Mathématiques
licence 2ème année

Table des matières

Première partie

Analyse

Chapitre 1

Intégration

Sommaire

1.1 Intégrale d'une fonction

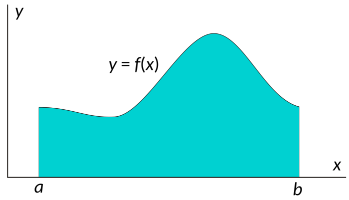
Dans tout ce cours, $a < b$ sont deux nombres réels. L'idée intuitive d'intégrale d'une fonction est celle "d'aire sous sa courbe" (au moins pour une fonction positive).

Soit f une fonction continue et positive définie sur un intervalle I de $\mathbb{R}$ et $[a, b] \subset I$. On souhaite calculer l'aire délimitée par :

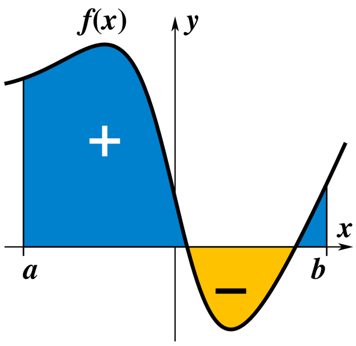
— l'axe réel ;

— le graphe de f ;

— les droites verticales $x = a$ et $x = b$.



Si f change de signe on compte positivement l'aire associée aux intervalles pour lesquels $f \geq 0$ et négativement l'aire associée aux intervalles pour lesquels $f \leq 0$.

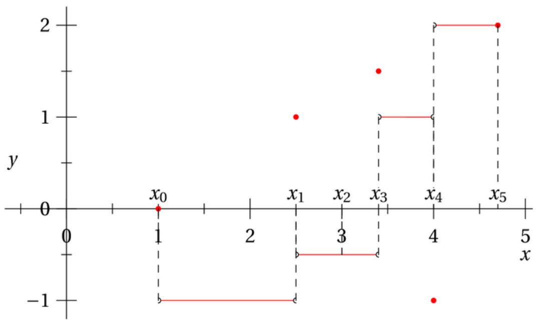


On va noter cette aire : $\mathcal{A} = \int_a^b f(x)dx$.

1.2 Intégrale des fonctions en escalier

Définition 1.1 (fonction en escalier)**.** Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite en escalier sur $[a, b]$ s'il existe $a = x_0 < x_1 < \ldots x_{n-1} < x_n = b$ tels que f soit contante sur chaque intervalle $]x_i, x_{i+1}[$, $i = 0, \ldots, n-1$.

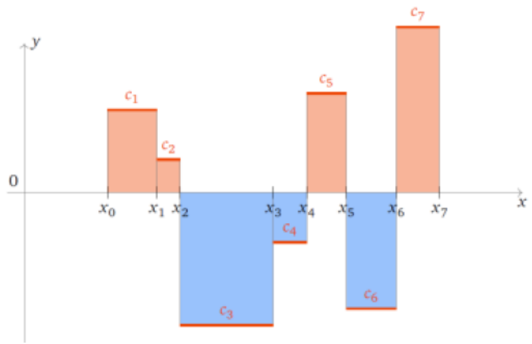
Exemple 1.



Exemple 2. *Soit* $f : [-1, 1] \to \mathbb{R}$ *donnée par*

$$f(x) = \begin{cases} -1 & si \ -1 \leq x \leq 0; \\ +1 & si \ 0 < x \leq 1. \end{cases}$$

L'intégrale d'une fonction en escalier revient à calculer des aires de rectangles :



Pour une fonction f en escalier sur $[a, b]$ et prenant la valeur y_i sur l'intervalle $]x_i, x_{i+1}[$, on a ainsi

$$\int_a^b f(x)dx = \sum_{i=0}^{n-1} y_i \times (x_{i+1} - x_i).$$

Exemple 3.

$$f(x) = \begin{cases} 1 & si \ x \in [0, 2], \\ 3 & si \ x \in [2, 3], \\ 2 & si \ x \in [3, 5], \\ 0 & sinon, \end{cases} \qquad \int_0^5 f(x)dx = 1 \times (2 - 0) + 3 \times (3 - 2) + 2 \times (5 - 3) = 9.$$

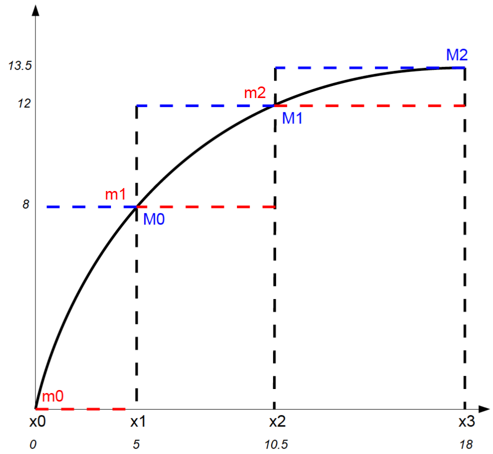
1.2.1 Intégrale des fonctions continues

Les résultats sont présentés ici pour les fonctions continues mais ils s'adaptent quand les fonctions sont seulement continues par morceaux.

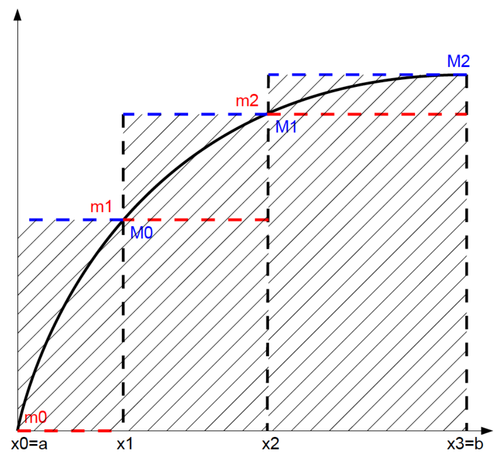
Considérons une fonction f continue sur $[a, b]$ et un ensemble de points $\{x_i\}$ tels que $a = x_0 < \cdots < x_n = b$.
Soit m_i la valeur minimale prise par f sur l'intervalle $[x_i, x_{i+1}]$ et M_i la valeur maximale de f sur ce même intervalle. On définit s_n et S_n par :

$$s_n = \sum_{i=0}^{n-1} m_i \times (x_{i+1} - x_i), \qquad S_n = \sum_{i=0}^{n-1} M_i \times (x_{i+1} - x_i).$$

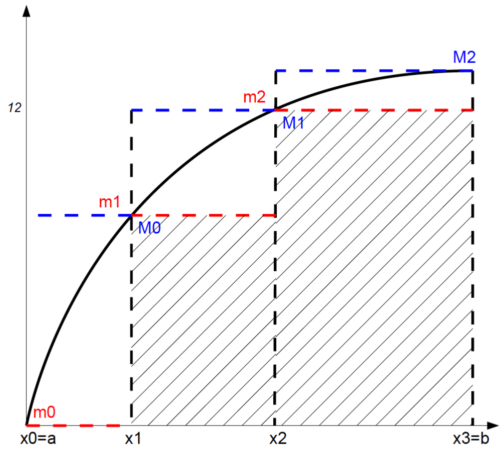
On les appelle respectivement petite et grande sommes de Darboux.



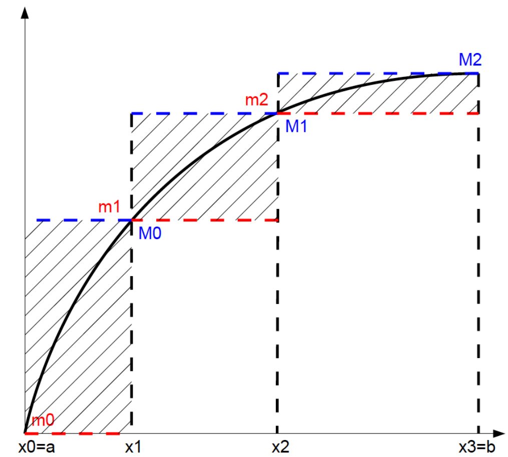
On obtient un encadrement de l'intégrale de la fonction f. La grande somme de Darboux S_n correspond à l'intégrale de la fonction en escalier bleue et elle majore l'intégrale de f.



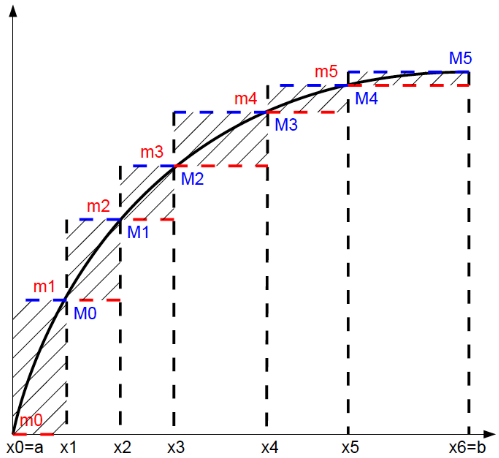
La petite somme de Darboux s_n correspond à l'intégrale de la fonction en escalier rouge et elle minore l'intégrale de f.



La différence entre les deux sommes est donnée par l'aire hachurée dans la figure suivante.



On raffine ensuite la subdivision, c'est-à-dire que l'on prend n de plus en plus grand avec $\max|x_{i+1} - x_i|$ qui tend vers 0 quand $n \to +\infty$.



Alors la différence entre s_n et S_n tend vers 0 et surtout ces deux suites tendent vers la même valeur quand $n \to +\infty$. Cette valeur est l'intégrale de f sur l'intervalle $[a, b]$. Elle est notée $\int_a^b f(x)dx$.

Plus généralement pour une fonction définie sur l'intervalle $[a, b]$ on a la définition suivante.

Définition 1.2. Une fonction f définie sur $[a, b]$ est dite intégrable si les deux sommes de Darboux s_n et S_n tendent vers une limite commune quand $n \to +\infty$. Leur limite commune est appelée intégrale de f sur $[a, b]$ et notée $\int_a^b f(x)dx$.

Comme on l'a vu plus haut

Proposition 1.3. *Une fonction continue sur $[a, b]$ est intégrable sur $[a, b]$.*

1.2.2 Propriétés de l'intégrale

De la limite des sommes de Darboux découle les propriétés suivantes.

Proposition 1.4. *Soient f et g deux fonctions intégrables sur $[a, b]$.*

1. *On a*

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

2. *Si $\lambda \in \mathbb{R}$ alors*

$$\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx.$$

3. *Relation de Chasles : pour tout $c \in [a,b]$, on a*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

4. *Relation d'ordre : Si $f \leq g$ sur $[a,b]$, c'est-à-dire, $f(x) \leq g(x)$ pour tout $x \in [a,b]$, alors*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Une conséquence du dernier point est l'inégalité suivante :

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

On remarquera que $\int_a^a f(x)dx = 0$ et la relation de Chasles permettent de considérer le cas de bornes dans l'ordre opposé à celui donné jusqu'à maintenant

$$\int_b^a f(x)dx = -\int_a^b f(x)dx.$$

1.2.3 Valeur moyenne d'une fonction

On appelle valeur moyenne d'une fonction sur $[a,b]$ la valeur suivante

$$\mu = \frac{1}{b-a}\int_a^b f(x)dx.$$

Une fonction constante sur $[a,b]$ ayant μ pour valeur a la même intégrale que f.

Des propriétés données plus haut on obtient

Si $m \leq f(x) \leq M$ pour tout $x \in [a,b]$ alors $m \leq \mu \leq M$.

1.3 Primitives

Définition 1.5. Soient f et F deux fonctions définies sur un intervalle I et à valeurs réelles. Si F est dérivable en tout point de I avec $F'(x) = f(x)$ pour tout $x \in I$ on dit que F est une primitive de f sur I.

Exemple 4.

— *$x \mapsto C$ avec $C \in \mathbb{R}$ est une primitive de la fonction nulle.*

— *$x \mapsto x^2/2$ est une primitive de $x \mapsto x$ sur $I = \mathbb{R}$;*

— *$x \mapsto x^2/2 + 3$ est* aussi *une primitive de $x \mapsto x$ sur $I = \mathbb{R}$.*

— $x \mapsto \exp(x)$ *est une primitive de* $x \mapsto \exp(x)$ *sur* $I = \mathbb{R}$.
— $x \mapsto x\ln(x) - x$ *est une primitive de* $x \mapsto \ln(x)$ *sur* $I =]0, +\infty[$.

On s'aperçoit que si F est une primitive de f alors $F + C$ où $C \in \mathbb{R}$ est aussi une primitive de f. Ainsi une fonction qui admet une primitive en admet une infinité.

En montrant que les fonctions contantes forment précisemment toutes les primitives de la fonction nulle on obtient le résultat suivant.

Proposition 1.6. *Soit f une fonction continue sur $I = [a, b]$. Deux fonctions F et G sont des primitives de f sur I* si et seulement si *$F - G$ est une fonction constante sur I.*

Donc si on connaît une primitive F de f on les connaît toutes : elles sont de la forme $F + C$ avec $C \in \mathbb{R}$.

Les notions d'intégrale et de primitive sont reliées comme expliqué par le théoréme suivant.

Théorème 1.7. *Soit f une fonction continue sur $I = [a, b]$.*

1. *Si F est une primitive de f alors*
$$\forall x, x' \in I, \int_x^{x'} f(t)dt = F(x') - F(x).$$
2. *La fonction $G(x) = \int_a^x f(t)dt$ est une primitive de f. Elle est donc dérivable et vérifie $G' = f$ sur I.*

Ce résultat permet en particulier d'affirmer que toute fonction continue admet des primitives.

On note
$$F(y) - F(x) = \Big[F\Big]_x^y. \tag{1.3.1}$$

Corollaire 1.8. *Soit f une fonction continue sur $I = [a, b]$, x dans I et un $y \in \mathbb{R}$ quelconque, alors il existe une primitive F de f telle que $F(x) = y$.*

Pour parler d'une primitive de la fonction f connue à une constante près, de fait du théorème 1.7, on écrit souvent
$$\int f(x)dx. \tag{1.3.2}$$

Mise en garde. Il faut se méfier de cette dernière notation (1.3.2). Elle désigne plus une classe de fonctions qui diffèrent d'une constante les unes des autres qu'une fonction particulière choisie.

Le théorème 1.7 explique pourquoi il est si utile de connaître un certain nombre de primitives usuelles. Dans le tableau suivant C dénote une constante réelle.

condition	fonction	primitive	intervalle
	0	C	$\mathbb{R}$
$a, b \in \mathbb{R}$	$ax+b$	$\frac{a}{2}x^2+bx+C$	$\mathbb{R}$
$n \in \mathbb{N}$	x^n	$\frac{x^{1+n}}{1+n}+C$	$\mathbb{R}$
$n \in \mathbb{N}$, $n \neq 1$, $n \neq 0$	x^{-n}	$\frac{x^{1-n}}{1-n}+C$	$]-\infty, 0[$ ou $]0, +\infty[$
	$x^{-1} = \frac{1}{x}$	$\ln\|x\|+C$	$]-\infty, 0[$ ou $]0, +\infty[$
$a > 0$, $a \notin \mathbb{N}$	x^a	$\frac{x^{1+a}}{1+a}+C$	$[0, +\infty[$
$a > 0$, $a \notin \mathbb{N}$	x^{-a}	$\frac{x^{1-a}}{1-a}+C$	$]0, +\infty[$
	$\exp(x)$	$\exp(x)+C$	$\mathbb{R}$
	$\ln(x)$	$x\ln(x)-x+C$	$]0, +\infty[$

En utilisant la formule de dérivation de deux fonctions composées

$$\frac{d}{dx}F(u(x)) = u'(x)F'(u(x)),$$

on peut compléter le tableau précédent.

condition	fonction	primitive
$n \in \mathbb{N}$	$u'(x)u(x)^n$	$\frac{u(x)^{1+n}}{1+n}+C$
$u(x) \neq 0$	$\frac{u'(x)}{u(x)}$	$\ln\|u(x)\|+C$
	$u'(x)e^{u(x)}$	$e^{u(x)}+C$

Proposition 1.9. *Soient f et g deux fonctions continues sur $I = [a, b]$ et F et G des primitives respectives.*

1. *$F + G$ est une primitive de $f + g$;*
2. *Si $\lambda \in \mathbb{R}$ alors λF est une primitive de λf.*

On prendra garde au fait que FG n'est pas une primitive de fg en général. L'exemple de $f = g = 1$ et $F = G = x$ permet de bien s'en rendre compte.

1.4 Quelques techniques pour le calcul d'intégrales

En pratique, dans ce que nous ferons dans ce cours, le calcul d'une intégrale repose sur l'utilisation de primitives et donc sur l'utilisation du théorème 1.7. On n'utilisera jamais la définition de l'intégrale comme limite des sommes de Darboux.

1.4.1 Intégration par parties

Prenons deux fonctions continues f et g sur l'intervalle $I = [a, b]$ et F, G des primitives respectives. On se fixe comme objectif le calcul de l'intégrale suivante

$$\int_a^b F(x)g(x)dx = \int_a^b F(x)G'(x)dx. \tag{1.4.1}$$

On rappelle la formule de dérivation d'un produit de deux fonctions :

$$(uv)' = u'v + uv'.$$

Ici, cela donne

$$(FG)' = fG + Fg.$$

Le deuxième terme est celui que l'on trouve dans l'intégrale dans (1.4.1). Si on écrit l'équation précédente de la manière suivante

$$Fg = (FG)' - fG,$$

après intégration cela donne, avec la notation introduite dans (1.3.1)

$$\int_a^b F(x)g(x)dx = \Big[FG\Big]_a^b - \int_a^b f(x)G(x)dx,$$

que l'on nomme *formule d'intégration par parties*, *IPP* pour faire court. On la retient plus sous la forme :

$$\int_a^b F(x)G'(x)dx = \Big[FG\Big]_a^b - \int_a^b F'(x)G(x)dx.$$

Si on utilise la notation introduite dans (1.3.2), la formule d'intégration par parties pour les primitives donne

$$\int F(x)G'(x)dx = F(x)G(x) - \int F'(x)G(x)dx.$$

Exemple 5. *Retrouver une primitive de* $x \mapsto \ln(x)$ *en faisant une intégration par parties.*

Exemple 6. *Calculs de*

1. $\displaystyle\int_1^3 x\ln(x)dx.$

$$\begin{aligned}\int_1^3 x\ln(x)dx &= \Big[\frac{x^2}{2}\ln(x)\Big]_1^3 - \int_1^3 \frac{x^2}{2x}dx = \frac{9}{2}\ln(3) - \frac{1}{2}\int_1^3 xdx \\ &= \frac{9}{2}\ln(3) - \frac{1}{4}\Big[x^2\Big]_1^3 = \frac{9}{2}\ln(3) - \frac{9-1}{4} = \frac{9}{2}\ln(3) - 2.\end{aligned}$$

2. $\displaystyle\int_0^1 xe^x dx.$

$$\int_0^1 xe^x dx = \Big[xe^x\Big]_0^1 - \int_0^1 e^x dx = e - \Big[e^x\Big]_0^1 = e - e + 1 = 1.$$

3. $\displaystyle\int_0^1 \frac{x}{\sqrt{1+2x}} dx.$

$$\begin{aligned}\int_0^1 \frac{x}{\sqrt{1+2x}} dx &= \Big[x\sqrt{1+2x}\Big]_0^1 - \int_0^1 \sqrt{1+2x} dx = \sqrt{3} - \Big[\frac{1}{3}(1+2x)^{3/2}\Big]_0^1 \\ &= \sqrt{3} - \sqrt{3} + \frac{1}{3} = \frac{1}{3}.\end{aligned}$$

1.4.2 Changements de variables

On peut se retrouver parfois avec une intégrale de la forme

$$\int_a^b f(\varphi(x))\varphi'(x) dx.$$

On a la proposition suivante.

Proposition 1.10 (formule de changement de variable)**.** *Si f est continue et φ est une fonction de classe $\mathscr{C}^1$ sur $[a,b]$ alors*

$$\int_a^b f(\varphi(x))\varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(y) dy.$$

Pour retenir la formule on pourra poser $y = \varphi(x)$. Alors, quand x varie dans $[a,b]$, y se « déplace » de $\varphi(a)$ à $\varphi(b)$. De plus on écrit

$$\frac{dy}{dx} = \varphi'(x),$$

ce qu'on écrit encore, de manière abusive,

$$dy = \varphi'(x)\, dx.$$

Cela a néanmoins du sens : à une variation infinitésimale de x, c'est-à-dire dx, correspond une variation infinitésimale de y qui est $\varphi'(x)\, dx$. C'est le sens même de la notion de dérivée d'une fonction.

Avec cette écriture on trouve

$$\int_a^b \underbrace{f(\varphi(x))}_{f(y)} \underbrace{\varphi'(x) dx}_{dy} = \int_{\varphi(a)}^{\varphi(b)} f(y) dy,$$

puisque la variable y varie entre $\varphi(a)$ et $\varphi(b)$.

N.B. Ce résultat explique aussi que « dx » apparaisse dans la notation de l'intégrale. Ce « dx » n'est pas juste une décoration inutile.

Démonstration de la formule de changement de variable. Soit F une primitive de f. La formule de dérivation de deux fonctions composées donne

$$\frac{d}{dx}F(\varphi(x)) = (F\circ\varphi)'(x) = \varphi'(x)F'(\varphi(x)) = \varphi'(x)f(\varphi(x)).$$

Ainsi $F\circ\varphi$ est une primitive de la fonction qui se trouve précisément dans l'intégrale. On a donc

$$\int_a^b f(\varphi(x))\varphi'(x)dx = \Big[F\circ\varphi(x)\Big]_a^b = F(\varphi(b)) - F(\varphi(a)) = \Big[F\Big]_{\varphi(a)}^{\varphi(b)} = \int_{\varphi(a)}^{\varphi(b)} f(y)dy.$$

□

Exemple 7. *Calculs de*

1. $\int_0^1 \frac{3x}{x^2+1}dx$.

$$\int_0^1 \frac{3x}{x^2+1}dx = \frac{3}{2}\int_0^1 \frac{2x}{x^2+1}dx = \frac{3}{2}\int_0^1 \frac{\varphi'(x)}{\varphi(x)}dx$$

avec $\varphi(x) = x^2+1$ *qui est strictement croissante sur* $[0,1]$. *On trouve donc*

$$\int_0^1 \frac{3x}{x^2+1}dx = \frac{3}{2}\int_1^2 \frac{1}{y}dy = \frac{3}{2}\Big[\ln(y)\Big]_1^2 = \frac{3}{2}(\ln(2)-\ln(1)) = \frac{3\ln(2)}{2}.$$

2. $\int_0^2 \frac{1}{\sqrt{x}+2x}dx$.

On pose $\varphi(x) = \sqrt{x}$. *On a* $\varphi' = \frac{1}{2\sqrt{x}}$. *On écrit*

$$\frac{1}{\sqrt{x}+2x} = \frac{2}{1+2\sqrt{x}} \times \frac{1}{2\sqrt{x}} = \frac{2}{1+2\varphi(x)}\varphi'(x)$$

On trouve alors

$$\int_0^2 \frac{1}{\sqrt{x}+2x}dx = \int_0^2 \frac{2}{1+2\varphi(x)}\varphi'(x)dx = \int_{\varphi(0)}^{\varphi(2)} \frac{2}{1+2y}dy = \int_0^{\sqrt{2}} \frac{2}{1+2y}dy$$
$$= \Big[\ln(1+2y)\Big]_0^{\sqrt{2}} = \ln(1+2\sqrt{2}).$$

Une autre facon d'écrire les choses est de poser $y = \sqrt{x}$. *Alors* $dy = \frac{1}{2\sqrt{x}}dx$ *et on écrit*

$$\frac{1}{\sqrt{x}+2x}dx = \frac{2\sqrt{x}}{\sqrt{x}+2x}dy = \frac{2y}{y+2y^2}dy = \frac{2}{1+2y}dy.$$

On trouve ainsi

$$\int_0^2 \frac{1}{\sqrt{x}+2x}dx = \int_0^{\sqrt{2}} \frac{2}{1+2y}dy,$$

et on poursuit avec le même calcul que précédemment.

3. $\int_1^e \frac{\ln(x)}{x+x(\ln(x))^2}dx$.
4. $\int_1^2 \frac{e^x}{e^x-1}dx$.

Deuxième partie

Algèbre

Chapitre 2

Systèmes linéaires et espaces vectoriels

Sommaire

2.1 Systèmes linéaires

2.1.1 Généralités

Définition 2.1. On appelle *système linéaire de n équations à m inconnues* $x_1, \ldots, x_m$ tout système d'équations de la forme

$$\mathcal{S} \quad \left\{ \begin{array}{ccccc} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,m}x_m & = & b_1 & & (L_1) \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + \cdots\cdots + a_{2,m}x_m & = & b_2 & & (L_2) \\ \vdots & = & \vdots & & \\ \vdots & = & \vdots & & \\ a_{n,1}x_1 + a_{n,2}x_2 + a_{n,3}x_3 + \cdots\cdots + a_{n,m}x_m & = & b_n & & (L_n) \end{array} \right.$$

où les $a_{i,j}$ et b_i pour $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, m\}$ sont des constantes données. Lorsque tous les b_i sont nuls, on dit que le système est homogène.

Résoudre le système $\mathcal{S}$, c'est déterminer (s'ils existent) les m-uplets $(x_1, \ldots, x_m)$ qui sont solution de toutes les équations du système.

Remarque 1. *Le m-uplet* $(0, \ldots, 0)$ *est une solution de tout système linéaire homogène.*

A priori, il est difficile de résoudre un système linéaire. Vous savez toutefois déjà très bien le faire dans certains cas particulier, comme les suivants.

Définition 2.2. Un système linéaire de n équations à m inconnues est dit *triangulaire* (ou *échelonné*) lorsque pour tout $i \in \{1, \ldots, n\}$ et pour tout $j \in \{1, \ldots, m\}$,

$$i > j \Rightarrow a_{i,j} = 0$$

Exemple 1. *1. Si $n = m$, un système triangulaire est de la forme*

$$\left\{\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,n}x_n & = & b_1 \\ a_{2,2}x_2 + a_{2,3}x_3 + \cdots\cdots + a_{2,n}x_n & = & b_2 \\ \vdots & = & \vdots \\ a_{n,n}x_n & = & b_n \end{array}\right.$$

2. Les systèmes suivants sont-ils échelonnés ?

$$\left\{\begin{array}{rcl} 2x + y & = & 0 \\ y & = & 4 \end{array}\right. \qquad \left\{\begin{array}{rcl} x + y + z + t & = & 0 \\ y + z + t & = & 1 \\ z + t & = & 4 \\ t & = & 2 \end{array}\right. \qquad \left\{\begin{array}{rcl} x + y + z & = & 0 \\ y + 2z & = & 1 \end{array}\right.$$

Exercice 1. *Résoudre les systèmes précédents.*

Les systèmes triangulaires sont donc très simples à résoudre, mais tous les systèmes linéaires ne sont pas triangulaires dans la nature. L'objectif de la suite va être premièrement de montrer comment réduire l'étude des solutions d'un système quelconque à celle d'un système triangulaire bien choisi.

2.1.2 L'ensemble des solutions d'un système linéaire

Le cas des systèmes linéaires homogènes est intéressant car leurs solutions possèdent des propriétés très particulières. En effet, étant donné un tel système

$$\mathcal{S} \quad \left\{\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,m}x_m & = & 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + \cdots\cdots + a_{2,m}x_m & = & 0 \\ \vdots & = & \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + a_{n,3}x_3 + \cdots\cdots + a_{n,m}x_m & = & 0 \end{array}\right.$$

on remarque que si $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ sont deux solutions de $\mathcal{S}$, alors leur somme $(x_1 + y_1, \ldots, x_n + y_n)$ est aussi une solution de $\mathcal{S}$. De même, si λ est un nombre réel, on vérifie aisément que $(\lambda x_1, \ldots, \lambda x_n)$ est aussi une solution de $\mathcal{S}$.

On peut donc "additionner" les solutions d'un système linéaire homogène, et on peut même les "multiplier par des nombres réels". Ces propriétés auront bientôt un nom : nous dirons que l'ensemble de ces solutions est "stable par addition" et "stable par multiplication par un scalaire".

Définition 2.3. Soit $\mathcal{S}$ un système linéaire de n équations à n inconnues. Le *système linéaire homogène* associé à $\mathcal{S}$, noté $\mathcal{S}_h$, est le système linéaire obtenu à partir de $\mathcal{S}$ en imposant que tous les seconds membres de ses équations soient nuls.

Exemple 2. *Le système homogène $\mathcal{S}_h$ associé à*

$$\mathcal{S} \quad \begin{cases} 2x+y &= 2 \\ 3x+y &= 4 \end{cases}$$

est le système

$$\mathcal{S}_h \quad \begin{cases} 2x+y &= 0 \\ 3x+y &= 0 \end{cases}$$

Soient $u=(x_1,\dots,x_n)$ et $v=(y_1,\dots,y_n)$ deux solutions d'un système linéaire $\mathcal{S}$ de n équations à n inconnues. Un simple calcul montre que le n-uplet

$$u-v=(x_1-y_1,\dots,x_n-y_n)$$

est une solution du système homogène $\mathcal{S}_h$ associé à $\mathcal{S}$.

De même, si $w=(z_1,\dots,z_n)$ est une solution de $\mathcal{S}_h$, alors

$$w+u=(z_1+x_1,\dots,z_n+x_n)$$

est une solution de $\mathcal{S}$. Les solutions du système $\mathcal{S}$ et du système homogène $\mathcal{S}_h$ associé sont donc intimement liées et on obtient sans peine le résultat suivant.

Théorème 2.4. *Un système linéaire $\mathcal{S}$ de n équations à n inconnues admet une unique solution si et seulement si le système homogène associé $\mathcal{S}_h$ admet lui-même une unique solution.*

2.1.3 Résolution d'un système linéaire

2.1.3.1 Opérations sur les lignes d'un système linéaire

On note

$$\mathcal{S} \quad \begin{cases} a_{1,1}x_1+a_{1,2}x_2+a_{1,3}x_3+\cdots\cdots+a_{1,m}x_m &= b_1 \qquad (L_1) \\ a_{2,1}x_1+a_{2,2}x_2+a_{2,3}x_3+\cdots\cdots+a_{2,m}x_m &= b_2 \qquad (L_2) \\ \vdots &= \vdots \\ \vdots &= \vdots \\ a_{n,1}x_1+a_{n,2}x_2+a_{n,3}x_3+\cdots\cdots+a_{n,m}x_m &= b_n \qquad (L_n) \end{cases}$$

un système linéaire de n équations à m inconnues.

Définition 2.5. On désigne les *opérations de Gauss* sur le lignes de $\mathcal{S}$ ainsi :

1. Échange de deux lignes (L_i) et (L_j), noté $L_i \leftrightarrow L_j$

2. Remplacement d'une ligne (L_i) par elle-même, multipliée par un nombre réel λ non-nul, noté
$$L_i \leftarrow \lambda L_i$$
3. Remplacement d'une ligne (L_i) par elle-même, à laquelle on ajoute un multiple $\lambda(L_j)$ d'une *autre* colonne (L_j), noté
$$L_i \leftarrow L_i + \lambda L_j$$

On dit que deux systèmes $\mathcal{S}$ et $\mathcal{S}'$ sont *équivalents* si l'on peut passer de l'un à l'autre par des opérations de Gauss sur leurs lignes.

Théorème 2.6. *Soient $\mathcal{S}$ et $\mathcal{S}'$ deux systèmes linéaires. Si $\mathcal{S}$ et $\mathcal{S}'$ sont équivalents, alors ils ont le même ensemble de solutions.*

Le théorème précédent est essentiel car il montre que dès qu'il s'agit d'étudier les solutions d'un système $\mathcal{S}$, on peut s'autoriser les opérations de Gauss pour tenter de le rendre plus simple. Notre travail précédent n'est donc pas inutile : nous allons montrer que l'on peut transformer tout système linéaire par les opérations des Gauss en un système *triangulaire*, que l'on sait très bien résoudre et qui possède le même ensemble de solutions par le théorème précédent.

2.1.3.2 L'algorithme de Gauss

L'algorithme de Gauss consiste à transformer un système $\mathcal{S}$ d'équations linéaires en un système linéaire triangulaire à partir des opérations définies cidessus. L'idée est de procéder à l'élimination successive des inconnues permettant de rendre le système triangulaire. Plus précisément, fixons

$$\mathcal{S} \quad \left\{ \begin{array}{lcll} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,m}x_m & = & b_1 & (L_1) \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + \cdots\cdots + a_{2,m}x_m & = & b_2 & (L_2) \\ \vdots & = & \vdots & \\ \vdots & = & \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + a_{n,3}x_3 + \cdots\cdots + a_{n,m}x_m & = & b_n & (L_n) \end{array} \right.$$

On va supposer pour simplifier que le coefficient $a_{1,1}$ n'est pas nul (si ce n'est pas le cas, il suffit d'échanger deux lignes pour s'y ramener). Sachant que $a_{1,1}$ n'est pas nul, on peut effectuer les opérations suivantes :

$$L_2 \leftarrow L_2 - \frac{a_{2,1}}{a_{1,1}} L_1$$

$$L_3 \leftarrow L_3 - \frac{a_{3,1}}{a_{1,1}} L_1$$

$$\vdots$$

$$L_n \leftarrow L_n - \frac{a_{n,1}}{a_{1,1}} L_1.$$

Ces opérations ont le bon goût d'éliminer l'inconnue x_1 des lignes L_2, L_3,...,L_n, si bien qu'après calcul, on obtient un nouveau système

$$\mathcal{S}' \quad \left\{ \begin{array}{rcll} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,m}x_m & = & b_1 & (L_1) \\ a'_{2,2}x_2 + a'_{2,3}x_3 + \cdots\cdots + a'_{2,m}x_m & = & b'_2 & (L'_2) \\ \vdots & = & \vdots & \\ \vdots & = & \vdots & \\ a'_{n,2}x_2 + a'_{n,3}x_3 + \cdots\cdots + a'_{n,m}x_m & = & b'_n & (L'_n) \end{array} \right.$$

qui est équivalent à $\mathcal{S}$. On peut désormais "oublier" la première ligne et faire exactement la même méthode pour le sous-système obtenu avec les lignes L'_2,...,L'_n. Plus précisément, on peut à nouveau supposer que $a'_{2,2}$ est non-nul, et les opérations

$$L'_3 \leftarrow L'_3 - \frac{a'_{3,1}}{a'_{2,2}} L'_2$$

$$L'_4 \leftarrow L'_4 - \frac{a'_{4,1}}{a'_{2,2}} L'_2$$

$$\vdots$$

$$L'_n \leftarrow L'_n - \frac{a'_{n,1}}{a_{2,2}} L'_2.$$

nous font aboutir à un nouveau système

$$\mathcal{S}'' \quad \left\{ \begin{array}{rcll} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots\cdots + a_{1,m}x_m & = & b_1 & (L_1) \\ a'_{2,2}x_2 + a'_{2,3}x_3 + \cdots\cdots + a'_{2,m}x_m & = & b'_2 & (L'_2) \\ a''_{2,3}x_3 + \cdots\cdots + a''_{2,m}x_m & = & b''_2 & (L''_3) \\ \vdots & = & \vdots & \\ a''_{n,3}x_3 + \cdots\cdots + a''_{n,m}x_m & = & b''_n & (L''_n) \end{array} \right.$$

pour lequel la variable x_2 est éliminé à partir de troisième ligne, et lui aussi équivalent à $\mathcal{S}$.

On peut continuer ainsi le processus pour construire un système triangulaire équivalent à $\mathcal{S}$, en au plus $n-1$ étapes, mais plutôt que de l'écrire jusqu'au bout ici, la meilleure façon de saisir l'algorithme de Gauss est de le voir en action et le pratiquer lors de l'exercice suivant.

Exercice 2. *Transformer les systèmes suivants en systèmes triangulaires grâce à l'algorithme de Gauss, puis déterminer leurs ensembles de solutions.*

$$\left\{\begin{array}{rcl} 2x-y+z & = & 0 \\ x+y-2z & = & 0 \\ 2x-y-2z & = & 0 \end{array}\right. \qquad \left\{\begin{array}{rcl} x+y+2z & = & 3 \\ x+2y+z & = & 1 \\ 2x+y+z & = & 0 \end{array}\right. \qquad \left\{\begin{array}{rcr} x+y+z-3t & = & 1 \\ 2x+y-z+t & = & -1 \end{array}\right.$$

2.2 Espaces vectoriels

Si $n \geq 2$, le produit cartésien de n copies de $\mathbb{R}$ est noté $\mathbb{R}^n$. C'est l'ensemble des n-uplets (ou vecteurs) $(x_1, \ldots, x_n)$, où $x_1, \ldots, x_n$ sont des nombres réels. Nous avons vus précédemment que $\mathbb{R}^n$ est muni des deux opérations suivantes :

1. *Addition des vecteurs.* Si $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ sont deux vecteurs de $\mathbb{R}^n$, on définit leur somme par
$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix}$$
2. *Multiplication par un scalaire.* Si $(x_1, \ldots, x_n)$ est un vecteur de $\mathbb{R}^n$ et si $\lambda \in \mathbb{R}$ est un scalaire, on définit leur multiplication par
$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

L'exemple de $\mathbb{R}^n$ muni des opérations ci-dessus est aussi simple qu'il est essentiel : nous allons voir très vite la définition d'*espace vectoriel* qui s'en inspire fortement. Pour préparer un petit peu le terrain, attardons-nous un peu sur $\mathbb{R}^n$.

Définition 2.7. Un sous-ensemble W de $\mathbb{R}^n$ est dit *stable par addition*, si la somme de deux vecteurs de W est elle aussi un élément de W.

De même, on dit que W est *stable par multiplication par un scalaire* si pour tout vecteur $w \in W$ et pour tout réel λ le vecteur λw appartient aussi à W.

Exemple 3.

1. *Le sous-ensemble* $W = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ *de* $\mathbb{R}^3$ *est stable par addition et par multiplication par un scalaire. Qu'en est-il de* $W' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ *?*

2. *Une droite* L *de* $\mathbb{R}^2$ *qui passe par* $(0,0)$ *est stable par addition et par multiplication par un scalaire.*

3. *L'ensemble des solutions d'un système homogène* $\mathcal{S}$ *est stable par addition et par multiplication par un scalaire.*

4. *Le sous-ensemble* $W = \{(x,y) \in \mathbb{R}^2, x = 0 \text{ ou } y = 0\}$ *est-il stable par addition ? par multiplication par un scalaire ?*

5. $A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, 7x + 8y = 0 \right\}$ *est-il stable par addition ? Par multiplication par un scalaire ?*

6. $B = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, y = x^2 \right\}$ *est-il stable par addition ? Par multiplication par un scalaire ?*

7. *L'ensemble des fonctions* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *qui vérifient l'équation différentielle* $f' = f$ *est-il stable par addition ? par multiplication par un scalaire ? (la fonction exponentielle par exemple est une solution de solution de cette équation).*

2.2.1 Définition des espaces vectoriels

Nous allons maintenant en venir à la définition d'un *espace vectoriel.* L'idée est d'essayer de construire des ensembles qui se comportent de manière assez semblable à $\mathbb{R}^n$, c'est-à-dire de considérer des ensembles dont on peut *additionner les éléments* et dont les éléments peuvent être *multipliés par un nombre réel.*

Il s'agit donc en un sens d'extraire la spécificité de $\mathbb{R}^n$, mais l'addition et la multiplication par un scalaire que nous avons définis pour $\mathbb{R}^n$ cache beaucoup de petits détails qui semblent évidents au premier abord car nous sommes habitués aux nombres réels, mais ne le sont en fait pas du tout ! Il faut donc prendre garde à ces détails, c'est pour cela que la définition suivante est un peu longue.

Définition 2.8. On dit qu'un ensemble V est un *espace vectoriel*, s'il est muni de deux lois particulières :

1. La première est l'addition "+" entre les élements de V : si v et w sont deux éléments de V, on peut considérer leur somme dans V, notée $v + w$;

2. La seconde est la multiplication "$\cdot$" des éléments de V par les nombres réels : si $v \in V$ et si $\lambda \in \mathbb{R}$, on a un élément λv de V qu'on appelle le multiple de v par λ.

En outre, ces lois doivent vérifier de nombreuses propriétés. Les voici pour "+" :

i) "+" est *commutative* : si $(v, w) \in V^2$, on a
$$v + w = w + v$$

ii) Il existe un élément $0_V \in V$ qui est *l'élément neutre* de la loi $+$, c'est-à-dire que pour tout $v \in V$,
$$v + 0_V = 0_V + v = v$$

iii) Tout élément $v \in V$ admet un *opposé*, c'est-à-dire qu'il existe un élément de V, noté $-v$, tel que
$$v + (-v) = (-v) + v = 0_V$$

iv) La loi "+" doit être *associative* : si v, w, et u sont des éléments de V, alors
$$(u + v) + w = u + (v + w)$$

La loi "$\cdot$" doit quant à elle vérifier les propriétés suivantes :

i) La loi "$\cdot$" est associative : si λ et μ sont des réels, et si $v \in V$, alors
$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$$

ii) Si $v \in V$, alors $1 \cdot v = v$ (eh oui, il faut l'imposer).

Enfin, les deux lois doivent les deux lois "+" et "$\cdot$" doivent vérifier deux propriétés dites de distributivité :

i) Si $v \in V$ et si λ et μ sont deux nombres réels, on a
$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$$

ii) Si $\lambda \in \mathbb{R}$ et si v et w sont deux éléments de V, on a
$$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$$

Pas de panique. Cette définition est très longue, fastidieuse et nécessaire pour la suite du cours mais en pratique, vous n'aurez pas à démontrer *toutes* ces propriétés pour construire votre premier espace vectoriel. Elles sont néanmoins toutes nécessaires (même celle qui stipule que $1 \cdot v = v$!) comme en témoigne le petit exercice suivant dont l'énoncé semble évident... Mais qui ne l'est en fait pas du tout.

Exercice 3. *Soit V un espace vectoriel, dont on note "+" et "·" les deux lois.*
Montrer si λ est un nombre réel, $\lambda \cdot 0_V = 0_V$.
Montrer de même que si $v \in V$, alors $0 \cdot v = 0_V$.
En déduire que si $v \in V$, $(-1) \cdot v = -v$.

Exemple 4.

1. *Bien entendu, $\mathbb{R}^n$ muni de l'addition et de la multiplication définies plus haut est bien un espace vectoriel.*

2. *Plus généralement, tous les sous-ensembles de $\mathbb{R}^n$ apparus dans l'exemple 3 ci-dessus qui sont stables par addition et par multiplication par un scalaire sont eux aussi des espaces vectoriels, pour les mêmes lois que $\mathbb{R}^n$.*

3. *Il existe des espaces vectoriels qui ne sont pas construits à partir de $\mathbb{R}^n$. Par exemple, l'ensemble $\mathcal{S}$ est suites réelles est un espace vectoriel. Pour le montrer, on peut introduire les lois suivantes :*
   i) la somme de deux suite $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ est la suite $u + v$ définie par
   $$(u + v)_n = u_n + v_n.$$
   ii) de même si $u = (u_n)_{n\in\mathbb{N}}$ est une suite réelle et λ est un réel, la suite $\lambda \cdot u$ est donnée par
   $$(\lambda \cdot u)_n = \lambda u_n.$$

2.3 Sous-espaces vectoriels

Comme vous l'imaginez, démontrer qu'un ensemble munies de deux lois "+" et "·" (on note souvent cela $(V, +, \cdot)$) est un espace vectoriel en vérifiant toutes les propriétés de la Définition 2.8 est très laborieux. Dans la quasi-totalité des cas, on parvient à éviter cela grâce à la notion de sous-espace vectoriel.

Définition 2.9. Soit $(V, +, \cdot)$ un espace vectoriel et $W \subset V$ un sous-ensemble non-vide de V. On dit que W est un sous-espace vectoriel de V (abrégé "sev de V") s'il est stable par addition ainsi que par multiplication par un scalaire, autrement dit

1. si $(v, w) \in W^2$, $v + w \in W$;
2. si $(\lambda, v) \in \mathbb{R} \times W$, $\lambda \cdot v \in W$.

Ces propriétés sont assez rapides à montrer et permettent de construire de nombreux espaces vectoriels. En effet, on a le résultat suivant.

Proposition 2.10. *Si $(V, +, \cdot)$ est un espace vectoriel et si on considère $W \subset V$ un sous ensemble non-vide sous-espace vectoriel de W au sens de la définition 2.9, alors $(W, +, \cdot)$ est automatiquement un espace vectoriel avec $0_W = 0_V$.*

Cela provient du fait que toutes les propriétés pénibles restantes (associativité, etc) sont automatiques, car déjà vraies dans V.

Il faut prendre garde pour rester rigoureux de ne pas oublier de vérifier que W n'est pas l'ensemble vide, lorsqu'on essaie de montrer qu'il est un sous-espace vectoriel. La petite observation suivante permet de voir qu'il suffit pour cela de considérer 0_V.

Remarque 2. *Si $W \subset V$ est un sous espace vectoriel de $(V, +, \cdot)$, alors $0_V \in W$. En effet W n'est pas vide, donc il contient un élément v. Par définition, W doit donc aussi contenir l'élément $(-1) \cdot v = -v$... Mais alors W contient même l'élément $v + (-v) = 0_V$.*

Exemple 5. *Si $(V, +, \cdot)$ est un espace vectoriel alors $\{0_V\}$ est un sous-espace vectoriel de V. Notons que V est lui-même un sous-espace vectoriel de V.*

Exemple 6. *On considère les deux sous-ensembles de $\mathbb{R}^2$*

$$A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \ 2x + y = 0 \right\}$$

et

$$B = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \ 2x + y = 1 \right\}.$$

Sont-ils des sous-espaces vectoriels de $\mathbb{R}^2$?

La notion de sous-espace vectoriel est stable par intersection.

Proposition 2.11. *Soient $(V, +, \cdot)$ un espace vectoriel et W_1 et W_2 deux sous-espaces vectoriels de V. Alors $W_1 \cap W_2$ est un sous-espace vectoriel de V.*

Si A est une partie de V, A n'est pas nécessairement stable par l'addition et la multiplication par un scalaire. Il devient alors intéressant de chercher le plus petit ensemble de V qui contient A et qui possède ces deux propriétés. Cette ensemble est alors le plus petit sous-espace vectoriel qui contient la partie A. La proposition 2.11 permet de conclure à la propriété suivante.

Proposition 2.12. *Soient $(V, +, \cdot)$ un espace vectoriel et $A \subset V$. Le plus petit sous-espace vectoriel qui contient A est donné par l'intersection de tous les sous-espaces vectoriels de V qui contiennent A.*

Notons que l'intersection réalisée dans la proposition précédente

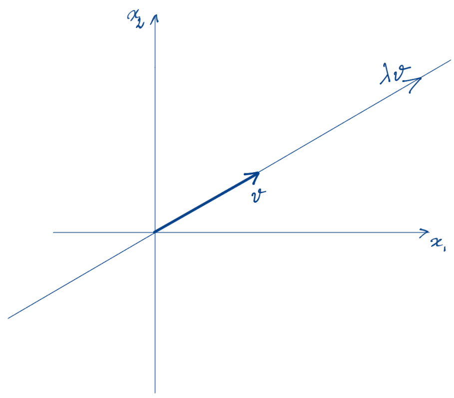
— implique bien au moins un sous-espace vectoriel puisque est V est lui-même un sous-espace vectoriel de V (voir l'exemple 5) ;

— est non vide puisque tous les sous-espace vectoriels de V contiennent au moins un élément en commun : le vecteur nul 0_V (voir la remarque 2).

2.3.1 Sous-espaces vectoriels engendrés

Nous avons vu précédemment que si D est une droite de $\mathbb{R}^2$ qui passe par $(0, 0)$, alors D en est un sous-espace vectoriel. En effet, on peut considérer un vecteur $v \in \mathbb{R}^2$ non-nul porté par D et observer que tous les éléments de D sont les multiples de v, c'est-à-dire plus formellement :

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



La droite D est donc un espace vectoriel qui contient v, et on voit même que D est *le plus petit* sous-espace vectoriel de $\mathbb{R}^2$ qui contienne v (car un sous-espace vectoriel qui contient v doit contenir tous ses multiples). La notion de sous-espace vectoriel engendré par une partie que nous présentons maintenant va coïncider avec celle de *plus petit sous-espace vectoriel* contenant cette partie que nous avons vue plus haut.

Nous commençons par le cas d'une famille finie de vecteurs, généralisant ainsi le cas d'une droite.

2.3.1.1 Sous-espace vectoriel engendré par une famille finie de vecteurs

Définition 2.13. Soit $(V, +, \cdot)$ un espace vectoriel et $\mathcal{F} = \{v_1, \ldots, v_k\}$ un ensemble fini de vecteurs de V. Une combinaison linéaire des vecteurs de $\mathcal{F}$ est un élément v de V de la forme

$$v = a_1 v_1 + \cdots + a_k v_k$$

où $a_1, \ldots, a_k$ sont des réels.

L'ensemble des combinaisons linéaires des vecteurs de $\mathcal{F} = \{v_1, \ldots, v_k\}$ est noté $\mathrm{Vect}(\{v_1, \ldots, v_k\})$ ou $\mathrm{Vect}(\mathcal{F})$.

Exemple 7. *Tout vecteur de $\mathbb{R}^2$ est une combinaison linéaire des deux vecteurs $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ puisqu'on peut toujours écrire*

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

En revanche, le vecteur $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ n'est pas une combinaison linéaire du vecteur $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Exercice 4. *Montrer que si $\mathcal{F} = \{v_1, \ldots, v_k\}$ est un ensemble de vecteurs d'un espace vectoriel V, alors $\mathrm{Vect}(\mathcal{F})$ est un sous-espace vectoriel de V.*

Grâce à cet exercice on obtient le résultat suivant.

Proposition 2.14. *Si $\mathcal{F} = \{v_1, \ldots, v_k\}$ est un ensemble de vecteurs d'un espace vectoriel V alors $\mathrm{Vect}(\mathcal{F}) = \mathrm{Vect}(\{v_1, \ldots, v_k\})$ est le plus petit sous-espace vectoriel de V qui contient les vecteurs $v_1, \ldots, v_k$.*

Par la proposition 2.12 $\mathrm{Vect}(\{v_1, \ldots, v_k\})$ est donc l'intersection de tous les sous-espaces vectoriels de V qui contiennent les vecteurs $v_1, \ldots, v_k$.

On dira alors que $\mathrm{Vect}(\{v_1, \ldots, v_k\})$ est l'*espace vectoriel engendré par les vecteurs $v_1, \ldots, v_k$*.

La notion d'espace vectoriel engendré se généralise sans difficulté à une famille infinie de vecteurs, comme lorsqu'on considère par exemple une partie quelconque de V comme cela est fait dans la proposition 2.12. Bien que nous n'utiliserons que très peu cela dans la suite, nous en faisons une brève description maintenant.

2.3.1.2 (∗) Sous-espace vectoriel engendré par une famille infinie de vecteurs

L'étoile (∗) qui précède le titre de la section souligne que cette section n'est pas essentielle et peut être omise en première lecture.

Si V est un espace vectoriel, une famille infinie de vecteurs de V peut être notée $\mathcal{F} = (v_i)_{i\in\mathcal{I}}$, avec I ensemble infini et chaque v_i dans V, ou bien tout simplement $\mathcal{F}$ une partie de V.

Une combinaison linéaire de vecteurs de cette famille $\mathcal{F}$ correspond

1. au choix d'un nombre fini $w_1, \dots, w_k$ de vecteurs de $\mathcal{F}$;
2. au choix de scalaires $a_1, \dots, a_k$.

La combinaison linéaire est alors $a_1w_1 + \cdots + a_kw_k$.

On note que l'on réalise uniquement des combinaisons linéaires *finies* de vecteurs de $\mathcal{F}$. On ne s'intéresse pas à une éventuelle somme infinie dont le sens ne serait pas clair du tout.

Une fois le concept de combinaison linéaire étendue aux familles infinies, on note $\mathrm{Vect}(\mathcal{F})$ l'ensemble des combinaisons linéaires que l'on peut réaliser à partir d'une telle famille $\mathcal{F}$ et on obtient le résultat suivant analogue aux précédents.

Proposition 2.15. *Soit $\mathcal{F} \subset V$. Alors $\mathrm{Vect}(\mathcal{F})$ est le plus petit sous-espace vectoriel de V qui contient les vecteurs de $\mathcal{F}$.*

C'est donc l'intersection de tous les sous-espaces vectoriels de V qui contiennent $\mathcal{F}$.

2.3.2 Famille génératrice

Soit $\mathcal{F}$ une famille de vecteurs de V (finie ou infinie).

Définition 2.16 (famille génératrice)**.** Si $V = \mathrm{Vect}(\mathcal{F})$ on dit que V est engendré par les vecteurs de $\mathcal{F}$ et que $\mathcal{F}$ est une famille génératrice de V.

Dans ce cas tout vecteur de V s'écrit comme une combinaison linéaire des vecteurs de la famille $\mathcal{F}$. (Dans le cas d'une famille $\mathcal{F}$ infinie la notion de combinaison linéaire est celle donnée dans la section 2.3.1.2.)

Exemple 8. *Nous avons vu précédemment que les vecteurs $\begin{pmatrix}1\\0\end{pmatrix}$ et $\begin{pmatrix}0\\1\end{pmatrix}$ engendrent $\mathbb{R}^2$ et il est clair que pour des raisons similaires, $\mathbb{R}^n$ est engendré par les n vecteurs*

$$\left\{ e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \cdots, e_n = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} \right\}.$$

Bien sûr, il existe des exemples plus exotiques et ils sont essentiels pour la suite de ce cours. Par exemple, on peut observer que les vecteurs $\begin{pmatrix}1\\0\end{pmatrix}$ *et* $\begin{pmatrix}1\\1\end{pmatrix}$ *engendrent eux aussi* $\mathbb{R}^2$*. On peut pour le montrer utiliser que*

$$\begin{pmatrix}0\\1\end{pmatrix} = -\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix}$$

et donc on a bien pour tout vecteur $(x, y) \in \mathbb{R}^2$*,*

$$\begin{pmatrix}x\\y\end{pmatrix} = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix} = x\begin{pmatrix}1\\0\end{pmatrix} + y\left(-\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix}\right) = (x-y)\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}1\\1\end{pmatrix}.$$

2.3.3 Familles libres, bases

Nous avons vu dans la partie précédente avec la notion de *sous-espace vectoriel engendré* comment on peut construire à partir d'une famille de vecteurs $\mathcal{F}$ de V un sous-espace vectoriel dont les éléments sont tous des *combinaisons linéaires* des vecteurs de $\mathcal{F}$. Nous allons désormais introduire une nouvelle propriété pour $\mathcal{F}$ qui permet, lorsqu'elle est vérifiée, de décrire $\mathrm{Vect}(\mathcal{F})$ avec encore plus de précision.

Définition 2.17 (famille libre – cas fini)**.** Soit V un espace vectoriel et $\mathcal{F} = \{v_1, \ldots, v_k\}$ une famille finie de vecteurs de V. On dit que $\mathcal{F}$ est *libre* si l'équation

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0_V$$

n'est vérifiée que si les réels $\lambda_1, \ldots, \lambda_k$ sont tous nuls.

Autrement dit, la seule combinaison linéaire des vecteurs de $\mathcal{F}$ qui donne le vecteur nul 0_V est celle où tous les coefficients scalaires sont nuls.

Exercice 5. *Montrer que la famille* $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$ *est libre.*

Observer que $\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}\right\}$ *n'est pas libre. Plus généralement, montrer qu'une famille* $\{v_1, v_2\}$ *constituée de deux vecteurs non-nuls est libre si et seulement si* v_1 *et* v_2 *ne sont pas colinéaires. (On rappelle que par définition,* v_1 *et* v_2 *sont dits colinéaires s'il existe un réel* λ *tel que* $v_1 = \lambda v_2$ *ou* $v_2 = \lambda v_1$*.)*

Comme pour la notion de famille génératrice, celle de famille libre peut-être étendue au cas des familles infinies, grâce à l'extension de la notion de combinaison linéaire faite pour de telles familles dans la section 2.3.1.2. Comme plus haut cette notion peut-être omise dans une première lecture.

Définition 2.18 ((∗) famille libre – cas infini)**.** Soit V un espace vectoriel et $\mathcal{F}$ une famille infinie de vecteurs. Cette famille est dite *libre* si pour tous $v_1, \dots, v_\ell$, vecteurs de la famille $\mathcal{F}$, et pour tous scalaires $\lambda_1, \dots, \lambda_\ell$ avoir

$$\lambda_1 v_1 + \cdots + \lambda_\ell v_\ell = 0_V,$$

implique nécessairement $\lambda_1 = \cdots = \lambda_\ell = 0$.

Autrement dit, $\mathcal{F}$ est une famille libre si toute sous-famille finie est libre au sens de la définition 2.17.

On définit maintenant un nouveau type de famille de vecteurs de V.

Définition 2.19 (Base d'un espace vectoriel)**.** Soit V un espace vectoriel et $\mathcal{F}$ une famille de vecteurs de V. Si $\mathcal{F}$ est une famille libre et si elle engendre V on dit que $\mathcal{F}$ est une *base* de V.

Proposition 2.20 (Décomposition unique dans une base)**.** *Si $\mathcal{F}$ est une base de V, alors tout vecteur de V s'écrit de manière* unique *comme combinaison linéaire de vecteurs de $\mathcal{F}$.*

Si $\mathcal{F} = \{v_1, \dots, v_k\}$ est une base finie de V et si

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

pour des réels $\lambda_1, \dots, \lambda_k$ on dira que $(\lambda_1, \dots, \lambda_k)$ sont les *coordonnées* de v dans la base $\mathcal{F}$.

Démonstration de la proposition 2.20. Pour simplifier on donne la démonstation dans le cas d'une famille finie uniquement. Le cas infini n'apporte rien à la compréhension de l'argument. On considère donc $\mathcal{F} = \{v_1, \dots, v_k\}$.

Soit $v \in V$. Puisque la famille $\mathcal{F}$ est une base de V, elle engendre V et on a donc des réels $\lambda_1, \dots, \lambda_k$ tels que

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k.$$

Il reste à montrer que cette représentation est unique, et c'est ici qu'on utilise le fait que $\mathcal{F}$ est libre. En effet, si on considère deux écritures

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = v = \mu_1 v_1 + \cdots + \mu_k v_k$$

de v comme combinaison linéaire des éléments de $\mathcal{F}$, on a alors

$$(\lambda_1 - \mu_1) v_1 + \cdots + (\lambda_k - \mu_k) v_k = v - v = 0_V$$

mais comme $\mathcal{F}$ est libre, cela implique que

$$\lambda_1 - \mu_1 = \cdots = \lambda_k - \mu_k = 0$$

et ainsi $\lambda_1 = \mu_1$, ..., $\lambda_k = \mu_k$. □

Exercice 6. *On a vu que les familles* $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$ *et* $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}$ *sont des bases de* $\mathbb{R}^2$*. Est-ce aussi le cas pour* $\left\{\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}$ *?*

De même, la famille $\left\{\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix}\right\}$ *est-elle une base de* $\mathbb{R}^3$ *?*

Exercice 7. *Montrer que le famille de vecteurs*

$$\left\{e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},e_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,e_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right\}$$

introduite dans l'exemple 8 est une base de $\mathbb{R}^n$*. On l'appelle la base* canonique *de* $\mathbb{R}^n$*.*

Le théorème suivant est très utile, mais il est loin d'être évident et nous allons donc l'admettre.

Théorème 2.21 (existence des bases)**.** *Si V est un espace vectoriel, alors V admet une base. En outre, toutes les bases de V possèdent le même cardinal (c'est-à-dire le même nombre d'éléments).*

En particulier, si V admet une base finie alors toutes ses bases sont finies. Et si V admet une base infinie alors toutes ses bases sont infinies

Définition 2.22 (Dimension d'un espace vectoriel)**.** Soit V un espace vectoriel. Le cardinal commun à toutes les bases de V est appelé la *dimension* de V, notée $\dim(V)$.

La dimension d'un espace vectoriel est soit un entier naturel $k \in \mathbb{N}$ soit ∞.

Nous n'avons pas donné d'exemple d'espace de dimension infinie jusqu'alors mais il en existe certains que l'on peut saisir sans trop de peine.

Exemple 9 (Fonctions d'une variable réelle et polynômes)**.** *L'analyse offre de nombreux exemples d'espaces de dimension infinie. Considérons I un intervalle de $\mathbb{R}$ et $\mathcal{E}$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ définies sur I.*

Pour ces fonctions nous avons une addition notée « + » : si $f,g \in \mathcal{E}$ alors $f+g$ donnée par $(f+g)(x) = f(x)+g(x)$ est encore dans $\mathcal{E}$.

Nous disposons aussi d'une multiplication par un scalaire, notée « · ». Si $\lambda \in \mathbb{R}$ et $f \in \mathcal{E}$ alors $\lambda \cdot f$ donnée par $(\lambda \cdot f)(x) = \lambda \times \big(f(x)\big)$ est encore dans $\mathcal{E}$.

On peut vérifier que toutes les propriétés de la définition 2.8 sont satisfaites. Ainsi $(\mathcal{E}, +, \cdot)$ *est un espace vectoriel. Nous ne le montrons pas ici mais c'est un espace vectoriel de dimension* infinie.

On s'intéresse plutôt à la partie $\mathcal{P}$ *de* $\mathcal{E}$ *formée des polynômes définis sur l'intervalle* I*. On a* $p \in \mathcal{P}$ *si* p *s'écrit* $p(x) = a_0 + a_1 x + \cdots + a_n x^n$*, avec* $a_0, a_1, \ldots, a_n$ *des nombres réels. On dit alors que* p *est un polynôme de degré inférieur à* n*. Ici* n *peut être aussi grand que l'on veut.*

On vérifie sans peine que la somme $p + q$ *de deux polynômes* p *et* q *est un polynôme. De même si* p *est un polynôme et* $\lambda \in \mathbb{R}$ *alors* $\lambda \cdot p$ *est un polynôme. Ainsi* $\mathcal{P}$ *est stable par addition et multiplication par un scalaire. C'est donc un sous-espace vectoriel de* $\mathcal{E}$*.*

Pour $j = 0, 1, \ldots,$ *considérons le polynôme* e_j *donné par* $e_j(x) = x^j$*. La famille* $\mathcal{F} = \{e_j ; j \in \mathbb{N}\}$ *est infinie et incluse dans* $\mathcal{P}$*. Clairement elle est génératrice.*

A titre d'exercice, montrez que c'est une famille libre.
Indication : on écrira une combinaison linéaire des e_j*,* $p = a_0 e_0 + a_1 e_1 + \cdots + a_k e_k$ *pour un certain* $k \in \mathbb{N}$*, qui donne la fonction nulle, c'est-à-dire,*

$$p(x) = a_0 + a_1 x + \cdots + a_k x^k = 0, \qquad \forall x \in I.$$

On poura utiliser le fait que seul le polynôme nul admet une infinité de racines ou bien calculer itérativement $p(0), p'(0), \ldots, p^{(k)}(0)$ *et en déduire que tous les coefficients* $a_0, a_1, \ldots, a_k$ *sont nuls.*

Ainsi $\mathcal{F} = \{e_j ; j \in \mathbb{N}\}$ *est une famille infinie libre et génératrice. L'espace vectoriel* $\mathcal{P}$ *est de dimension infinie.*

Exercice 8. *En lien avec l'exemple précédent, expliquer pourquoi l'ensemble des polynômes de degré inférieur à* n *est un espace vectoriel. De plus, expliquer pourquoi sa dimension est* $n + 1$*.*

La proposition suivante est importante. Elle permet de déduire qu'un sous-espace vectoriel est en fait égal à l'espace vectoriel complet (dans le cas de la dimension finie).

Proposition 2.23. *Si* $W \subset V$ *est un sous-espace vectoriel d'un espace vectoriel* V*, alors* $\dim(W) \leq \dim(V)$*.*

Sous ces conditions, si de plus $\dim(V) < \infty$ *on a*

$$\dim(W) = \dim(V) \quad \Rightarrow \quad W = V.$$

La proposition est souvent utile en dimension finie pour excure que certaines familles soient libres ou génératrice.

Proposition 2.24. *Soit V un espace vectoriel de dimension $n < \infty$ et $\mathcal{F}$ une famille de vecteurs.*

1. *Si $\mathcal{F}$ est une famille libre alors* $\mathrm{Cardinal}(\mathcal{F}) \leq n$.
2. *Si $\mathcal{F}$ est une famille génératrice alors* $\mathrm{Cardinal}(\mathcal{F}) \geq n$.

Enfin le théorème suivant est une conséquence du théorème 2.21. Il porte de nom de « théorème de la base incomplète ».

Théorème 2.25 (base incomplète)**.** *Si V est un espace vectoriel et $\mathcal{F}_1$ une famille libre de vecteurs de V. Alors il existe $\mathcal{F}_2$, une seconde famille libre de vecteurs, telle que, réunies, $\mathcal{F}_1$ et $\mathcal{F}_2$ forment une base de V.*

En dimension finie le résultat suivant devient très important en pratique.

Proposition 2.26. *Soit V un espace vectoriel de dimension $n < \infty$. Si $\mathcal{F}$ est une famille de n vecteurs d eV on a l'équivalence*

$$\mathcal{F} \text{ est une famille libre} \quad \Leftrightarrow \quad \mathcal{F} \text{ est une famille génératrice} \quad \Leftrightarrow \quad \mathcal{F} \text{ est une base.}$$

2.4 Matrices

2.4.1 Un nouvel exemple d'espace vectoriel

Les matrices que vous connaissez certainement très bien déjà sont d'autres exemples fondamentaux d'espaces vectoriels qui généralisent $\mathbb{R}^n$.

Définition 2.27. Soient m et n deux entiers positifs. Une matrice $m \times n$ est un tableau à m lignes et n colonnes dans lequel sont disposés mn réels $a_{11}, \ldots, a_{mn}$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}$$

Remarque 3. *Pour condenser, on notera parfois $(a_{ij})_{i\in[1,m],j\in[1,n]}$ une matrice de la forme ci-dessus.*

L'ensemble $\mathcal{M}_{mn}$ des matrices $m \times n$ est un espace vectoriel de manière très similaire aux vecteurs de $\mathbb{R}^n$ via les opérations suivantes.

1. La somme $A+B$ de deux matrices $A = (a_{ij})_{i\in[1,m],j\in[1,n]}$ et $B = (b_{ij})_{i\in[1,m],j\in[1,n]}$ est définie par
$$A + B = (a_{ij} + b_{ij})_{i\in[1,m],j\in[1,n]}$$
.

2. Le multiple $\lambda \cdot A$ d'une matrice $A = (a_{ij})_{i\in[1,m],j\in[1,n]}$ par un réel λ est donné par
$$\lambda \cdot A = (\lambda a_{ij})_{i\in[1,m],j\in[1,n]}.$$

L'espace vectoriel $\mathcal{M}_{mn}$ des matrices $m \times n$ est de dimension égale à mn.

Exercice 9. *Trouver une base de* $\mathcal{M}_{mn}$ *et justifier que* $\dim(\mathcal{M}_{mn}) = mn$.
Indication : on pourra s'inspirer de la base canonique de $\mathbb{R}^n$ *donnée dans l'exercice 7 pour penser à une base de* $\mathcal{M}_{mn}$.

Si $m = n$ on parle de matrice carrée. On note alors $\mathcal{M}_n(\mathbb{R})$ ou simplement $\mathcal{M}_n$ l'espace vectoriel des matrices carrées $n \times n$.

2.4.2 Produits de matrices

Une des spécificités des matrices est qu'elles ne sont pas *que* des espaces vectoriels. Il existe en effet dans de nombreux cas une loi supplémentaire : on peut *multiplier* les matrices entre elles.

Vous avez sûrement déjà vu ce produit, mais il est très important de vous le rappeler pour la suite du cours. Nous allons donc définir le produit AB de deux matrices A et B lorsque le nombre de *colonnes* de A est le même que le nombre de *lignes* de B. Commençons par le cas le plus simple, celui des vecteurs ligne et colonne.

Si on a $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, leur produit AB est défini par

$$AB = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n.$$

Plus généralement pour les matrices rectangulaires, on applique la même formule à chaque ligne de A et chaque colonne de B. Plus formellement, si on a $A = (a_{ij})_{i\in[1,m],j\in[1,n]}$ et $B = (b_{ij})_{i\in[1,n],j\in[1,p]}$ (on prend bien garde de ne pas oublier que le nombre de lignes de B doit correspondre au nombre de colonnes de A), on définit le produit AB par la formule

$$AB = (c_{ij})_{i\in[1,m],j\in[1,p]}, \text{ où } c_{ij} = a_{i1}b_{ij} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Bien sûr, il ne s'agit pas d'apprendre cette formule par coeur mais de la comprendre et pour cela, il faut avant tout la pratiquer. Elle signifie simplement que le réel c_{ij} qui apparait dans AB correspond au produit de la i-ème ligne de A par la j-ième ligne de B, pour le produit des vecteurs ligne et colonne défini ci-dessus.

Exercice 10. *Calculer les produits de matrices suivantes, lorsqu'ils existent.*

1. $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 1 \end{pmatrix}$;
2. $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 4 & 2 & 1 \\ 1 & 2 & 5 \end{pmatrix}$;
3. $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$;
4. $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 2 & 5 \end{pmatrix}$;

2.4.3 Matrices particulières

Nous allons utiliser à de nombreuses reprises des matrices particulières dans la suite de ce cours. On ne considérera ici que les matrices carrées, mais ces définitions sont tout à fait légitimes, même pour les matrices rectangulaires.

Définition 2.28. Soit $A = (a_{ij})_{i\in[1,n],j\in[1,n]}$ une matrice de $\mathcal{M}_n(\mathbb{R})$.

1. On dit que A est triangulaire supérieure, si $a_{ij} = 0$ pour tout $i > j$. Une telle matrice est alors de la forme
$$A = \begin{pmatrix} * & \cdots & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}$$
où $*$ désigne une valeur quelconque.
2. Si on a $a_{ij} = 0$ pour tout $j > i$, on dit que A est triangulaire inférieure. Une telle matrice est de la forme
$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ * & \cdots & \cdots & * \end{pmatrix}$$
où $*$ désigne une valeur quelconque.

3. Si $a_{ij} = 0$ lorsque $i \neq j$, A est diagonale et donc de la forme

$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & * \end{pmatrix}$$

   où $*$ désigne une valeur quelconque.

Exemple 10. *Un exemple essentiel de matrice diagonale est celui des matrices identité : pour tout $n \geq 1$, on note I_n la matrice diagonale, pour laquelle tous les coefficients non-nuls sont égaux à 1. Vous pouvez vérifier que cette matrice est très particulière : si $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice, alors on a $AI_n = I_nA = A$. Cette propriété peut même servir de définition équivalente à la matrice I_n.*

La classe de matrice que nous définissons désormais est essentielle, nous y reviendrons très vite dans le cours.

Définition 2.29. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite inversible, s'il existe une autre matrice $B \in \mathcal{M}_n(\mathbb{R})$, telle que $AB = BA = I_n$.

Exemple 11. *Bien sûr, la matrice I_n est elle-même inversible. Toutes les matrices ne sont cependant pas inversibles : par exemple, la matrice nulle (dont tous les coefficients sont égaux à 0) ne peut pas être inversible, puisque son produit avec toute matrice ne peut être que la matrice nulle.*

Chapitre 3

Applications linéaires

Définition 3.1. Soient A, B deux ensembles et $f : A \longrightarrow B$ une application.

1. On dit que f est injective si pour tout couple $(x, y) \in A^2$, on a

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

2. On dit que f est surjective si tout élément y de B, il existe un élément $x \in A$ tel que $f(x) = y$.
3. Si f est à la fois injective et surjective, on dit que f est bijective.

Remarque 1. *On peut montrer qu'une application $f : A \longrightarrow B$ est bijective, si et seulement s'il existe une autre application $f^{-1} : B \longrightarrow A$ tel que l'on ait*

$$f \circ f^{-1} = id_B \text{ et } f^{-1} \circ f = id_A$$

(id_A et id_B désignent ici les applications identités de A et B).

Nous avons défini précédemment la notion d'espace vectoriel, il est désormais temps de nous intéresser aux applications qui respectent la structure d'espace vectoriel, c'est à dire celles qui se comportent convenablement vis-à-vis des additions et des multiplications par un scalaire.

3.1 Définition, exemples et premières propriétés

Définition 3.2 (application linéaire)**.** Soient V et W deux espaces vectoriels et une application $f : V \to W$. On dit que f est linéaire si pour tout $v, v' \in V$ et $\lambda \in \mathbb{R}$ on a

1. $f(v + v') = f(v) + f(v')$;
2. $f(\lambda v) = \lambda f(v)$.

Quand $W = \mathbb{R}$ on parle de forme linéaire.

Remarque 2. *Si on veut être très rigoureux on doit plutôt écrire* $(V, +_V, \cdot_V)$ *et* $(W, +_W, \cdot_W)$ *pour désigner les deux espaces vectoriels. Alors* $f : V \to W$ *est linéaire si on a*

1. $f(v +_V v') = f(v) +_W f(v')$.
2. $f(\lambda \cdot_V v) = \lambda \cdot_W f(v)$.

pour tous $v, v' \in V$ *et* $\lambda \in \mathbb{R}$.

Dans la suite nous ne ferons pas toujours un tel effort de rigueur.

Exemple 1. $f : \mathbb{R}^n \to \mathbb{R}^p$ *donnée par* $f(x_1, \ldots, x_n) = (0, \ldots, 0) = 0_{\mathbb{R}^p}$. *On dit que* f *est l'application nulle de* $\mathbb{R}^n$ *dans* $\mathbb{R}^p$.

Exemple 2. *La forme linéaire* $f : \mathbb{R}^2 \to \mathbb{R}$ *donnée par* $f(x_1, x_2) = x_1 + x_2$.

Exemple 3.

1. *L'application*

$$f : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array}$$

n'est pas linéaire.

2. *L'application*

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & 2x - y \end{array}$$

est linéaire.

3. *L'application*

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & xy \end{array}$$

n'est pas linéaire.

4. *L'application*

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x, y) & \longmapsto & (3x + y, -x + 2y) \end{array}$$

est linéaire.

5. *On note* A *la matrice* $n \times n$ *dont tous les coefficients sont nuls, sauf le coefficient* a_{11} *qui vaut* 1. *L'application*

$$f : \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ B & \longmapsto & AB \end{array}$$

est linéaire.

Exemple 4. *Sur* $V = \mathscr{C}^0(\mathbb{R})$ *l'application* $f : V \ni g \mapsto g(0)$ *est une forme linéaire.*

Exercice 1. *Si* $f : V \longrightarrow W$ *est une application linéaire, montrer que* $f(0_V) = 0_W$*. En déduire que l'application*

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & 2x - y + 1 \end{array}$$

n'est pas linéaire.

De la définition 3.2 il découle que si $u_1, \ldots, u_k$ sont des vecteurs de V et $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ alors

$$f(\lambda_1 u_1 + \cdots + \lambda_k u_k) = \lambda_1 f(u_1) + \cdots + \lambda_k f(u_k). \tag{3.1.1}$$

L'image d'une combinaison linéaire est donc la combinaison linéaire des images.

3.2 Caractérisation au moyen d'un base et représentation matricielle

Proposition 3.3. *Soit* V *un espace vectoriel et* $\mathcal{V}$ *une base de* V*. Une application linéaire* f *de* V *dans* W *un second espace vectoriel est caractérisée par les images des éléments de la base* $\mathcal{V}$*.*

Autrement dit il suffit uniquement de connaître les images des élément d'une base par une application linéaire pour complètement connaître l'application linéaire.

Démonstration. Supposons les images des vecteurs de la base $\mathcal{V}$ connus. Soit $x \in V$. Alors, il existe $v_1, \ldots, v_k$ des vecteurs de $\mathcal{V}$ et $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tels que $x = \lambda_1 v_1 + \cdots + \lambda_k v_k$. Ainsi par (3.1.1) on a nécessairement

$$f(x) = \lambda_1 f(v_1) + \cdots + \lambda_k f(v_k),$$

et les $f(v_1), \ldots, f(v_k)$ sont connus. □

Focalisons maintenant la discussion sur le cas de la dimension finie. Ainsi supposons que V et W sont de dimensions finies n et p respectivement. Soient aussi $\mathcal{V} = (v_1, \ldots, v_n)$ une base de V et $\mathcal{W} = (w_1, \ldots, w_p)$ une base de W.

On a $f(v_1) \in W$ qui se décompose donc suivant la base $\mathcal{W}$. On note a_{i1}, avec $i = 1, \ldots, p$ ses coordonnées :

$$f(v_1) = a_{11} w_1 + \cdots + a_{p1} w_p.$$

On fait de même avec tous les vecteurs de la base $\mathcal{V}$:

$$f(v_j) = a_{1j}w_1 + \cdots + a_{pj}w_p.$$

On définit ainsi une matrice A à p lignes et n colonnes. La j-ème colonne de cette matrice donne les coordonnées de $f(v_j)$ dans la base $\mathcal{W}$.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pn} \end{pmatrix}. \tag{3.2.1}$$

Si $x \in V$ et $x_1, \ldots, x_n$ sont ses coordonnées dans $\mathcal{V}$ alors $x = x_1v_1 + \cdots + x_nv_n$. On forme X le vecteur colonne

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{3.2.2}$$

On calcule

$$\begin{aligned} f(x) &= x_1f(v_1) + \cdots + x_nf(v_n) \\ &= \big(a_{11}x_1 + \cdots + a_{1n}x_n\big)w_1 + \cdots + \big(a_{p1}x_1 + \cdots + a_{pn}x_n\big)w_p \end{aligned}$$

Ainsi la coordonnée de $f(x)$ selon le vecteur w_i de la base $\mathcal{W}$ est donnée par

$$a_{i1}x_1 + \cdots + a_{in}x_n.$$

Par la définition du produit de deux matrices c'est précisement la i-ème coordonnée du vecteur colonne AX qui est un vecteur à p composantes. Ainsi le vecteur colonne AX donne les composantes de $f(x)$ suivant la base $\mathcal{W}$.

On conclut que l'existence des bases et la caractérisation d'une application linéaire par les images des vecteurs d'une base permet de complètement représenter une application linéaire et son action sur un vecteur au moyen d'une matrice et du produit matriciel.

Remarque 3. *La notion de matrice associée à une application linéaire est d'autant plus utile et naturelle qu'elle respecte la composition des applications.*

Soient E, F et G trois espaces vectoriels dont on fixe trois bases respectives $\mathcal{E}$, $\mathcal{F}$ et $\mathcal{G}$. On considère deux applications linéaires $f : E \longrightarrow F$, $g : F \longrightarrow G$ et on note A la matrice de f dans les bases $\mathcal{E}$ et $\mathcal{F}$ (resp. B la matrice de g dans les bases $\mathcal{F}$ et $\mathcal{G}$).

Alors la matrice de l'application $g \circ f : E \longrightarrow G$ dans les bases $\mathcal{E}$ et $\mathcal{G}$ est donnée par le produit BA.

3.3 Noyau et image

Définition 3.4. Soit $f : V \longrightarrow W$ une application linéaire. Le noyau de f noté $\ker(f)$ est défini par
$$\ker(f) = \{v \in V,\ f(v) = 0_W\}.$$
L'image de f, notée $Im(f)$, est quant à elle définie par
$$Im(f) = \{f(v),\ v \in V\}.$$

Le petit résultat suivant le montre, Les noyaux et images d'applications linéaires sont des espaces vectoriels. Nous verrons dans la suite que ces exemples sont fondamentaux et très utiles.

Proposition 3.5. *Si $f : V \longrightarrow W$ est une application linéaire, alors $\ker(f)$ est un sous-espace vectoriel de V. De même, l'image de f est un sous-espace vectoriel de W.*

Démonstration.

1. On commence par montrer le résultat pour le noyau $\ker(f)$, qui n'est pas vide puisqu'il contient 0_V. Si v et v', sont deux éléments de $\ker(f)$, on a bien
$$f(v + v') = f(v) + f(v') = 0_W + 0_W = 0_W$$
car f est une application linéaire, et de même, si $\lambda \in \mathbb{R}$, on a
$$f(\lambda v) = \lambda \cdot f(v) = \lambda \cdot 0_W = 0_W.$$
2. Le sous-ensemble $Im(f)$ n'est évidemment pas vide, reste à montrer que lui aussi est stable par addition et multiplication par un scalaire. Si $w = f(v)$ et $w' = f(v')$ sont deux éléments de $Im(f)$, on a
$$w + w' = f(v) + f(v') = f(v + v') \in Im(f).$$
De même si λ est un réel,
$$\lambda \cdot w = \lambda \cdot f(v) = f(\lambda \cdot v) \in Im(f).$$

□

Par définition, une application $f : V \longrightarrow W$ est surjective, si et seulement si $Im(f) = W$. Si f est une application linéaire, son noyau permet aussi de vérifier son injectivité.

Proposition 3.6. *Soit $f : V \longrightarrow W$ une application linéaire. L'application f est injective si et seulement si son noyau $\ker(f)$ est réduit à 0_V.*

Démonstration. On a vu précédemment que $f(0_V) = 0_W$, on a donc toujours $0_V \in \ker(f)$. Supposons d'abord que f est injective et montrons que $\ker(f) = \{0_V\}$. Soit $x \in \ker(f)$. On a $f(x) = 0_W = f(0_V)$ donc par injctivité de f, $x = 0_V$ et on a bien $\ker(f) = \{0_V\}$.
Réciproquement, supposons que $\ker(f) = \{0_V\}$ et montrons que f est injective. Soient $(v, v') \in V^2$ tels que $f(v) = f(v')$. Comme f est linéaire, on a

$$f(v - v') = f(v) - f(v') = 0_V,$$

on a donc $v - v' \in \ker(f) = \{0_V\}$ donc $v - v' = 0_V$ et $v = v'$. □

Déterminer le noyau ou l'image d'une application linéaire est donc essentiel et une partie important de la suite de ce cours consiste à élaborer des techniques pour ce faire.

Focalisons maintenant la discussion sur le cas de la dimension finie. Ainsi supposons que V et W sont de dimensions finies n et p respectivement. On choisit des bases $\mathcal{V} = (v_1, \dots, v_n)$ et $\mathcal{W} = (w_1, \dots, w_p)$ de V et W respectivement. Alors si $f : V \longrightarrow W$ est une application linéaire on peut lui associer une matrice $A \in \mathcal{M}_{pn}$. De plus si x a pour coordonnées le vecteur colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{V}$ alors $f(x)$ admet AX comme vecteur colonne de coordonnées dans la base $\mathcal{W}$. Ainsi, si $x \in \ker(f)$ alors $AX = 0$. Cela s'écrit

$$\begin{cases} a_{11}x_1 & + & \cdots & + a_{1n}x_n & = 0 \\ \quad\vdots & & & \quad\vdots & \\ a_{p1}x_1 & + & \cdots & + a_{pn}x_n & = 0 \end{cases}$$

Ainsi la détermination d'un noyau d'application linéaire correspond à la résolution d'un système d'équations linéaires homogènes.

3.4 Le théorème du rang

Le théorème du rang est l'énoncé suivant qui est essentiel, très spécifique à l'algèbre linéaire et valable uniquement pour les espaces vectoriels de dimension finie.

Théorème 3.7. *On considère $f : V \longrightarrow W$ une application linéaire, où V et W sont des espaces vectoriels de dimension finie. Alors on a l'égalité*

$$\dim(V) = \dim(\ker(f)) + \dim(Im(f))$$

Démonstration. On note $k = \dim(\ker(f))$ et $n = \dim(V)$. Dans un premier temps fixons une base $\mathcal{B} = \{e_1, ..., e_k\}$ de $\ker(f)$ grâce au théorème 2.21. Dans un second temps, on la complète avec des vecteurs $\{e_{k+1}, ..., e_n\}$ pour obtenir une base

$$\mathcal{C} = \{e_1, ..., e_k, e_{k+1}, ..., e_n\}$$

de V grâce au théorème de la base incomplète (théorème 2.25).

On montre qu'avec ces données, la famille $\mathcal{I} = \{f(e_{k+1}), ..., f(e_n)\}$ est une base de $Im(f)$:

1. Elle engendre en effet $Im(f)$: si $x = f(y) \in Im(f)$, alors en décomposant y dans la base $\mathcal{C}$ on obtient

$$\begin{aligned} x = f(y) = & \ f(y_1e_1 + ... + y_ne_n) \\ = & \ y_1f(e_1) + ... + y_kf(e_k) + y_{k+1}f(e_{k+1}) + ... + y_nf(e_n) \\ = & \ y_{k+1}f(e_{k+1}) + ... + y_nf(e_n) \end{aligned}$$

   Car les vecteurs e_1,..., e_k appartiennent au noyau de f, donc $x \in \mathrm{Vect}(\mathcal{I})$.

2. On montre que la famille $\mathcal{I}$ est libre : soit $\lambda_{k+1}, ..., \lambda_n$ est réels tels que

$$\lambda_{k+1}f(e_{k+1}) + ... + \lambda_nf(e_n) = 0_W.$$

   En utilisant à nouveau la linéarité de f, cette inégalité implique que

$$f(\lambda_{k+1}e_{k+1} + ... + \lambda_ne_n) = 0_W,$$

   en particulier l'élément $\lambda_{k+1}e_{k+1} + ... + \lambda_ne_n$ appartient au noyau de f, mais $\mathcal{B}$ en est une base par définition, on a donc

$$\lambda_{k+1}e_{k+1} + ... + \lambda_ne_n = \mu_1e_1 + ... + \mu_ke_k$$

   pour certains réels μ_1,...,μ_k, ce qui s'écrit encore

$$-\mu_1e_1 - \cdots - \mu_ke_k + \lambda_{k+1}e_{k+1} + \cdots + \lambda_ne_n = 0_V.$$

   La famille $\mathcal{C}$ étant une base de V, les coefficients μ_1,..., μ_k, λ_{k+1},..., λ_n sont tous nuls.

On a donc montré que $\mathcal{I}$ est une base de $Im(f)$, ce qui donne directement que $\dim(Im(f) = n - k$, et il devient immédiat que

$$\dim(\ker(f)) + \dim(Im(f) = k + (n - k) = n = \dim(V).$$

□

3.5 Opérations de Gauss sur les colonnes d'une matrice

3.5.1 Opérations élémentaires

Tout l'attirail technique mis en place avec le calcul matriciel va nous être très utile pour développer divers algorithmes à partir des opérations de Gauss. Il s'agit ici de mimer l'algorithme de Gauss pour les systèmes linéaires présenté précédemment dans le cas des matrices, mais il va vous falloir toutefois faire un petit effort : plutôt que de réduire suivant les lignes, nous allons cette fois considérer les colonnes.

Les raisons qui nous font privilégier les colonnes sont trop complexes pour être démontrées dans ce cours. Pour résumer, la réduction sur les colonnes nous permettront d'obtenir, étant donné une application linéaire f, des bases pour l'image de f ainsi que pour son noyau en un seul algorithme, ce qui n'est pas le cas pour la réduction sur les lignes. Il s'agit donc désormais la formulation matricielle de la réduction de Gauss.

Dans toute cette partie, fixe une matrice $A = (a_{ij})_{i\in[1,m],j\in[1,n]}$.

Définition 3.8. On note C_1,...,C_n les colonnes de la matrice A. Les opérations de Gauss sur les colonnes de A sont définies de manière similaire aux opérations sur le slignes des systèmes linéaires :

1. Échange de deux lignes C_i et C_j, noté $C_i \leftrightarrow C_j$

2. Remplacement d'une ligne C_i par elle-même, multipliée par un nombre réel λ non-nul, noté
$$C_i \leftarrow \lambda C_i$$

3. Remplacement d'une ligne (C_i) par elle-même, à laquelle on ajoute un multiple $\lambda(C_j)$ d'une *autre* colonne C_j, noté
$$C_i \leftarrow C_i + \lambda C_j$$

La remarque fondamentale est que ces opérations ont une interprétation matricielle d'une efficacité redoutable. Pour ces exemples, notons $\{e_1, ..., e_n\}$ la base canonique de $\mathbb{R}^n$.

1. Si $i \neq j$, on note $f_{ij} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ définie par
$$f_{ij}(e_k) = \begin{cases} e_j & \text{si } k = i \\ e_i & \text{si } k = j \\ e_k & \text{sinon.} \end{cases}$$

La matrice de l'application f_{ij} dans la bases canonique est très simple, il s'agit de la matrice identité de taille n, où on inverse les colonnes i et j.

Exercice 2. *(a) Vérifier cela en exprimant à la main la matrice P_{23} de l'application $f_{23} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.*

(b) Considérons par exemple la matrice $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Calculer BP_{23}. Que s'est-il passé sur les colonnes de B après cette multiplication ?

Nous espérons avec l'exercice précédent vous en convaincre : la première opération de Gauss sur les colonnes $C_i \leftrightarrow C_j$ correspond simplement à la multiplication à droite par la matrice P_{ij}. Montrons que les autres opérations de Gauss sur les colonnes sont tout aussi sympathiques.

2. Si $\lambda \in \mathbb{R}^*$ est un réel non-nul, considérons l'application $p_{i,\lambda} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ définie par

$$p_{i,\lambda}(e_k) = \begin{cases} \lambda e_i & \text{si } k = i \\ e_k & \text{sinon.} \end{cases}$$

La matrice de $p_{i,\lambda}$ dans la base canonique de $\mathbb{R}^n$ est très simple, il s'agit de la matrice identité pour laquelle le "1" de la i-ème colonne est remplacé par "λ".

Exercice 3. *(a) Vérifier cela en exprimant à la main la matrice C de l'application $p_{2,\lambda} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.*

(b) Considérons par exemple la matrice $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Calculer BC. Que s'est-il passé sur les colonnes de B après cette multiplication ?

L'histoire se répète : vous venez de vérifier que l'opération $C_i \leftarrow \lambda C_i$ est obtenue matriciellement en multipliant à droite par la matrice associée à $p_{i,\lambda}$.

3. Pour l'opération $C_i \leftarrow C_i + \lambda C_j$, étant données deux entiers $i \neq j$ et un réel λ, on considère l'application $g_{i,j,\lambda} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ définie par

$$g_{i,j,\lambda}(e_k) = \begin{cases} e_i + \lambda e_j & \text{si } k = i \\ e_k & \text{sinon.} \end{cases}$$

Sa matrice n'est pas plus compliquée que les précédentes, il s'agit de la matrice identité, pour laquelle on remplace un des "0" par λ".

Exercice 4. *(a) Vérifier cela en exprimant à la main la matrice D de l'application $p_{2,3,\lambda} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.*

(b) *Considérons par exemple la matrice* $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. *Calculer* BD. *Que s'est-il passé sur les colonnes de* B *après cette multiplication ?*

Vous l'aurez compris, la multiplication à droite par la matrice associée à $g_{i,j,\lambda}$ est l'interprétation matricielle de l'opération de Gauss $C_i \leftarrow C_i + \lambda C_j$.

Ces interprétations matricielles sont particulièrement performantes, et en mimant l'algorithme de réduction sur les lignes vu pour les systèmes linéaires, on obtient le résultat suivant.

Définition 3.9. Soient A et B deux matrices de même taille. On dit que A et B sont équivalentes, si l'on peut obtenir B en effectuant des opérations de Gauss sur les colonnes de A.

Théorème 3.10. *Si* $A = (a_{ij})_{i \in [1,m], j \in [1,n]}$ *est une matrice, alors* A *est équivalente à une matrice échelonnée de la forme*

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & \cdots & 0 \\ * & * & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & 0 & 0 & \cdots & 0 \\ & & & * & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ * & \cdots & & * & 0 & \cdots & 0 \end{pmatrix}$$

Exprimé dans le langage matriciel, toute matrice A *devient échelonnée et triangulaire inférieure après multiplication à droite par une matrice inversible.*

Démonstration. Il s'agit de l'algorithme explicité précédemment pour les systèmes linéaires, à la différence près que cette fois, au lieu d'annuler les termes $a_{2,1}, \ldots, a_{n,1}$ en faisant des opérations sur les lignes, on annule les termes $a_{1,2}, \ldots, a_{1,n}$ en faisant les opérations sur les colonnes.

Comme vu précédemment, les opérations de Gauss sont toutes réalisées en multipliant la matrice par une matrice inversible. Les opérations élémentaire pour parvenir à une matrice triangulaire et échelonnées peuvent être nombreuses, mais les effectuer toutes correspond toujours par la multiplication par une matrice inversible, donnée par le produit des matrices associées à chacune des opérations élémentaires. □

Exemple 5. *On considère la matrice* $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}$.

1. *Effectuer la réduction de Gauss sur les colonne de* A.
2. *Construire une matrice* P *inversible, telle que* AP *soit triangulaire supérieure.*

3.5.2 Noyaux et images d'une application linéaire

La réduction de Gauss sur les colonnes d'une matrice est un outil très utile pour déterminer de bases au noyau et à l'image d'une application linéaire, comme le montre le résultat suivant.

Théorème 3.11. *Soit $A = (a_{ij})_{i\in[1,m],j\in[1,n]}$ la matrice associée à une application linéaire $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. On note $A' = AP$ une matrice échelonnée et triangulaire inférieure obtenue grâce à la réduction de Gauss sur les colonnes de A. En notant*

$$f' : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

l'application linéaire telle que $Mat(f')_{\mathcal{B}} = A'$ on a alors

1. *$Im(f) = Im(f')$.*
2. *si C_k, C_{k+1},..., C_n sont les colonnes nulles de A', alors*

$$\ker(f) = \text{Vect}\{P_k, ..., P_n\}$$

où P_k,...,P_n désignent les $n - k + 1$ dernières colonnes de P.

Démonstration. 1. On sait que $Im(f)$ est l'espace vectoriel engendré par les colonnes de A et que $Im(f')$ est celui engendré par les colonnes de A'.

Les colonnes de A' sont obtenues en effectuant des opérations de Gauss sur celles de A. Les opérations de Gauss ne modifient pas l'espace vectoriel engendré par une famille, on a donc bien $Im(f) = Im(f')$.

2. On a
$$v \in \ker(f') \Leftrightarrow A'(v) = 0 \Leftrightarrow AP(v) = 0 \Leftrightarrow P(v) \in \ker(f).$$

Comme A' est triangulaire et échelonnée, son noyau est très simple : une base en est donnée par Vect$\{e_k, ..., e_k\}$, si les colonnes nulles de A' sont C_k,...,C_n.

Dès lors, la famille $\{P(e_k), ..., P(e_n)\}$ constituée des $n - k + 1$ dernières colonnes de P engendre $\ker(f)$ par la série d'équivalence ci-dessus, et c'en est même une base car P est inversible.

□

Corollaire 3.12. *En gardant les notations du théorème 3.11, Les colonnes C_1,...,C_{k-1} de A' donnent une base de $Im(f)$.*

Démonstration. On a vu que $Im(f) = Im(f')$ et que $\{C_1, ..., C_{k-1}\}$ engendre $Im(f')$, il s'agit donc de montrer que cette famille est libre, ce qui est une conséquence du fait que la matrice A' est échelonnée.

Soient $\lambda_1, \ldots, \lambda_{k-1}$ des réels tels que

$$\lambda_1 C_1 + \cdots + \lambda_{k-1} C_{k-1} = 0.$$

Considérons la première coordonnée non-nulle de C_1. Cette coordonnée est nulle pour les autres vecteurs, l'égalité ci-dessus implique donc que $\lambda_1 = 0$. On procède alors de la même manière avec les autres vecteurs, à la manière d'un système linéaire triangulaire, pour montrer en cascade que λ_2 est nul, puis que λ_3 est nul, etc, jusqu'à aboutir à $\lambda_1 = \lambda_2 = \cdots = \lambda_{k-1} = 0$. □

Exercice 5. *Le théorème précédent fournit ainsi un algorithme très efficace pour déterminer des bases à l'image et au noyau d'une application linéaire.*

1. *Soit $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ l'application linéaire définie par*

   $$f(x, y, z, t) = (x + 2y + 3z + 4t, x + 3y + 6z + 6t, 3x + 8y + 15z + 16t).$$

   On pose la matrice A associée à f dans les bases canoniques, à coté de laquelle on place une matrice identité dont le nombre de colonnes est identique à celui de A :

   $$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 1 & 3 & 6 & 6 & 0 & 1 & 0 & 0 \\ 3 & 8 & 15 & 16 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right)$$

2. *Appliquer l'algorithme de Gauss pour effectuer la réduction de la matrice de gauche, tout en faisant au fur et à mesure les mêmes opérations pour la matrice de droite.*

   Lire sur les matrices obtenues en fin d'algorithme une base de $\ker(f)$ *et une base de $Im(f)$.*

3.5.3 Rang d'une famille de vecteurs

L'algorithme donné plus haut donnant une base de l'image $Im(f)$ d'une application linéaire f, mais est tout aussi utile pour construire pour n'importe quelle famille de vecteurs $\mathcal{F}$ une base de $\mathrm{Vect}(\mathcal{F})$, ce qui est par exemple utile pour déterminer sa dimension (qu'on appelle aussi le *rang* de $\mathcal{F}$).

En effet si $\mathcal{F} = \{v_1, \ldots, v_k\}$, pour $j = 1, \ldots, k$ on considère la matrice

$$A = \left(\begin{array}{c|c|c} V_1 & \cdots & V_k \end{array} \right)$$

dont la j-ième colonne V_j est formée des coordonnées de v_j dans une base choisie. Le théorème 3.11 se traduit alors de la manière suivante.

Proposition 3.13. *On note A' la matrice obtenue à partir de*

$$A = \left(V_1 \mid \cdots \mid V_k \right)$$

après réduction de Gauss sur les colonnes du théorème 3.11.

Si C_1,...,C_k sont les colonnes non-nulles de A', alors $\{C_1, ..., C_k\}$ est une base de $\mathrm{Vect}(\mathcal{F})$, qui est donc de dimension k.

3.6 Déterminants et matrices inversibles

Dans la suite de ce cours, les matrices seront généralement toutes carrées. Vous avez étudié en première année le *déterminant* d'une matrice carrée, qui en est un invariant très importants, voici quelques rappels à ce sujet.

Le déterminant d'une matrice est un *nombre* qu'on lui associe. Il est le plus souvent défini pour toutes les matrices par récurrence sur la taille, on commence donc par les "petites" matrices.

Définition 3.14. Soit $A = (a_{11})$ une matrice 1×1. On définit le *déterminant* de de la matrice A par $\det(A) = a_{11}$.

Si A est une matrice 2×2, le déterminant de A est donné par

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

Exemple 6. *Soit*

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

alors $\det A = 1 \times 4 - 2 \times 3 = 4 - 6 = -2$.

Plus généralement, on définit le déterminant des matrices $n \times n$ à partir du déterminant des matrices $(n-1)\times(n-1)$ (avec $n \geq 2$) grâce à la formule compliquée suivante.

Définition 3.15. Soit A une matrice de dimension $n \times n$ et i un nombre entre 1 et n. On appelle déterminant de A la quantité :

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}.$$

où Δ_{ij} est le déterminant de la matrice déduite de A en supprimant la i-ième ligne et la j-ième colonne de A.

Remarque 4. 1. *Il faut fixer la valeur de i au début du calcul et ne plus en changer, on parle alors de* développement du déterminant par rapport à la i-ème ligne.. *Bien sûr, que vous décidiez de développer le déterminant par une ligne ou une autre, le résultat sera le même.*

2. *Dans la définition on a fait le choix de développer le caclul du déterminant par rapport à une ligne. Cela peut aussi se faire par rapport à une colonne : on choisit j un nombre entre 1 et n et on démontre que l'on a aussi*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}.$$

3. *On note indifféremment le déterminant d'une matrice A par* $\det(A)$*, ou par la matrice A en remplaçant ses parenthèses par des traits verticaux.*

Exercice 6. *On fixe une matrice de taille* 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

dont on va calculer le déterminant en le développant par rapport à la première ligne.

Si $1 \leq i, j \leq 3$*, on note A_{ij} la matrice* 2×2 *obtenue à partir de A en lui ôtant la ligne i et la colonne j.*

1. *Déterminer les matrices A_{11}, A_{12} et A_{13} ainsi que leurs déterminants respectifs Δ_{11}, Δ_{12} et Δ_{13}.*

2. *À l'aide de la matrice*

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix},$$

utiliser la formule de la définition précédente pour montrer que

$$\det(A) = a_{11}\Delta_{11} - a_{12}\Delta_{12} + a_{13}\Delta_{13}.$$

3. *En déduire que*

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Bien sûr, il existe des formules générales qui expriment le déterminant d'une matrice de taille quelconque, mais il vaut mieux pour notre santé mentale ne pas nous risquer à les écrire (elles sont trop gourmandes en calculs, même pour un ordinateur).

3.6.1 Premières propriétés du déterminant

Nous avons vu précédemment que l'on ne modifie pas le rang d'une famille de vecteurs en faisant des opérations de Gauss sur les colonnes qu'ils forment. Ces mêmes opérations de Gauss modifient le déterminant, mais de manière très contrôlée.

Propriété 3.16. *Par convention, on pose* $\det(I_n) = 1$*.*

On fixe A une matrice carrée de taille n.

1. *Si A' est la matrice obtenue à partir de A en faisant l'opération de Gauss $C_i \leftrightarrow C_j$, alors*
$$\det(A') = -\det(A).$$
2. *Si A' est la matrice obtenue à partir de A en faisant l'opération de Gauss $C_i \leftarrow \lambda C_i$ (le réel λ peut être éventuellement nul ici) alors*
$$\det(A') = \lambda \det(A).$$
3. *Si A' est la matrice obtenue à partir de A en faisant l'opération de Gauss $C_i \leftarrow C_i + \lambda C_j$, alors*
$$\det(A) = \det(A).$$
4. *On a toujours* $\det({}^tA) = \det(A)$*, en particulier les résultats ci-dessus sont parfaitement similaires en considérant les opérations de Gauss sur les lignes à la place de celles sur les colonnes.*
5. *Le déterminant se comporte très bien avec le produit matriciel : si A et B sont deux matrices carrées de même taille, alors*
$$\det(AB) = \det(A)\det(B).$$

Corollaire 3.17. 1. *Si A est une matrice carrée qui contient une ligne ou une colonne de* 0*, alors* $\det(A) = 0$*.*

2. *Si A est une matrice carrée de taille n et si λ est un réel, alors*
$$\det(\lambda A) = \lambda^n \det(A).$$

Exercice 7. *Montrer par récurrence sur la taille des matrices que le déterminant d'une matrice diagonale est le produit de ses termes diagonaux. Remarquer que le résultat est le même pour les matrices triangulaires inférieures et supérieures.*

3.6.2 Matrices inversibles

Nous avons vu précédemment qu'une matrice carrée $A \in \mathcal{M}_n$ est inversible, s'il existe une autre matrice B de même taille, telle que $AB = BA = I_n$. Dans le langage des applications linéaires, les matrices inversibles correspondent aux *applications bijectives*, ces matrices sont donc très importantes et particulières. Nous avons désormais les moyens d'identifier assez facilement les matrices inversibles, et même de calculer leur inverses.

Remarque 5. *On peut montrer qu'étant donnée une matrice carrée A, s'il existe une matrice carrée B telle que $AB = I_n$, alors on a automatiquement que $BA = I_n$, autrement dit A est inversible. Il s'agit d'une conséquence du théorème du rang.*

Théorème 3.18. *Soit A une matrice carrée $n \times n$. Alors A est inversible* si et seulement si $\det(A) \neq 0$.

Ce critère très utile se traduit alors de la manière suivante dans le langage des applications linéaires et celui des espaces vectoriels engendrés.

Corollaire 3.19. *Une application linéaire $f : V \to V$ est bijective si et seulement si pour toute base $\mathcal{V}$ de V, la matrice de f dans la base $\mathcal{V}$ est de déterminant non-nul.*

Corollaire 3.20. *Soient V un espace vectoriel de dimension n et $\{u_1, \ldots, u_n\}$ une famille de n vecteurs de V. Soit A la matrice $n \times n$ formée des vecteurs colonnes des coordonnées de $u_1, \ldots, u_n$ dans une base de V. Alors le rang de cette famille de vecteurs est égal à n* si et seulement si $\det(A) \neq 0$.

Étant donnée une matrice carrée A, déterminer si A est inversible ou non revient ainsi simplement à calculer son déterminant. Notons qu'une fois assurés de l'inversibilité de A, l'algorithme de réduction de Gauss sur les colonnes de A que nous avons vu au théorème 3.11 permet même de déterminer l'inverse de A, comme le montre le théorème suivant.

Théorème 3.21. *Si $A = (a_{ij})_{i \in [1,n], j \in [1,n]}$ est une matrice carrée de taille n inversible, alors A est équivalente à la matrice I_n. En particulier, poursuivre l'algorithme de l'exercice 5 permet de déterminer l'inverse de A.*

Démonstration. Par le théorème 3.11, on sait que A est équivalente à une matrice triangulaire inférieure B dont les termes diagonaux sont non-nuls. En effet si B possédait un terme diagonal nul, son déterminant serait nul, mais les relations de la propriété 3.16 impliqueraient que $\det(A) = 0$, ce qui contredirait l'inversibilité de A.

Puisque les termes diagonaux de B sont non-nuls, nous avons loisir de les supposer égaux à 1 divisant les colonnes par ceux-ci même. La matrice A est donc équivalente à une matrice triangulaire inférieure avec des 1 sur la diagonales

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

Montrer qu'une telle matrice est équivalente à l'identité est facile en effectuant des opérations de Gauss sur ses colonnes, grâce aux termes 1 de la diagonale qui permettent de rendre tous les termes $*$ égaux à 0.

Nous avons donc montré qu'une matrice inversible de taille n est équivalente à I_n, c'est à dire qu'il existe une matrice inversible P telle que $AP = I_n$. La matrice P est alors l'inverse de A et il s'agit simplement de la matrice obtenue à droite, après avoir continué l'algorithme de l'exercice 5 jusqu'à transformer A en la matrice identité. □

Exercice 8. *On considère la matrice*

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

1. *Calculer le déterminant de A. En déduire que A est inversible.*
2. *Utiliser l'algorithme du théorème 3.21 pour déterminer l'inverse de A.*

3.7 Diagonalisation des endomorphismes

3.7.1 Changements de base et effet sur la matrice d'une application linéaire

Dans V espace vectoriel de dimension n ($V = \mathbb{R}^n$ par exemple) si on se donne une base $\mathcal{E} = \{e_1, \ldots, e_n\}$ alors tout vecteur v de V s'écrit de manière unique $v = x_1e_1 + \cdots + x_ne_n$. On rappelle que $x_1, \ldots, x_n$ sont les coordonnées de v dans la base $\mathcal{E}$. Si on se donne une seconde base $\mathcal{F} = \{f_1, \ldots, f_n\}$ alors ce même vecteur v s'écrit de manière unique $v = x'_1f_1 + \cdots + x'_nf_n$. Les coordonnées $x_1, \ldots, x_n$ et les coordonnées $x'_1, \ldots, x'_n$ sont bien sûr différentes mais quel lien y a-t-il entre elles ?

On note U et U' les vecteurs colonnes contenant les coordonnées de U dans les bases $\mathcal{E}$ et $\mathcal{F}$:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

On dira que $\mathcal{E}$ est l'ancienne base et $\mathcal{F}$ est la nouvelle base.

Pour $j = 1, \ldots, n$ le vecteur f_j s'écrit lui aussi $f_j = \sum_{i=1}^n p_{ij} e_i$. On note

$$F_j = \begin{pmatrix} p_{ij} \\ \vdots \\ p_{ij} \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} F_1 & | & \cdots & | & F_n \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} = (p_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Ainsi on a

$$\begin{aligned} u = x'_1 f_1 + \cdots + x'_n f_n &= x'_1 \Big(\sum_{i=1}^n p_{i1} e_i \Big) + \cdots + x'_n \Big(\sum_{i=1}^n p_{in} e_i \Big) \\ &= \sum_{1 \leq j \leq n} \sum_{1 \leq i \leq n} p_{ij} x'_j e_i = \sum_{1 \leq i \leq n} \Big(\sum_{1 \leq j \leq n} p_{ij} x'_j \Big) e_i \\ &= \Big(\sum_{1 \leq j \leq n} p_{1j} x'_j \Big) e_1 + \cdots + \Big(\sum_{1 \leq j \leq n} p_{nj} x'_j \Big) e_n. \end{aligned}$$

Ainsi, de part l'unicité de la décomposition dans un base, on a

$$x_i = \sum_{1 \leq j \leq n} p_{ij} x'_j.$$

Matriciellement cela s'écrit

$$X = PX'.$$

La matrice opére donc sur les nouvelles coordonnées de u et donne ses anciennes coordonnées. Pour cette raison la matrice P est appelée la *matrice de passage* associée au changement de base.

Si on écrit $e_j = \sum_{i=1}^n q_{ij} f_i$ alors la matrice $Q = (q_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est telle que $X' = QX$. Ainsi $X = PQX$, ceci pour tout X. Cela signifie que $PQ = I_n$ et donc que P est inversible et $Q = P^{-1}$.

Exercice 9. *Mettons en pratique tout cela sur un exemple simple : prenons les deux bases* $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ *et* $\mathcal{F} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$ *de* $\mathbb{R}^2$.

1. *Exprimer les coordonnées de* $\begin{pmatrix}0\\1\end{pmatrix}$ *et* $\begin{pmatrix}2\\1\end{pmatrix}$ *dans la base* $\mathcal{E}$*. En déduire la matrice de passage* P *associée au changement de base de* $\mathcal{E}$ *à* $\mathcal{F}$*.*
2. *Grâce à* P*, montrer que si* $v \in \mathbb{R}^2$ *a pour coordonnées* (x, y) *dans la base* $\mathcal{F}$*, alors ses coordonnées dans la base* $\mathcal{E}$ *sont* $(y - x, x + y)$*.*

On veut maintenant aller dans l'autre sens : trouver les coordonnées d'un vecteur dans $\mathcal{F}$ *à partir de ses coordonnées dans* $\mathcal{E}$*.*

1. *Déterminer l'inverse de la matrice* P*.*
2. *Grâce à* P^{-1}*, montrer que si* $w \in \mathbb{R}^2$ *a pour coordonnées* (a, b) *dans la base* $\mathcal{E}$*, alors ses coordonnées dans la base* $\mathcal{F}$ *sont* $(\frac{-x+y}{2}, \frac{x+y}{2})$*.*

Nous allons maintenant regarder l'effet d'un changement de base sur la représentation matricielle d'une application linéaire de V dans lui-même. Soit donc $f : V \to V$ une telle application linéaire. On note A sa matrice lorsque l'on utilise la base $\mathcal{E}$. C'est une matrice $n \times n$ et les coordonnées de $\varphi(u)$ dans la base $\mathcal{E}$ sont données par AX (voir la section 3.2).

Si on considère la seconde base $\mathcal{F}$, à φ est associée une matrice B elle aussi $n \times n$. Les coordonnées de de $f(u)$ dans la base $\mathcal{F}$ sont données par BX'.

Et ainsi $AX = PBX'$. Mais comme $X' = QX = P^{-1}X$ on trouve

$$AX = PBP^{-1}X.$$

Comme cela est vrai pour tout vecteur colonne X on conclut que PBP^{-1} est la matrice de f dans la base $\mathcal{E}$. Ainsi, la formule est

$$A = PBP^{-1}$$

et elle exprime l'effet d'un changement de base sur la représentation matricielle d'une application linéaire d'un espace vectoriel dans lui-même.

Exercice 10. *On considère l'application linéaire*

$$\begin{array}{rccc} f : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x, y) & \longmapsto & (x + y, x - y) \end{array}$$

ainsi que les bases $\mathcal{E}$ *et* $\mathcal{F}$ *de l'exercice 9.*

1. *Montrer que la matrice de* f *dans la base* $\mathcal{E}$ *est donnée par*
$$Mat(f)_{\mathcal{E}} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$
2. *Montrer que la matrice de* f *dans la base* $\mathcal{F}$ *est donnée par*
$$Mat(f)_{\mathcal{F}} = \begin{pmatrix} -3/2 & -1/2 \\ 1/2 & 3/2 \end{pmatrix}$$
3. *Retrouver* $Mat(f)_{\mathcal{E}}$ *a partir de* $Mat(f)_{\mathcal{F}}$ *grâce à la matrice de changement de base déjà donnée à l'exercice 9.*

3.8 Espaces propres et diagonalisation

Soit V un espace vectoriel de dimension n. L'application linéaire la plus simple est certainement l'application identité. Une fois une base $\mathcal{V} = \{v_1, \ldots, v_n\}$ de V choisie, sa matrice est bien sûr la matrice carrée $n \times n$ identité I_n.

Moins simple mais toujours très simple à calculer et à manipuler est le cas d'une application linéaire f associée à une matrice diagonale

$$D = \begin{pmatrix} d_1 & 0 & \ldots & 0 & 0 \\ 0 & d_2 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & d_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & d_n \end{pmatrix}.$$

Pour une telle application linéaire on a

$$f(v_1) = d_1 v_1, \quad \ldots, \quad f(v_n) = d_n v_n$$

et si $v = x_1 v_1 + \cdots + x_n v_n$ on a $f(x) = d_1 x_1 v_1 + \cdots + d_n x_n v_n$, ce que l'on retrouve en faisant le calcul DX avec $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$:

$$DX = \begin{pmatrix} d_1 x_1 \\ \vdots \\ d_n x_n \end{pmatrix}.$$

Si $r \geq 2$, alors nous avons vu à la remarque 3 que la matrice D^r est la matrice associée à $f^r = f \circ \cdots \circ f$ où f est composée r fois. Un calcul immédiat montre que

$$D^r = \begin{pmatrix} d_1^r & 0 & \ldots & 0 & 0 \\ 0 & d_2^r & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & d_{n-1}^r & 0 \\ 0 & 0 & \ldots & 0 & d_n^r \end{pmatrix}.$$

L'itération de l'application linéaire f se calcule donc de manière trés simple quand la matrice est sous cette forme diagonale.

Ces observations motivent la question suivante : peut-on trouver une base de V dans laquelle la représentation matricielle de f est une matrice diagonale ?

3.8.1 Vecteurs propres et valeurs propres

Avec la base $\mathcal{V}$ choisie plus haut, on rappelle que parler de f ou de la matrice A est équivalent. De même on a maintenant clairement en tête l'identification faite entre un vecteur v de V et le vecteur colonne $X \in \mathbb{R}^n$ de ses coordonnées.

Suite à la discussion dans la section précédente on cherche maintenant s'il existe une valeur $\lambda \in \mathbb{R}$ et un vecteur X tel que $AX = \lambda X$. Bien sûr cela n'a d'intérêt que si le vecteur colonne X est non nul. Cette propriété s'écrit

$$AX - \lambda X = 0 \text{ c'est à dire } X \in \ker(A - \lambda I_n).$$

Ici $A - \lambda I_n$ est une matrice $n \times n$. La formule du rang du théorème 3.7 donne cependant l'équivalence entre

$$\ker(A - \lambda I_n) = \{0\} \;\Leftrightarrow\; A - \lambda I_n \text{ est une matrice inversible.}$$

Donc pouvoir trouver $X \neq 0$ tel que $AX - \lambda X = 0$ est équivalent à avoir $A - \lambda I_n$ non inversible, ce qui s'écrit $\det(A - \lambda I_n) = 0$ par le théorème 3.18.

Définition 3.22. Soit A une matrice $n \times n$. Un réel λ est une *valeur propre* de A s'il existe un vecteur non nul X tel que

$$AX = \lambda X.$$

La discussion précédente montre que λ est une valeur propre de A si et seulement si λ est une racine du polynôme

$$p_A(\lambda) = \det(A - \lambda I_n),$$

qu'on appelle le *polynôme caractéristique* de A.

Exemple 7. *On considère la matrice*

$$A = \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix}.$$

1. *Calculer* $\begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *et en déduire une première valeur propre de A.*
2. *Calculer le polynôme caractéristique de A. En déduire toutes les valeurs propres de A.*

Vous connaissez probablement le résultat célèbre suivant, dont la preuve mêle des arguments d'algèbre et d'analyse :

Un polynôme de degé n à coefficients réels[1] admet au plus n racines réelles (comptées avec leur multiplicité).

Ainsi, une matrice $n \times n$ a ainsi au plus n valeurs propres, car son polynôme caractéristique est de degré n.

Après avoir montré avec la discussion précédente comment déterminer les valeurs propres d'une matrice, on s'intéresse désormais aux vecteurs naturellement associés à celles-ci.

Définition 3.23. Soit A une matrice $n \times n$ et λ une valeur propre de A. Un *vecteur propre* de A associé à la valeur propre λ est un vecteur non-nul $X \in \mathbb{R}^n$ tel que

$$AX = \lambda X.$$

De manière équivalente, $X \in \mathbb{R}^n$ est un *vecteur propre* de A associé à la valeur propre λ s'il est non-nul et si

$$X \in \ker(A - \lambda I_n).$$

Vous avez donc tous les outils en mains pour déterminer les vecteurs propres d'une matrice : il s'agit ici de calculer un noyau.

Exercice 11. *On considère à nouveau la matrice*

$$A = \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix}.$$

Déterminer les vecteurs propres associés aux valeurs propres de A.

3.8.2 Diagonalisation effective

La définition suivante qui justifie ce qui précède est une simple reformulation.

Définition 3.24. Soit A une matrice $n \times n$. On dit que A est diagonalisable, s'il existe une matrice diagonale D et une matrice inversible P telles que

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

Dit d'une autre manière, A est diagonalisable si elle admet une base constituée de vecteurs propres.

1. Ce résultat sur les polynômes à coefficients réels est certainement une des raisons historiques de l'introduction des nombres complexes que certains d'entre vous connaissent. On a en effet le résultat suivant qui est encore plus satisfaisant : un polynôme de degré n à coefficients réels ou complexes admet précisemment n racines complexes (comptées avec leur multiplicité). Les nombres complexes ont peu d'applications en économie. Nous choisissons donc de ne considérer que des espaces vectoriels liés aux nombres réels ici et donc en particulier des matrices à coefficients réels.

Toutes les matrices ne sont pas diagonalisables, mais la partie précédent nous donne les outils pour déterminer si une matrice fixée l'est ou non : il suffit de déterminer toutes les valeurs propres et les vecteurs propres de A, puis de voir si l'on peut en extraire une base. Les résultats suivants sont utiles pour gagner du temps.

Si A est une matrice $n \times n$ et λ en est une valeur propre, on note

$$E_\lambda = \ker(A - \lambda I_n).$$

Remarque 6. *On a très envie de dire qu'il s'agit des vecteurs propres de A associés à λ, mais ce n'est que presque le cas : on leur a ici ajouté le vecteur nul.*

Proposition 3.25. *Soit A une matrice $n \times n$ et λ et μ deux valeurs propres distinctes de A. Les sous-ensembles E_λ et E_μ sont des sous-espace vectoriels de $\mathbb{R}^n$. De plus, si $\mathcal{A}$ est une base de E_λ et si $\mathcal{B}$ est une base de E_μ, alors la famille $\mathcal{A} \cup \mathcal{B}$ est libre.*

Démonstration. Les sous-ensembles E_λ et E_μ sont clairement des sous-espaces vectoriels, ce sont des noyaux.

Notons $\mathcal{A} = \{a_1, ..., a_k\}$ et $\mathcal{B} = \{b_1, ..., b_s\}$ et considérons une combinaison linéaire nulle

$$x_1a_1 + ... + x_ka_k + x_{k+1}b_1 + ... + x_{k+s}b_s = 0$$

En multipliant ce vecteur par A on obtient une seconde égalité

$$\lambda(x_1a_1 + ... + x_ka_k) + \mu(x_{k+1}b_1 + ... + x_{k+s}b_s) = 0$$

et en soustrayant λ-fois la première à la seconde, on obtient alors

$$(\mu - \lambda)(x_{k+1}b_1 + ... + x_{k+s}b_s) = 0.$$

Maintenant, $\lambda \neq \mu$ donc on peut très bien diviser cette équation par $\mu - \lambda$, on obtient alors

$$x_{k+1}b_1 + ... + x_{k+s}b_s = 0$$

ce qui implique que x_{k+1},...,x_{k+s} sont tous nuls car $\mathcal{B}$ est une base. Le même raisonnement montre que x_1,...,x_k le sont aussi. □

Corollaire 3.26. *Soit A une matrice $n \times n$, dont on note λ_1,...,λ_k les valeurs propres. La matrice A est diagonalisable si et seulement si*

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

Démonstration. En effet par la proposition 3.25, on voit que l'on obtient une base de vecteurs propres pour A si et seulement si l'union des bases de tous les espaces propres est de cardinal n. □

Ce dernier résultat est admis mais permet souvent de déterminer si une matrice est diagonalisable sans même avoir à calculer ses vecteurs propres.

Proposition 3.27. *Soit A une matrice $n \times n$. Le polynôme caractéristique de A s'écrit de la forme*

$$p_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}(\lambda - \lambda_k)^{n_k}Q(\lambda),$$

où $Q(\lambda)$ est un polynôme qui n'a pas de solution réelle.

De plus pour tout $i = 1, ..., k$, on a

$$1 \leq \dim(E_{\lambda_i}) \leq n_i,$$

autrement dit : "la multiplicité algébrique des valeurs propres est supérieure à leur multiplicité géométrique".

Corollaire 3.28. *Si A une matrice $n \times n$ a n valeurs propres distinctes, alors elle est diagonalisable.*

Démonstration. Comme le polynôme caractéristique p_A est de degré n, le fait que A ait n valeurs propres distinctes λ_1...,λ_n implique que

$$p_A = (\lambda - \lambda_1)...(\lambda - \lambda_n)$$

Mais dans ce cas on est obligés d'avoir par la proosition précédente

$$\dim(E_{\lambda_1}) = ... = \dim(E_{\lambda_n}) = 1.$$

On a alors bien $n = 1 + ... + 1 = \dim(E_{\lambda_1}) = ... = \dim(E_{\lambda_n})$ et le corollaire 3.25 garantit que A est diagonalisable. □

Exercice 12. *On considère la matrice*

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

1. *Déterminer les valeurs propres de A.*
2. *En déduire que A est diagonalisable.*

Exercice 13. *On considère la matrice*

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

1. *Déterminer les valeurs propres de A.*
2. *Déterminer les sous-espaces E_{λ_1} et E_{λ_2} associés aux valeurs propres de A et montrer que A est diagonalisable.*
3. *À l'aide des questions précédentes, trouver une matrice P inversible et une matrice D diagonale tels que*

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

4. *En déduire la valeur de A^{17}.*