $\tilde{A} = I_n$, alors A est inversible et $A^{-1} = \tilde{I}$.

Preuve: Rappelons que, pendant le pivot de Gauss, on applique des transformations à A qui correspondent à des multiplications à gauche par des matrices inversibles. Notons ces matrices $P_1, \dots, P_k$. Ainsi $\tilde{A} = P_k \dots P_1 A$ et $\tilde{I} = P_k \dots P_1 I_n = P_k \dots P_1$.

Si $\tilde{A}$ n'est pas égale à I_n, alors elle contient au moins une ligne nulle, donc n'est pas inversible. Comme $P = P_k \dots P_1$ est inversible, cela implique que A n'est pas inversible non-plus (car le produit de deux matrices inversibles est forcement inversible – voir la proposition 2.17).

Si $\tilde{A} = PA = I_n$ alors $A = P^{-1}$ est inversible et $A^{-1} = P = P_k \dots P_1 - \tilde{I}$. □

Exercice 2.26.
Déterminer si les matrices suivantes sont inversibles. Si c'est le cas, calculer leur inverse. Multiplier la matrice par le résultat pour vérifier.

$$a) \begin{pmatrix} 1 & 2 & 3 \\ 3 & 7 & 6 \\ 2 & 8 & -5 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}); \qquad b) \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_5(\mathbb{R});$$

$$c) \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}_5(\mathbb{R}); \qquad d) \begin{pmatrix} 1 & -1 & \dots & -1 & -1 \\ 0 & 1 & \dots & -1 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R});$$

$$e) \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 2 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n-1 & n-2 & \dots & 1 & 0 \\ n & n-1 & \dots & 2 & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}); \qquad f) \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

2.4.4 Familles de vecteurs et pivot de Gauss

Soit $C_1, \dots, C_n \in \mathcal{M}_{m,1}$ une famille de vecteurs de $\mathbb{R}^m$ écrits en format colonne. A l'aide du pivot de Gauss on peut déterminer si la famille $C_1, \dots, C_n$ est libre; plus généralement on peut calculer son rang.

En effet, posons $A = \begin{pmatrix} C_1, \dots, C_n \end{pmatrix} \in \mathcal{M}_{m,n}$ la matrice formée des colonnes $C_1, \dots, C_n$. Alors $\mathrm{rang}(C_1, \dots, C_n) = \mathrm{rang}(A) = \mathrm{rang}(\tilde{A})$, ou $\tilde{A}$ est la matrice échélonnée obtenue à partir de A par le pivot de Gauss. En particulier, la famille $C_1, \dots, C_n$ est libre si et seulement si $\mathrm{rang}(\tilde{A}) = n$.

Mentionnons qu'on peut aussi déterminer les scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que

$$\lambda_1 C_1 + \dots + \lambda_n C_n = 0. \tag{2.22}$$

Si on pose $\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$, alors (2.22) s'ecrit

$$A\Lambda = \lambda_1 C_1 + \cdots + \lambda_n C_n = 0,$$

ce qui revient à dire que $\Lambda \in \mathrm{Ker}(A) = \mathrm{Ker}(\tilde{A})$.

En conclusion, une famille $\lambda_1, \ldots, \lambda_n$ satisfait (2.22) si et seulement si le vecteur colonne qu'elle forme est dans $\mathrm{Ker}(A) = \mathrm{Ker}(\tilde{A})$. On peut donc, de façon alternative, voir si la famille $C_1, \ldots, C_n$ est libre en calculant $\mathrm{Ker}(\tilde{A})$. La famille est libre si et seulement si l'unique famille de scalaires satisfaisant (2.22) est nulle, donc si et seulement si $\mathrm{Ker}(A) = \mathrm{Ker}(\tilde{A}) = \{0\}$.

À retenir

Applications linéaires:

- Pour deux espaces vectoriels E, F, une application $u : E \to F$ est linéaire si $u(\lambda x + \mu y) = \lambda u(x) + \nu u(y)$ pour tout $x, y \in E$ et $\lambda, \nu \in \mathbb{R}$. L'ensemble des applications linéaires est noté $\mathcal{L}(E, F)$.

- Les applications peuvent être additionnées entre elles et multiplies par des scalaires. Avec ces opérations, $\mathcal{L}(E, F)$ est un espace vectoriel.

- Pour E, F, G des e.v. et $u \in \mathcal{L}(E, F)$, $v \in \mathcal{L}(F, G)$, la composition de v et u, notée $v \circ u$, est donnée par $x \mapsto v(u(x))$. C'est une application linéaire de $\mathcal{L}(E, G)$.

- Dans $\mathcal{L}(E) = \mathcal{L}(E, E)$ on a trois opérations: $+$, $\cdot$ et $\circ$. $\mathcal{L}(E)$ est un algèbre.

- Une application $u \in \mathcal{L}(E, F)$ est inversible, si et seulement si elle est bijective. Si elle est inversible, alors $u^{-1} \in \mathcal{L}(F, E)$ est l'unique application telle que $u \circ u^{-1} = u^{-1} \circ u = \mathrm{id}$.

- Pour $u \in \mathcal{L}(E, F)$ on pose
$$\mathrm{Ker}(u) = \{x \in E : u(x) = 0\} \quad \text{et} \quad \mathrm{Im}(u) = \{y \in F : \exists x \in E \text{ t.q. } u(x) = y\}.$$
Ce sont des sous-espaces vectoriels de E et F, respectivement.

- Si E est de dimensions finie, alors $\mathrm{Im}(u)$ l'est aussi. On note $\mathrm{rang}(u) = \dim(\mathrm{Im}(u))$ et on a
$$\mathrm{rang}(u) + \dim(\mathrm{Ker}(u)) = \dim(E).$$

Matrices:

- Une matrice de taille $m \times n$ est un tableau rectangulaire de $m \cdot n$ scalaires noté $A = (a_{i,j})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$. On écrit $\mathcal{M}_{m,n}(\mathbb{R})$ pour l'ensemble des matrices de taille $m \times n$.

- On peut additionner deux matrices de même taille et multiplier une matrice par un scalaire. $\mathcal{M}_{m,n}(\mathbb{R})$ est un espace vectoriel.

- On peut multiplier une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ avec $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et on obtient $AB \in \mathcal{M}_{m,p}(\mathbb{R})$. Avec cette règle de multiplication $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$ est une algèbre.
La multiplication des matrices n'est pas commutative!

- L'élément neutre pour la multiplication dans $\mathcal{M}_n(\mathbb{R})$ est la matrice $I_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$.
Une matrice A est inversible s'il existe $A^{-1} \in \mathcal{M}_n(\mathbb{R})$ telle que $AA^{-1} = A^{-1}A = I_n$.

Applications linéaires représentées par des matrices:
Soient E et F deux espaces vectoriels de dimensions n et m respectivement. Soient $\mathcal{E}$ une base de E et $\mathcal{F}$ une base de F.

- Les matrices colonne de $\mathcal{M}_{n,1}(\mathbb{R})$ représentent les vecteurs de E dans la base $\mathcal{E}$. Plus précisément, $V_{\mathcal{E}} : E \to \mathcal{M}_{n,1}(\mathbb{R})$ est une application linéaire bijective. De la même façon $\mathcal{M}_{m,1}(\mathbb{R})$ représente les vecteurs de F dans la base $\mathcal{F}$.

- Les matrices de $\mathcal{M}_{m,n}(\mathbb{R})$ représente les applications linéaires de $\mathcal{L}(E,F)$ dans les bases $\mathcal{E}$ et $\mathcal{F}$. Plus précisément, $M_{\mathcal{F},\mathcal{E}} : \mathcal{L}(E,F) \to \mathcal{M}_{m,n}(\mathbb{R})$ est une application linéaire bijective. Pour tout $u \in \mathcal{L}(E,F)$ et $x \in E$:
$$V_{\mathcal{F}}\big(u(x)\big) = \mathrm{Mat}_{\mathcal{F},\mathcal{E}}(u)V_{\mathcal{E}}(x).$$

- La multiplication des matrices correspond à la composition des applications. Une application linéaire est inversible si et seulement si la matrice associée l'est.

- La représentation des vecteurs et des applications linéaires par des matrices dépend des bases $\mathcal{E}$ et $\mathcal{F}$. Pour passer d'une base à une autre on utilise la matrice de changement de base. Pour deux bases $\mathcal{E}$ et $\mathcal{B}$ de E et $x \in E$ et $u \in \mathcal{L}(E)$,
$$V_{\mathcal{B}}(x) = \mathrm{P}_{\mathcal{B},\mathcal{E}}V_{\mathcal{E}}(x) \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{P}_{\mathcal{E},\mathcal{B}}^{-1}\mathrm{Mat}_{\mathcal{E}}(u)\mathrm{P}_{\mathcal{E},\mathcal{B}}.$$

Image, noyaux et rang des matrices:

- Comme pour les applications, on pose pour une matrice $M \in \mathcal{M}_{m,n}$,
$$\begin{aligned}&\mathrm{Ker}(M) = \{X \in \mathcal{M}_{n,1} : MX = 0\} \quad \text{et}\\ &\mathrm{Im}(M) = \{Y \in \mathcal{M}_{m,1} : \exists X \in \mathcal{M}_{n,1} \text{ avec } MX = Y\}.\end{aligned}$$
Ce sont des sous-espaces vectoriels de $\mathbb{R}^n$ et $\mathbb{R}^m$, respectivement.

- On défini le rang d'une matrice $M \in \mathcal{M}_{m,n}(\mathbb{R})$ par $\mathrm{rang}(M) = \dim(\mathrm{Im}(M))$. C'est également le rang de la famille des lignes de M (comme vecteurs de $\mathbb{R}^n$) et celui de la famille des colonnes (comme vecteurs de $\mathbb{R}^m$).

- Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement si $\mathrm{rang}(M) = n$.

- Un système linéaire est une équation matricielle $AX = B$ ou $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{m,1}(\mathbb{R})$ sont les coefficients et $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est l'inconnue. L'ensemble des solutions $\mathcal{S} = \{X \in \mathcal{M}_{n,1}(\mathbb{R}) : AX = B\}$ peut être vide (si $B \notin \mathrm{Im}(A)$) ou peut être une espace affine de dimension $\dim(\mathrm{Ker}(A)) = n - \mathrm{rang}(A)$. Dans le second cas
$$\mathcal{S} = X_0 + \mathrm{Ker}(A) = \{X_0 + Y : Y \in \mathcal{M}_{n,1}(\mathbb{R}) \text{ est telle que } AY = 0\},$$
ou X_0 est une solution particulière de $AX = B$.

- Les solutions d'un système linéaire peuvent être trouvées par l'algorithme du pivot de Gauss. L'inverse d'une matrice peut aussi être calculé par ce même algorithme.

Chapter 3

Matrices diagonalisables; valeurs et vecteurs propres

Le chapitre va porter entièrement sur des matrices carrées. Fixons pour l'intégralité de ce chapitre une matrice $A \in \mathcal{M}_n(\mathbb{R})$. Comme avant, on identifie $\mathcal{M}_{n,1}(\mathbb{R})$ avec $\mathbb{R}^n$; on va appeler parfois les éléments de $\mathcal{M}_{n,1}(\mathbb{R})$ des vecteurs.

3.1 Valeurs et vecteurs propres

Définition 3.1. *On dit que* $\lambda \in \mathbb{R}$ *est une* valeur propre[(i)] *de* A *s'il existe un vecteur* $X \in \mathcal{M}_{n,1}(\mathbb{R})$, $X \neq 0$, *tel que*

$$AX = \lambda X.$$

Dans ce cas on dit que X *est un* vecteur propre[(ii)] *de* A *pour la valeur propre* λ. *L'ensemble des valeurs propres de* A *s'appelle le* spectre[(iii)] *de* A *et est noté* $Sp(A)$. *Pour* $\lambda \in Sp(A)$, *on note*

$$E_\lambda(A) = \big\{X \in \mathcal{M}_{n,1}(\mathbb{R}) : AX = \lambda X\big\}.$$

Proposition 3.2.

(i) *Pour chaque* $\lambda \in Sp(A)$, $E_\lambda(A)$ *est un sous-espace vectoriel de* $\mathbb{R}^n$ *qu'on appelle* l'espace propre[(iv)] *de* A *associé à* λ.

(ii) *Si* $\lambda_1, \dots, \lambda_k \in Sp(A)$ *sont deux à deux distinctes et* $X_1 \in E_{\lambda_1}, \dots, X_k \in E_{\lambda_k}$ *sont des vecteurs non-nuls, alors la famille* $(X_1, \dots, X_k)$ *est libre.*

(iii) *Le spectre de* A *a au plus* n *éléments.*

(i) Eigenwert

(ii) Eigenvector

(iii) Spektrum

Preuve: (*i*) Soit $\lambda \in Sp(A)$. Alors $X \in E_\lambda(A)$ si et seulement si $AX - \lambda X = (A - \lambda I_n)X = 0$. Ainsi

$$E_\lambda(A) = \mathrm{Ker}(A - \lambda I_n).$$

En particulier on déduit que $E_\lambda(A)$ est un s.e.v. de $\mathbb{R}^n$.

(*ii*) Ce point est admis.

(*iii*) Soient $\lambda_1, \ldots, \lambda_k \in Sp(A)$ deux à deux distincts. On peut alors choisir des vecteurs non-nuls $X_1 \in E_{\lambda_1}, \ldots, X_k \in E_{\lambda_k}$. Par le point précèdent $(X_1, \ldots, X_k)$ est une famille libre de $\mathbb{R}^n$, donc $k \leq n$. Ainsi $|Sp(A)| \leq n$. □

Remarque 3.3. On insiste sur l'importance du point (*i*) et de sa preuve. Vu que

$$E_\lambda(A) = \mathrm{Ker}(A - \lambda I),$$

les valeurs propres de A sont exactement les $\lambda \in \mathbb{R}$ pour lesquels $A - \lambda I$ n'est pas inversible.

Les mêmes notions de valeur propre, vecteur propre, et espace propre se généralisent aux applications linéaires. Dans ce cours on va se limiter aux matrices. Il peut toutefois être intéressant de remarquer le liens entre les deux.

Proposition 3.4. *Soient E un e.v. de dimension finie, $\mathcal{B}$ une base de E, $u \in \mathcal{L}(E)$ et $x \in E$, $x \neq 0$. Alors on a équivalence de:*

(i) *x est vecteur propre de u avec valeur propre λ (c.-à-d. $u(x) = \lambda x$);*

(ii) $\mathrm{Mat}_\mathcal{B}(u)V_\mathcal{B}(x) = \lambda V_\mathcal{B}(x)$.

La preuve suit directement des propriétés de $\mathrm{Mat}_\mathcal{B}$ et $V_\mathcal{B}$ et on ne la donne pas ici.

On observe que, quand on parle de vecteur propre pour les applications linéaires on ne fixe pas une base de l'espace vectoriel. Ainsi, il est naturel que les valeur propres d'une matrice soient invariantes par changement de base.

Proposition 3.5. *Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible. Alors*

(i) *$Sp(A) = Sp(P^{-1}AP)$;*

(ii) *pour tout $\lambda \in \mathbb{R}$, $X \in E_\lambda(P^{-1}AP)$ si et seulement si $PX \in E_\lambda(A)$.*

Preuve: On commence par le point (*ii*). Soit $\lambda \in \mathbb{R}$ et $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Supposons que $X \in E_\lambda(P^{-1}AP)$. Alors $P^{-1}APX = \lambda X$. Si on multiplie cette égalité par P à gauche, on obtient $APX = \lambda PX$, donc $PX \in E_\lambda(A)$.

Inversement si $PX \in E_\lambda(A)$, alors $P^{-1}APX = \lambda P^{-1}PX = \lambda X$, donc $X \in E_\lambda(A)$.

Passons au point (*i*). Soit $\lambda \in Sp(P^{-1}AP)$ et $X \in E_\lambda(P^{-1}AP)$ non nul. Alors, par le point précédent, $PX \in E_\lambda(A)$. De plus, comme P est inversible, $PX \neq 0$, donc $E_\lambda(A) \neq \{0\}$. On en déduit que $\lambda \in Sp(A)$, donc que $Sp(P^{-1}AP) \subset Sp(A)$.

Inversement, si $\lambda \in Sp(A)$ et $X \in E_\lambda(A)$ non nul, alors $P^{-1}X \in Sp(P^{-1}AP)$. A nouveau $P^{-1}X \neq 0$ car P^{-1} est inversible, donc $\lambda \in Sp(P^{-1}AP)$. Il s'en suit que $Sp(A) \subset Sp(P^{-1}AP)$. La double inclusion montre que $Sp(A) = Sp(P^{-1}AP)$. □

(v) Eigenraum

3.2 Matrices diagonalisables

Définition 3.6. *On dit que A est une matrice* diagonalisable[vi] *(dans $\mathcal{M}_n(\mathbb{R})$) s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$ telle que $P^{-1}AP$ est une matrice diagonale.*

L'intérêt de cette notion peut s'expliquer par l'observation suivante. Dans beaucoup de situations pratiques il est intéressant de calculer les puissances d'une matrice carrée. Comment peut-on donc calculer A^k pour une grande matrice A, ou k est une très grande valeur, sans faire trop de calculs?

Si A est diagonalisable, le calcul est simple. Soit $P \in \mathcal{M}_n(\mathbb{R})$ inversible, telle que la matrice $D = P^{-1}AP$ soit diagonale. Alors $A = PDP^{-1}$, et

$$A^k = (PDP^{-1})^k = \underbrace{PD\cancel{P^{-1}}\cancel{P}D\cancel{P^{-1}}\ldots\cancel{P}DP^{-1}}_{k \text{ fois}} = PD^kP^{-1}.$$

Mais D est une matrice diagonale de la forme $D = \begin{pmatrix} d_{1,1} & 0 & \ldots & 0 \\ 0 & d_{2,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_{n,n} \end{pmatrix}$. On peut donc facilement calculer

$$D^k = \begin{pmatrix} d_{1,1}^k & 0 & \ldots & 0 \\ 0 & d_{2,2}^k & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_{n,n}^k \end{pmatrix}.$$

Ainsi, le calcul des puissances des matrices diagonalisables est facile, surtout si on connait la matrice P.

3.2.1 Lien avec les vecteurs propres

Théorème 3.7. *La matrice A est diagonalisable si et seulement s'il existe une base $\mathcal{E} = (X_1, \ldots, X_n)$ de $\mathbb{R}^n$ formée entièrement de vecteurs propres de A.*

Preuve: Avant de commencer, définissons, pour $i \in 1, \ldots, n$, le vecteur $Y_i \in \mathcal{M}_{n,1}$ comme ayant l'entrée 1 à la position i et 0 partout ailleurs. ($Y_1, \ldots, Y_n$ forment la base canonique de $\mathbb{R}^n$.) Alors, pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$, MY_i est la $i$eme colonne de M (cela résulte directement des règles de la multiplication matricielle).

[vii]diagonalisierbar

Supposons pour commencer que A est diagonalisable, et soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible, telle que

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

pour des valeurs $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (pas forcement distinctes). Posons, pour $i \in \{1, \dots, n\}$, $X_i = PY_i \in \mathcal{M}_{n,1}$; ainsi X_i est la $i$eme colonne de P. La forme diagonale de $P^{-1}AP$ entraine que chaque Y_i en est un vecteur propre, avec λ_i la valeur propre correspondante. Ainsi

$$\lambda_i Y_i = P^{-1}APY_i = P^{-1}AX_i, \qquad \text{pour chaque } 1 \leq i \leq n.$$

En multipliant cette équation par P à gauche, on obtient $\lambda_i X_i = AX_i$. On conclut $X_1, \dots, X_n$ est une famille de vecteurs propres de A.

De plus, comme P est inversible,

$$n = \text{rang}(P) = \text{rang}(X_1, \dots, X_n).$$

Ainsi $(X_1, \dots, X_n)$ est une base de $\mathbb{R}^n$.

Supposons maintenant qu'il existe une base $(X_1, \dots, X_n)$ de $\mathbb{R}^n$ de vecteurs propres de A. Soient $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ les valeurs propres associées.

On défini la matrice $P \in \mathcal{M}_n(\mathbb{R})$ comme la matrice dont les colonnes sont données par $X_1, \dots, X_n$. Alors, pour chaque $1 \leq i \leq n$, $PY_i = X_i$. Comme $(X_1, \dots, X_n)$ est une base de $\mathbb{R}^n$,

$$n = \text{rang}(X_1, \dots, X_n) = \text{rang}(P),$$

donc P est inversible. On peut alors calculer, pour chaque $1 \leq i \leq n$,

$$P^{-1}APY_i = P^{-1}AX_i = \lambda_i P^{-1}X_i = \lambda_i P^{-1}PY_i = \lambda_i Y_i.$$

On en déduit que les colonnes de $P^{-1}AP$ sont $\lambda_1 Y_1, \dots, \lambda_n Y_n$, respectivement. On peut donc écrire

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

et on conclut que A est diagonalisable. □

Remarque 3.8. Dans les faites, la matrice P n'est rien d'autre que la matrice de changement de base de la base canonique de $\mathbb{R}^n$ à la base de vecteurs propres de A. On aimerait insister sur le fait que la base de vecteurs propres (et donc la matrice P qui diagonalise A) ne sont pas uniques!

Si $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable et P est une matrice inversible telle que $P^{-1}AP = D$ est diagonale, alors les valeurs propres de A sont exactement les éléments de la diagonale de D. De plus, pour $\lambda \in Sp(A)$, $\dim(E_\lambda(A))$ est simplement le nombre d'apparitions de λ sur la diagonale de D.

Le critère de diagonalisation suivant va nous servir par la suite; il découle du théorème précédent.

Proposition 3.9. *La matrice A est diagonalisable si et seulement si*

$$\sum_{\lambda \in Sp(A)} \dim(E_\lambda(A)) = n.$$

Preuve: On va admettre ce résultat. $\square$

Malheureusement pas toutes les matrices sont diagonalisables.

Lemme 3.10. *La matrice $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'est pas diagonalisable.*

Preuve: On va procéder par l'absurde. Supposons que N est diagonalisable. Soit $P \in \mathcal{M}_2(\mathbb{R})$ inversible, telle que $P^{-1}NP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ soit une matrice diagonale. Alors

$$\begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} = (P^{-1}NP)^2 = P^{-1}N^2P = P^{-1}0P = 0,$$

donc $\lambda_1 = \lambda_2 = 0$. Cela implique que $N = P^{-1}0P = 0$, ce qui est clairement faux. $\square$

Comme les matrices diagonalisables sont souvent plus facile à traiter, il est intéressant d'avoir des critères pour les reconnaitre. On commence par un critère suffisant, mais pas nécessaire; il suit facilement des résultats déjà mentionnés.

Corollaire 3.11. *Si A est telle que $Sp(A)$ contient n valeurs propres distinctes, alors A est diagonalisable.*

Preuve: Si $Sp(A)$ contient n valeurs propres distinctes $\lambda_1, \dots, \lambda_n$, alors

$$\dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_n}) \geq n.$$

Mais on a vu que cette somme est toujours plus petite que n, donc

$$\dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_n}) = n.$$

La proposition 3.9 implique que A est diagonalisable. $\square$

Matrices symétriques

Théorème 3.12. *Si $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique (c.-à-d. $A^T = A$) alors A est diagonalisable. De plus ses valeurs propres sont toutes réelles et il existe une matrice $P \in \mathcal{M}_n(\mathbb{R})$ inversible, telle que $P^{-1} = P^T$, avec*

$$A = P^T DP,$$

ou D est une matrice diagonale.

Une matrice $P \in \mathcal{M}_n(\mathbb{R})$ telle que $P^{-1} = P^T$ est dite orthogonale. Ce type de matrices a une signification particulière dans l'interprétation géométrique des espaces vectoriels. On va admettre ce théorème.

Matrices positives: Perron-Frobenius

Théorème 3.13. *Soit $A = (a_{i,j})_{1 \le i,j \le n} \in \mathcal{M}_n(\mathbb{R})$ une matrice avec $a_{i,j} > 0$ pour tout $1 \le i, j \le n$. Alors il existe $\lambda_0 \in (0, +\infty)$ une valeur propre de A telle que:*

(i) *pour toute autre valeur propre $\lambda \in Sp(A) \setminus \{\lambda_0\}$, on a $|\lambda| < \lambda_0$;*

(ii) $\dim E_{\lambda_0}(A) = 1$ *(on dit que la valeur propre λ_0 est simple);*

(iii) *il existe $X \in E_{\lambda_0}(A)$ avec toutes les entées de X réelles et strictement positives;*

(iv) *si $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est un vecteur propre de A avec toutes les entrées positives ou nulles, alors $X \in E_{\lambda_0}(A)$.*

On va admettre ce théorème (il s'agit d'un théorème difficile).

Mentionnons que ce résultat se généralise à certaines matrices aux entrées positives ou nulles (pas à toutes). On peut par exemple montrer que si $A \in \mathcal{M}_n(\mathbb{R})$ est telle que $a_{i,j} \ge 0$ pour tout i, j et si toutes les entrées de A^n sont strictement positives pour un certain $n \in \mathbb{N}$, alors la conclusion du théorème est encore valable.

En pratique on rencontre souvent des matrices dont les entrées sont positives. Par exemple, les matrices qui décrivent les évolutions de populations sont souvent de ce type (voir exemple dans la partie 3.2.2). Le théorème de Perron Frobenius nous dit qu'il existe $\lambda_0 \ge 0$ et $X \in \mathcal{M}_{n,1}((0, +\infty))$ avec $AX = \lambda_0 X$.

Si on suppose que A est diagonalisable, avec valeurs propres $\lambda_0, \dots, \lambda_{n-1}$ et vecteurs propres $(X_1, \dots, X_n)$. Alors, pour tout vecteur $X = \sum_i \alpha_i X_i$ avec $\alpha_0 \neq 0$, on a

$$\frac{1}{\lambda_0^n} A^n X = \sum_i \alpha_i \left(\frac{\lambda_i}{\lambda_0} \right)^n X_i \to_n \alpha_0 X_0.$$

Ainsi le "taux" de croissance de $A^n X$ est positif et on observe une distribution (après normalisation) proportionnelle à un vecteur positif.

Le même résultat asymptotique peut être montré même si A n'est pas diagonalisable. On en parlera plus dans l'exemple donné dans la partie 3.2.2.

Perron-Frobenius pour les matrices stochastiques Supposons que $A \in \mathcal{M}_n$ est une matrice stochastique à gauche. Alors tous les coefficients de A sont positifs ou nuls. Supposons de plus qu'il existe $N \ge 1$ tel que tous les coefficients de A^N soit strictement positifs (on dit alors que A est irréductible et apériodique). On peut alors appliquer le théorème de Perron-Frobenius à A (une version un peu plus générale du théorème quand $N > 1$)

et déduire l'existence d'un vecteur propre $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ d'entrées strictement positives et de la valeur propre λ associée, qui est strictement positive et satisfait les points (i) et (ii) du théorème.

Du fait que $AX = \lambda X$ on peut déduire

$$\lambda \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} (AX)_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_j = \sum_{j=1}^{n} \Big(x_j \sum_{i=1}^{n} a_{i,j} \Big) = \sum_{j=1}^{n} x_j.$$

Dans la dernière égalité on a utilisé que $\sum_{i=1^n} a_{i,j} = 1$ car A est stochastique à gauche. En fin, les x_i sont tous strictement positifs, donc $\sum_{i=1}^n x_i = \sum_{j=1}^n x_j > 0$. On peut simplifier par cette quantité pour trouver $\lambda = 1$.

On conclut que toutes les autres valeurs propres de A sont de module strictement plus petit que 1 et que tout vecteur propre de valeur propre 1 est proportionnel à X. Si on choisie X de sorte que $\sum_{i=1}^n x_i = 1$ (ce qui le détermine uniquement) on peut voir X comme une probabilité sur $\{1, \dots, n\}$ (avec une probabilité x_i associée à i).

De la discussion qui précède, on conclut que cette probabilité est

- *invariante*: car $AX = X$;
- *asymptotique*: car pour tout vecteur Y d'entrées positives se sommant à 1, on a

$$A^N Y \xrightarrow[N \to \infty]{} X.$$

De plus, chaque propriété détermine X de manière unique.

Exercice 3.1.
Soient $A \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonalisable, avec n valeurs propres distinctes. Soit $B \in \mathcal{M}_n(\mathbb{R})$ une matrice qui commute avec A (c.-à-d. telle que $AB = BA$).

Montrer que B est aussi diagonalisable. De plus, montrer que B admet les mêmes vecteurs propres que A (mais pas forcement les mêmes valeurs propres).
Est-ce que le résultat reste vrai si les valeurs propres de A ne sont pas supposées distinctes?

Exercice 3.2.
Trouver une matrice diagonalisable A, et plusieurs matrices inversibles P telles que $P^{-1}AP$ est diagonale.
Que dire de A si $P^{-1}AP$ est diagonale pour toute matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$.

3.2.2 Application: matrices de Leslie

Les matrices de Leslie sont un modèle d'évolution d'une population classée selon l'âge. On considère une population qui évolue en temps discret (on peut par exemple supposer qu'on la mesure tout les ans/mois) dont les individus peuvent avoir un âge compris entre 1 et $d \in \mathbb{N}$.

Ainsi, la population à tout moment $t \in \mathbb{N}$ est représenté par un vecteur $N(t) = \begin{pmatrix} n_1(t) \\ \vdots \\ n_d(t) \end{pmatrix}$.

Supposons que l'évolution suit la dynamique suivante:

$$\begin{aligned} n_{k+1}(t+1) &= s_k n_k(t), \qquad\qquad \text{pour } 1 \le k < d, \text{ et} \\ n_1(t+1) &= a_1 n_1(t) + \dots + a_d n_d(t). \end{aligned}$$

ou $s_1 \dots s_{d-1} \in (0,1]$ et $a_1, \dots a_d \in (0,+\infty)$ sont des paramètres fixés. Ainsi $1-s_k$ représente la proportion d'individus d'âge k qui meurent avant d'arriver à l'âge $k+1$; a_k représente le nombre moyen d'enfants d'un individu de la génération k.

On peut écrire l'évolution sous forme matricielle comme suit:

$$N(t+1) = \begin{pmatrix} a_1 & a_2 & \dots & a_{d-1} & a_d \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{d-1} & 0 \end{pmatrix} N(t) = AN(t),$$

ou $A \in \mathcal{M}_d(\mathbb{R})$ est la matrice affichée. Même si la matrice A n'est pas strictement positive, A^d l'est, et on peut appliquer le théorème de Perron Frobenius à A.

Cherchons λ (different de 0) et $N \in \mathcal{M}_d(\mathbb{R}_+)$ tels que $AN = \lambda N$. Cela reviens à

$$\begin{aligned} n_{k+1} &= \frac{1}{\lambda} s_k n_k = \dots = \frac{1}{\lambda^k} s_1 \dots s_k n_1, \qquad\qquad \text{pour } 1 \le k < d, \text{ et} \\ n_1 &= \frac{1}{\lambda}(a_1 n_1 + \dots + a_d n_d) = \Big(\sum_{k=1}^{d} \frac{1}{\lambda^k} s_1 \dots s_{k-1} a_k \Big) n_1 \end{aligned}$$

Ainsi, λ est valeur propre de A (avec $n_1 \neq 0$) si et seulement si

$$\phi(\lambda) := \sum_{k=1}^{d} \frac{1}{\lambda^k} s_1 \dots s_{k-1} a_k = 1;$$

de plus pour cette valeur propre, un vecteur propre associé est donné par $n_k = \frac{s_1 \dots s_{k-1}}{\lambda^{k-1}}$ pour $k \ge 1$ (en particulier $n_1 = 1$).

On voit bien que $\phi(\lambda)$ admet exactement une racine positive, qu'on va noter λ_1.

Supposons que A est diagonalisable, avec $Sp(A) = \{\lambda_1, \dots, \lambda_d\}$. Soient $X_1, \dots, X_d$ des vecteurs propres associés; ils forment une base de $\mathbb{R}^d$ qu'on note $\mathcal{X}$. Alors, si on suppose que λ_1 est la valeur propre positive,

$$|\lambda_i| < \lambda_1 \quad \forall i \ge 2.$$

Soit $N(0)$ une distribution de population initiale. On écrit alors $N(0) = \sum_i \alpha_i X_i$ ou

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = V_{\mathcal{X}}(N(0)) \in \mathcal{M}_{d,1}(\mathbb{C}).$$

Un calcule immédiat montre que

$$N(t) = A^t N(0) = \sum_{i=1}^{d} \alpha_i \lambda_i^t X_i.$$

Ainsi

$$\lambda_1^{-t} N(t) = \alpha_1 X_1 + \sum_{i=2}^{d} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^t X_i \xrightarrow[t\to\infty]{} \alpha_1 X_1. \tag{3.1}$$

Cela doit être compris comme suit. Si $\alpha_1 > 0$, alors la population à un taux de croissance λ_1 et si on la normalise, elle devient proportionnelle à une distribution suivant l'âge donnée par X_1.

Une analyse plus approfondie des espaces propres des matrices montre que la convergence de (3.1) est vraie même si A n'est pas diagonalisable.

3.3 Le déterminant

On a vu qu'une propriété importante des matrices est l'inversibilité. Le déterminant nous offre un critère pratique pour vérifier si une matrice est inversible.

3.3.1 Définition et propriétés de base

Le déterminant associe à une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est un scalaire qu'on va noter $\det(A) \in \mathbb{R}$. Il y a plusieurs façons de définir le déterminant, on choisie celle par récurrence. On aura besoin de la notation suivante:

Définition 3.14. *Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $1 \le i, j \le n$. Le* mineur *$A_{i,j}$ de A est la matrice de $\mathcal{M}_{n-1}(\mathbb{R})$ obtenue en éliminant la i^{eme} ligne et j^{eme} colonne de A.*

Définition 3.15. *Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$.*

- *Si $n = 1$, $\det A = a_{1,1}$.*
- *Si $n \ge 2$,*

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

Donnons quelques conséquences de cette définition.

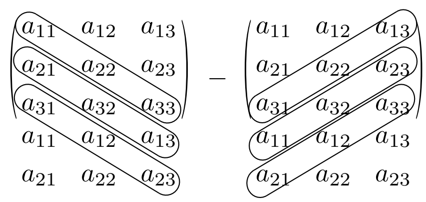
- Pour les matrices de $\mathcal{M}_2(\mathbb{R})$ le déterminant est donné par :

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc;$$

- Pour les matrices de $\mathcal{M}_2(\mathbb{R})$ le déterminant est donné par :

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = \tag{3.2}$$
$$= a_{1,1}(a_{2,2}a_{3,3} - a_{3,2}a_{2,3}) + a_{1,2}(a_{2,3}a_{3,1} - a_{2,1}a_{3,3}) + a_{1,3}(a_{2,1}a_{3,2} - a_{3,1}a_{2,2})$$
$$= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{3,1}a_{2,2}a_{1,3} - a_{2,1}a_{1,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3}.$$

  Un moyen de se souvenir de cette formule est par l'image suivante.



- Plus généralement, le déterminant d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est une somme de produits, chaque produit contenant n entrées, une par ligne et une par colonne. Ainsi il y a $n!$ produits dans l'expression du déterminant (certains ayant un signe $+$, certains un signe $-$). Formellement

$$\det(A) = \sum_{\sigma} \epsilon(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)}, \tag{3.3}$$

  ou la somme porte sur toutes les bijections $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ et ou ϵ vaut $+1$ ou -1 en fonction de σ.

Généralement le calcul du déterminant d'une matrice est compliqué. Pour les matrices 3×3 le nombre de termes dans la somme de (3.2) est 6. On peut écrire une formule similaire pour les matrices 4×4; elle va contenir 24 termes ...

Néanmoins, dans certains cas particulier, le déterminant se calcule facilement.

Proposition 3.16. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice triangulaire supérieure ou inférieure. Alors,*

$$\det A = a_{1,1}a_{2,2} \dots a_{n,n}.$$

Cette proposition se montre par récurrence sur la taille de A; on la laisse en exercice.

3.3.2 Déterminant et inversibilité

On va donner ici quelques propriétés essentielles du déterminant.

Proposition 3.17. *Soient $A, B \in \mathcal{M}_n(\mathbb{R})$. Alors,*

$$\det(AB) = \det A \cdot \det B.$$

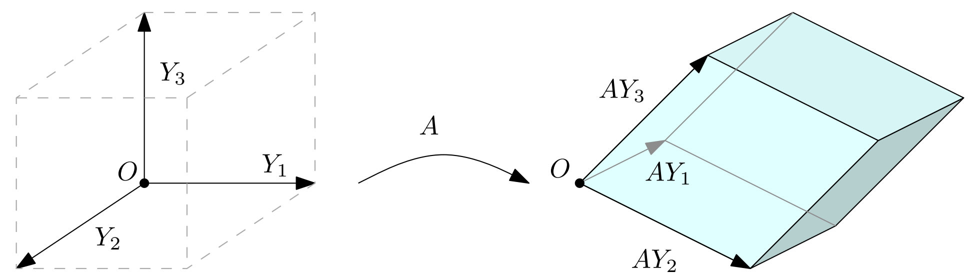


Figure 3.1: Une matrice A transforme les trois vecteurs de la base canonique de $\mathbb{R}^3$. L'image du cube unité par A est marquée en bleu; son volume est (à signe près) le déterminant de A.

On va admettre cette proposition.

Théorème 3.18. *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors, A est inversible si et seulement si* $\det A \neq 0$.

Une preuve rapide est basée sur le pivot de Gauss; on la donne dans la partie suivante.

Une façon géométrique de voir le déterminant d'une matrice A est comme le volume du polyèdre déterminé par les vecteurs colonne de A. Rappelons nous de la base canonique de $\mathbb{R}^n$ formé des vecteurs $Y_1, \dots, Y_n$. Le polyèdre déterminé par les vecteurs colonne de A est alors l'image du cube $[0,1]^n = \{\sum_{i=1}^n \lambda_i Y_i : \lambda_1, \dots, \lambda_n \in [0,1]\}$ par A. Ainsi

$$|\det(A)| = \mathrm{Vol}\Big\{\sum_{i=1}^n \lambda_i AY_i : \lambda_1, \dots, \lambda_n \in [0,1]\Big\}.$$

On a déjà vu que A n'est pas inversible si et seulement si les vecteurs colonne de A, à savoir $AY_1, \dots, AY_n$, sont liées. Cela reviens à $\mathrm{rang}(AY_1, \dots, AY_n) < n$, donc à ce que le polyèdre mentionné fasse partie d'un hyperplan de $\mathbb{R}^n$. On voit bien que alors son volume est nul.

Exemple: Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Alors A est inversible si et seulement si $ad - bc \neq 0$. De plus, dans ce cas

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

3.3.3 Opérations sur lignes; pivot de Gauss

Rappelons nous des opérations sur les lignes utilisées pour le pivot de Gauss. Ces opérations changent le déterminant comme suit.

Proposition 3.19. *Pour* $i \neq j$ *et* $\lambda \neq 0$,

- $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$ *ne change pas le déterminant,*
- $\mathrm{Mult}_{\lambda,\mathrm{i}}$ *multiplie le déterminant par* λ,
- $\mathrm{Ech}_{\mathrm{i},\mathrm{j}}$ *change le signe du déterminant (c.à d. le multiplie par* -1*).*

Rappelons nous que ces opérations correspondent à des multiplications à gauche par des matrices spécifiques, qu'on a notées aussi $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i},\mathrm{j}}$. Un calcul directe mène au lemme suivant.

Lemme 3.20. *Pour* $i \neq j$ *et* $\lambda \neq 0$,

$$\det(\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}) = 1 \qquad et \qquad \det(\mathrm{Mult}_{\lambda,\mathrm{i}}) = \lambda \qquad et \qquad \det(\mathrm{Ech}_{\mathrm{i},\mathrm{j}}) = -1.$$

Comme le déterminant est multiplicatif, la proposition suit directement du lemme.

Pour une matrice carrée de taille n, le calcul du déterminant par la formule récursive qui le définit est extrêmement long. On peut voir par récurrence qu'il a une complexité algorithmique d'ordre $n!$. Un moyen beaucoup plus rapide (de complexité n^2) est offert par le pivot de Gauss.

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Rappelons nous du théorème 2.45 qui nous dit qu'on peut transformer la matrice A en une matrice échelonnée $\tilde{A}$ en utilisant les opérations sur les lignes $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i},\mathrm{j}}$. Vu qu'on connait l'effet de ces opérations sur le déterminant, il suffit de savoir calculer le déterminant d'une matrice carrée échelonnée.

Ce calcul est particulièrement facile vu la proposition 3.16. Une matrice carrée échelonnée $\tilde{A} \in \mathcal{M}_n(\mathbb{R})$ est forcement triangulaire supérieure. Ses coefficients sur la diagonale valent soit 0 soit 1. Ainsi le déterminant de $\tilde{A}$ vaut soit 1 (si tous les coefficient diagonaux valent 1), soit 0. Rappelons également que l'unique matrice carrée échelonnée contenant que des 1 sur la diagonale est la matrice identité.

Ainsi, on arrive à la conclusion suivante.

Corollaire 3.21. *Soit* $A \in \mathcal{M}_n(\mathbb{R})$ *et* $\tilde{A} = P_k \dots P_1 A$ *la matrice échelonnée obtenue à partir de* A *en utilisant les opérations* $P_1, \dots, P_k$ *de type* $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ *et* $\mathrm{Ech}_{\mathrm{i},\mathrm{j}}$. *Alors,*

$$\det(A) = \begin{cases} 0, & si\ \tilde{A} \neq I_n; \\ \frac{1}{\det(P_k)\dots\det(P_1)}, & si\ \tilde{A} = I_n \end{cases}$$

Comme promis, on donne maintenant la preuve du théorème 3.18.

Preuve: [Théorème 3.18] Soit $A \in \mathcal{M}_n(\mathbb{R})$ et $\tilde{A} = P_k \dots P_1 A$ la matrice échelonnée obtenue à partir de A en utilisant les opérations $P_1, \dots, P_k$ de type $\mathrm{Add}_{\mathrm{i};\lambda,\mathrm{j}}$, $\mathrm{Mult}_{\lambda,\mathrm{i}}$ et $\mathrm{Ech}_{\mathrm{i},\mathrm{j}}$. On a vu dans la partie 2.4.3 que A est inversible, si et seulement si $\tilde{A} = I_n$. Mais le corollaire 3.21 nous dit que cela est équivalent à $\det(A) \neq 0$. □

Remarque 3.22. En général on n'est pas obligé de ramener A sous forme échelonnée pour calculer son déterminant. Il suffit d'utiliser $\mathrm{Add}_{i;\lambda,j}$, $\mathrm{Mult}_{\lambda,i}$ et $\mathrm{Ech}_{i,j}$ pour la ramener sous une forme triangulaire supérieure ou inférieure.
Dans certains cas on peut montrer par des moyens plus simple que le rang de A est strictement plus petit que n. Cela implique que A n'est pas inversible, donc que $\det(A) = 0$.

3.3.4 Compléments

Pour complétude, on mentionne le résultat suivant qui peut être utile en pratique. Il s'applique également à l'écriture par lignes. On va admettre ce résultat, même s'il suit directement de (3.3).

Proposition 3.23. *Soient $C_1, \dots, C_n, C_i' \in \mathcal{M}_{n,1}(\mathbb{R})$ une famille de colonnes et $\lambda \in \mathbb{R}$. Alors,*

$$\det\Big(C_1 \dots C_{i-1} \quad C_i + C_i' \quad C_{i+1} \dots C_n\Big)$$
$$= \det\Big(C_1 \dots C_{i-1} \quad C_i \quad C_{i+1} \dots C_n\Big) + \det\Big(C_1 \dots C_{i-1} \quad C_i' \quad C_{i+1} \dots C_n\Big),$$
$$\det\Big(C_1 \dots C_{i-1} \quad \lambda C_i \quad C_{i+1} \dots C_n\Big) = \lambda \det\Big(C_1 \dots C_{i-1} \quad C_i \quad C_{i+1} \dots C_n\Big).$$

On dit que le déterminant est une forme multi-linéaire *de la famille des colonnes.*

On viens de voir l'effet sur le déterminant de certaines opérations sur les lignes (proposition 3.19) et sur les colonnes (proposition 3.23). Les deux sont reliées par le lemme suivant. On va admettre ce résultat (il peut être montré en utilisant (3.3)).

Lemme 3.24. *Pour tout $A \in \mathcal{M}_n(\mathbb{R})$, $\det(A^T) = \det(A)$.*

3.4 Polynôme caractéristique

Une façon de déterminer les valeurs propres d'une matrice est donnée par le critère suivant. Fixons $A \in \mathcal{M}_n(\mathbb{R})$.

Proposition 3.25. *$\lambda \in \mathbb{R}$ est valeur propre de A si et seulement si $\det(\lambda I_n - A) = 0$.*

On peut vérifier (par exemple par récurrence sur la taille de A) que la fonction $\lambda \mapsto \det(\lambda I_n - A)$ est un polynôme de degré n en λ (et de coefficient dominant 1). Il est appelé le *polynôme caractéristique*[(viii)] de A. On le note $\chi_A(\lambda)$.

La proposition 3.25 nous dit que les valeurs propres de A sont exactement les racines de χ_A.

[(viii)] Charakteristisches Polynom

Preuve: On a déjà vu (remarque 3.3) que $\lambda \in \mathbb{R}$ est une valeur propre de A si et seulement si $A - \lambda I_n$ n'est pas inversible. Par ailleurs, le théorème 3.18 nous dit que cela est équivalent à $\det(\lambda I_n - A) = 0$. $\square$

Exercice 3.3.
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ une matrice triangulaire supérieure. Calculer χ_A et montrer que

$$Sp(A) = \{a_{1,1}, \dots, a_{n,n}\}.$$

3.4.1 Application: recherche de vecteurs propres

On est maintenant en mesure de décrire un algorithme pour chercher les valeurs et vecteurs propres d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$.

On commence par calculer le polynôme caractéristique de A, Ensuite on trouve ses racines, notons les $\lambda_1, \dots, \lambda_k$ et notons $m_1, \dots, m_k \in \mathbb{N}$ leur multiplicité.

Pour chaque $1 \leq i \leq k$, en utilisant le pivot de Gauss, on résout l'équation matricielle $(\lambda_i I_n - A)X = 0$, pour $X \in \mathcal{M}_n(\mathbb{R})$. L'ensemble des solutions est $E_{\lambda_i} = \mathrm{Ker}(\lambda_i I_n - A)$. Vu que $\lambda_i I_n - A$ n'est pas inversible, on va trouver un espace de solutions de dimension au moins 1. On a ainsi trouver les valeurs propres $\lambda_1, \dots, \lambda_k$ et les vecteurs propres associés.

Si on trouve des espaces propres avec $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = n$, alors A est diagonalisable (voir la proposition 3.9). De plus, la preuve de la proposition 3.9, décrit comment créer une base de $\mathbb{R}^n$ formée de vecteurs propres de A. Le théorème 3.7 (plus précisément sa preuve) nous dit alors comment créer une matrice $P \in \mathcal{M}_n(\mathbb{R})$ qui diagonalise A. Si par contre, $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) < n$, alors A n'est pas diagonalisable (voir à nouveau la proposition 3.9). On distingue trois situations:

(i) χ_A admet n racines distinctes (c.à-d. si $k = n$ et $m_1 = \cdots = m_n = 1$). Alors la matrice A est diagonalisable et chaque espace propre est de dimension 1. (Rappelons nous du corollaire 3.11 qui nous dit que si $|Sp(A)| = n$, alors A est diagonalisable.)

Exemple: Posons

$$A = \begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}.$$

On trouve par un calcul directe $\chi_A(\lambda) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = (\lambda - 1)(\lambda - 2)(\lambda - 3)$. Ainsi on déduit que A est bien diagonalisable et que ses valeurs propres sont $1, 2$ et 3. Un calcul rapide nous offre une base de vecteurs propres associés, notamment

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

(ii) χ_A est *scindé*, c.à-d. $m_1 + \cdots + m_k = n$, mais certaines racines ont des multiplicités plus grandes que 1. Alors, pour chaque $i = 1, \ldots, k$, on calcule $E_i = \mathrm{Ker}(\lambda_i I_n - A)$ et on trouve

$$1 \leq \dim(E_i) \leq m_i.$$

(Le fait que $\dim(E_i) \leq m_i$ va être admis). La matrice est diagonalisable si et seulement si $\dim(E_i) = m_i$ pour chaque i.

Exemple: Posons

$$B = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}.$$

Par calcul directe $\chi_B(\lambda) = (\lambda - 1)^2(\lambda - 3)$ et $\chi_C(\lambda) = (\lambda - 3)(\lambda + 1)^2$. En utilisant le pivot de Gauss, on peut calculer les espaces propres de B. On trouve

$$\mathrm{Ker}(I_3 - B) = \mathrm{Vect}\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right] \quad \text{et} \quad \mathrm{Ker}(3I_3 - B) = \mathrm{Vect}\left[\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right].$$

On en déduit que B est diagonalisable; une base de vecteurs propres est donnée par les bases de $\mathrm{Ker}(I_3 - B)$ et $\mathrm{Ker}(I_3 - B)$. Plus précisément, si on pose

$$P = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{alors} \quad P^{-1}BP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Le même type de calcul pour C nous permet d'obtenir ses espaces propres, à savoir

$$\mathrm{Ker}(3I_3 - C) = \mathrm{Vect}\left[\begin{pmatrix} 1/2 \\ 1 \\ 1 \end{pmatrix}\right] \quad \text{et} \quad \mathrm{Ker}(-I_3 - C) = \mathrm{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right].$$

Ainsi C n'est pas diagonalisable.

(iii) Si on travaille sur $\mathbb{R}$ il se peut que le polynôme χ_A n'ai pas toutes ses racines réelles. Si c'est le cas (c.à-d. si $m_1 + \cdots + m_k < n$) alors A n'est pas diagonalisable dans $\mathbb{R}$, c.à-d. il n'existe pas $P \in \mathcal{M}_n(\mathbb{R})$ inversible telle que $P^{-1}AP$ est diagonale. (Pour voir si A est diagonalisable dans $\mathbb{C}$, voir les points (i) et (ii).)

Exemple: Posons

$$D = \begin{pmatrix} 3 & -1 & 0 \\ 6 & -3 & 2 \\ 8 & -6 & 5 \end{pmatrix}.$$

Un calcul directe montre que $\chi_D(\lambda) = \lambda^3 - 5\lambda^2 + 9\lambda - 5 = (\lambda-1)(\lambda^2-4\lambda+5)$. En utilisant la résolution des équations d'ordre 2, on voir que $\lambda^2 - 4\lambda + 5$ n'as pas de racines réelles. Ainsi D n'est pas diagonalisable par une matrice de $\mathcal{M}_n(\mathbb{R})$.

Si on prend en compte les racines complexes, alors presque tout polynôme admet n racines distinctes. Il faut avoir quelques notions d'analyse mathématique pour donner un sens précis à cette affirmation. Informellement, on peut dire que, si $A \in \mathcal{M}_n(\mathbb{C})$ est prise "au hasard", alors elle est diagonalisable. Ou encore que toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ peut être approchée par des matrices diagonalisables.

En général, quand $n \geq 5$, les racines de χ_A n'admettent pas d'expression exacte. Ainsi, on doit se contenter d'approximations numériques pour trouver $Sp(A)$.

À retenir

- Pour une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ on dit qu'un vecteur non nul $X \in \mathbb{R}^n$ est vecteur propre de A si $AX = \lambda X$ pour un $\lambda \in \mathbb{R}$. Le scalaire λ s'appelle valeur propre de A. L'ensemble des valeurs propres est le spectre de A, noté $Sp(A)$.

- L'ensemble des vecteurs propres pour une valeur propre λ est le s.e.v.
$$E_\lambda(A) = \mathrm{Ker}(A - \lambda I_n).$$

- $\lambda \in Sp(A)$ si et seulement si $A - \lambda I_n$ n'est pas inversible.

- La matrice A est dite diagonalisable si elle est semblable à une matrice diagonale, c.-à-d. s'il existe $P \in \mathcal{M}_n(\mathbb{R})$ inversible, telle que $P^{-1}AP$ est diagonale.
Pas toutes les matrices sont diagonalisables!

- A est diagonalisable si et seulement s'il existe une base de $\mathbb{R}^n$ formée entièrement de vecteurs propres de A.

- A est diagonalisable si et seulement si $\sum_{\lambda \in Sp(A)} \dim(E_\lambda) = n$.

- Si A est diagonalisable par une matrice inversible P,
$$P^{-1}AP = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}, \quad \text{alors} \quad A^k = P \begin{pmatrix} \lambda_1^k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^k \end{pmatrix} P^{-1}.$$

- Les matrices symétriques sont diagonalisables dans $\mathcal{M}_n(\mathbb{R})$.

- Pour vérifier si une matrice est inversible, on utilise le déterminant: $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement si $\det(A) \neq 0$.

- Pour $A, B \in \mathcal{M}_n(\mathbb{R})$, $\det(AB) = \det(A)\det(B)$,

- L'effet des transformations sur les lignes:
  - multiplier une ligne de la matrice par λ multiplie le déterminant par λ;
  - échanger deux lignes multiplie le déterminant par -1;
  - rajouter une ligne à une autre ne change par le déterminant.

  Le déterminant peut se calculer par le pivot de Gauss.

- Le déterminant d'une matrice diagonale est le produit des coefficients diagonaux.

- Le polynôme caractéristique d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est $\chi_A(\lambda) = \det(\lambda I_n - A)$. Ses racines sont exactement les valeurs propres de A.

- La procédure de la partie 3.4.1 pour vérifier si une matrice est diagonalisable.

Références

[1] Gilbert, S., *Linear Algebra and Its Applications*, Academic Press (1976).

[2] Loser, F., *Un premier cours de logique*. http://www.math.ens.fr/~loeser/cours_26_01_2010.pdf.

[3] https://fr.wikipedia.org/wiki/Table_de_symboles_mathématiques
https://de.wikipedia.org/wiki/Liste_mathematischer_Symbole