Université de Rennes 1

L1-AL1 – Introduction à l'algèbre linéaire

Extraits adaptés en français du livre d'Otto Bretscher, Linear Algebra with Applications, 3rd edition, Prentice Hall (2004)

March 8, 2013

Chapter 1

Equations linéaires

1.1 Introduction aux systèmes linéaires

1.1.1 Un exemple de système linéaire dû aux Chinois il y a 2000 ans.

- 1 ballot de riz de catégorie inférieure (x) mélangé avec 2 ballots de riz de catégorie moyenne (y) et 3 de catégorie supérieure (z) ont un volume total de 39 Dou (un Dou est une unité de volume représentant environ 2 litres).

- 1 ballot de riz de catégorie inférieure (x) mélangé avec 3 ballots de riz de catégorie moyenne (y) et 2 de catégorie supérieure (z) ont un volume total de 34 Dou.

- 3 ballots de riz de catégorie inférieure (x) mélangé avec 2 ballots de riz de catégorie moyenne (y) et 1 de catégorie supérieure (z) ont un volume total de 26 Dou.

On se demande quel est le volume d'un ballot de riz de chaque catégorie.

Le système "linéaire" correspondant s'écrit sous la forme (I) suivante

$$\left|\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 39 \\ x & + & 3y & + & 2z & = & 34 \\ 3x & + & 2y & + & z & = & 26 \end{array}\right|$$

où

x est le volume d'un ballot de catégorie inférieure,

y est le volume d'un ballot de catégorie moyenne,

z est le volume d'un ballot de catégorie supérieure.

Pour résoudre ce système, on doit le transformer dans un système diagonal de la forme

$$\left|\begin{array}{ccccc} x & & & = & ... \\ & y & & = & ... \\ & & z & = & ... \end{array}\right|$$

En d'autres termes, il faut éliminer les termes non diagonaux (encadrés) et rendre les coefficients des termes diagonaux égaux à 1,

$$\left|\begin{array}{rcrcrcl} x & + & \boxed{2y} & + & \boxed{3z} & = & 39 \\ \boxed{x} & + & 3y & + & \boxed{2z} & = & 34 \\ \boxed{3x} & + & \boxed{2y} & + & z & = & 26 \end{array}\right|$$

On le fait pas à pas, une variable à la fois. On peux éliminer la variable x de la deuxième équation en soustrayant la première équation (L1) de la deuxième (L2), c'est-à-dire (L2) $\leftarrow$ (L2) $-$ (L1) :

$$\left|\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 39 \\ x & + & 3y & + & 2z & = & 34 \\ 3x & + & 2y & + & z & = & 26 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 39 \\ & & y & - & z & = & -5 \\ 3x & + & 2y & + & z & = & 26 \end{array}\right|$$

Pour éliminer la variable x de la troisième équation, on soustrait la première équation de la troisième 3 fois. En multipliant la première équation par 3 on obtient

$$3x + 6y + 9z = 117 \quad (3 \times (\text{L1}))$$

que l'on soustrait de la troisième équation, soit (L3) $\leftarrow$ (L3) $-$ 3 $\times$ (L1):

$$\left|\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 39 \\ & & y & - & z & = & -5 \\ 3x & + & 2y & + & z & = & 26 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 39 \\ & & y & - & z & = & -5 \\ & - & 4y & - & 8z & = & -91 \end{array}\right|$$

De la même manière on élimine la variable y en dessous et au dessus de la diagonale en faisant successivement les opérations

$$(\text{L1}) \leftarrow (\text{L1}) - 2 \times (\text{L2}) \quad \text{et} \quad (\text{L3}) \leftarrow (\text{L3}) + 4 \times (\text{L2}),$$

ce qui conduit à

$$\left|\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 39 \\ & & y & - & z & = & -5 \\ & - & 4y & - & 8z & = & -91 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & & & + & 5z & = & 49 \\ & & y & - & z & = & -5 \\ & & & - & 12z & = & -111 \end{array}\right|$$

Avant d'éliminer la variable z au dessus de la diagonale, on rend le coefficient en z de la diagonale égal à 1 en divisant la dernière équation par -12, ce qui revient à l'opération (L3) $\leftarrow$ (L3) $\div$ 12,

$$\left|\begin{array}{ccccccc} x & & & + & 5z & = & 49 \\ & & y & - & z & = & -5 \\ & & & - & 12z & = & -111 \end{array}\right| \rightarrow \left|\begin{array}{ccccccc} x & & & + & 5z & = & 49 \\ & & y & - & z & = & -5 \\ & & & & z & = & 9.25 \end{array}\right|$$

Enfin, on élimine la variable z au dessus de la diagonale par les opérations

$$(\text{L1}) \leftarrow (\text{L1}) - 5 \times (\text{L3}) \quad \text{et} \quad (\text{L2}) \leftarrow (\text{L2}) + (\text{L3})$$

ce qui donne

$$\left|\begin{array}{ccccccc} x & & & + & 5z & = & 49 \\ & & y & - & z & = & -5 \\ & & & & z & = & 9.25 \end{array}\right| \rightarrow \left|\begin{array}{ccccc} x & & & = & 2.75 \\ & y & & = & 4.25 \\ & & z & = & 9.25 \end{array}\right|$$

En conclusion :

- 1 ballot de riz de catégorie inférieure a un volume 2.75 Dou
- 1 ballot de riz de catégorie moyenne a un volume 4.25 Dou
- 1 ballot de riz de catégorie supérieure a un volume de 9.25 Dou

On vérifie **a posteriori** que $x = 2.75$, $y = 4.25$ et $z = 9.25$ satisfont bien

$$\begin{array}{rcrcrcr} 2.75 & + & 2 \times 4.25 & + & 3 \times 9.25 & = & 39 \\ 2.75 & + & 3 \times 4.25 & + & 2 \times 9.25 & = & 34 \\ 3 \times 2.75 & + & 2 \times 4.25 & + & 9.25 & = & 26 \end{array}$$

Il faut prendre l'habitude de vérifier ses résultats, ce qui est toujours faisable en algèbre linéaire

1.1.2 Interprétation géométrique

Chaque équation du système (I) définit un plan dans l'espace à trois dimensions x, y, z. Les solution de (I) sont donc les points qui sont dans les trois plans simultanément.

Nous avons trouvé que l'ensemble des solution était constitué d'un unique point de coordonnées $(2.75, 4.25, 9.25)$.

On note que l'intersection de trois plans quelconque est en général un point.

Dans certains cas particuliers cette intersection peut être vide, ou encore une droite entière ou encore tout un plan.

Donc un système linéaire en trois variables peut avoir une solution unique, pas de solution du tout ou une infinité de solutions.

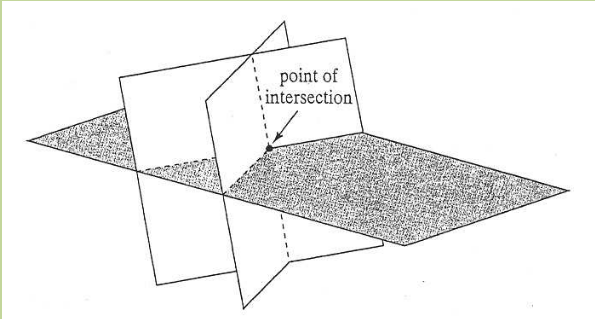


Figure 1.1: Trois plans avec un seul point d'intersection

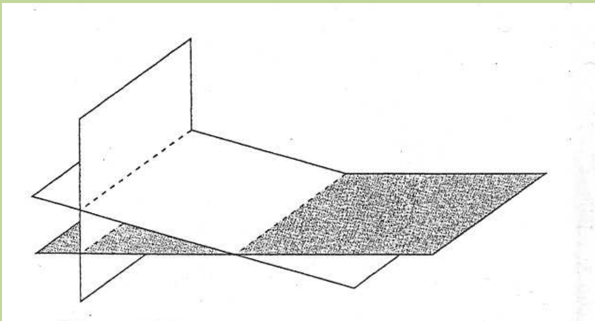


Figure 1.2: Trois plans d'intersection vide

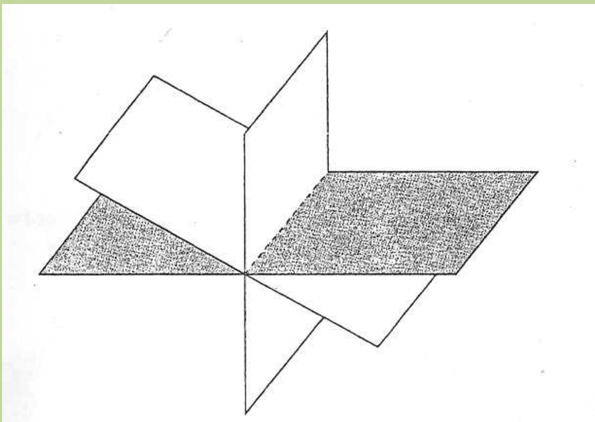


Figure 1.3: Trois plans avec une droite entière d'intersection

1.1.3 Un exemple d'un système qui n'a pas de solutions

Considérons le système (II) suivant :

$$\left|\begin{array}{ccccccc} 2x & + & 4y & + & 6z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 0 \end{array}\right|$$

On utilise la même méthode de résolution par élimination :

$$(\mathrm{L1}) \leftarrow (\mathrm{L1}) \div 2$$

$$\left|\begin{array}{ccccccc} 2x & + & 4y & + & 6z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 0 \end{array}\right| \rightarrow \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 0 \end{array}\right|$$

$$(\mathrm{L2}) \leftarrow (\mathrm{L2}) - 4(\mathrm{L1}), \quad (\mathrm{L3}) \leftarrow (\mathrm{L3}) - 7(\mathrm{L1})$$

$$\rightarrow \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ & - & 3y & - & 6z & = & 3 \\ & - & 6y & - & 12z & = & 0 \end{array}\right|$$

$$(\mathrm{L2}) \leftarrow (\mathrm{L2}) \div (-3)$$

$$\rightarrow \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ & & y & + & 2z & = & -1 \\ & - & 6y & - & 12z & = & 0 \end{array}\right|$$

$$(\mathrm{L1}) \leftarrow (\mathrm{L1}) - 2(\mathrm{L2}), \quad (\mathrm{L3}) \leftarrow (\mathrm{L3}) + 6(\mathrm{L2})$$

$$\rightarrow \left|\begin{array}{ccccccc} x & & & - & z & = & 2 \\ & & y & + & 2z & = & -1 \\ & & & & 0 & = & -6 \end{array}\right|$$

Quelque soient les valeurs prises pour x, y et z, l'équation $0 = -6$ ne peut jamais être satisfaite. Le système (II) est **inconsistant**, c'est-à-dire qu'il n'a aucune solution.

1.1.4 Un exemple de système avec une infinité de solutions

Considérons le système (III) suivant:

$$\left|\begin{array}{ccccccc} 2x & + & 4y & + & 6z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 6 \end{array}\right|$$

On utilise la même méthode de résolution par élimination :

$$(\text{L1}) \leftarrow (\text{L1}) \div 2$$

$$\left|\begin{array}{ccccccc} 2x & + & 4y & + & 6z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 6 \end{array}\right| \to \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ 4x & + & 5y & + & 6z & = & 3 \\ 7x & + & 8y & + & 9z & = & 6 \end{array}\right|$$

$$(\text{L2}) \leftarrow (\text{L2}) - 4(\text{L1}), \quad (\text{L3}) \leftarrow (\text{L3}) - 7(\text{L1})$$

$$\to \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ & - & 3y & - & 6z & = & 3 \\ & - & 6y & - & 12z & = & 6 \end{array}\right|$$

$$(\text{L2}) \leftarrow (\text{L2}) \div (-3)$$

$$\to \left|\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ & & y & + & 2z & = & -1 \\ & - & 6y & - & 12z & = & 6 \end{array}\right|$$

$$(\text{L1}) \leftarrow (\text{L1}) - 2(\text{L2}), \quad (\text{L3}) \leftarrow (\text{L3}) + 6(\text{L2})$$

$$\to \left|\begin{array}{ccccccc} x & & & - & z & = & 2 \\ & & y & + & 2z & = & -1 \\ & & & & 0 & = & 0 \end{array}\right| \to \left|\begin{array}{ccccccc} x & & & - & z & = & 2 \\ & & y & + & 2z & = & -1 \end{array}\right|$$

Une fois que l'on a omis la dernière équation $0 = 0$, il nous reste deux équations avec trois inconnues. L'ensemble des solutions de (III) est l'intersection de deux plans non parallèles, i.e. une droite. Ces deux équations peuvent s'́ecrire sous la forme

$$\left|\begin{array}{cccc} x & & = & z+2 \\ & y & = & -2z-1 \end{array}\right|$$

et x et y sont entièrement définis pas la valeur de z.

Pour chaque choix d'une valeur de z , on obtient une solution. Par exemple,

- Si on choisit $z = 1$, alors $x = z + 2 = 3$ et $y = -2z - 1 = -3$. La solution est
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}.$$

- Si on choisit $z = 7$, alors $x = z + 2 = 9$ et $y = -2z - 1 = -15$. La solution est
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -15 \\ 7 \end{bmatrix}.$$

- Plus généralement si on prend $z = t$ de manière arbitraire, on obtient $x = t + 2$ et $y = -2t - 1$. La solution générale est
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t+2 \\ 2t-1 \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

  qui est la droite affine dans l'espace passant par le point $(2, -1, 0)$ et dirigée par le "**vecteur** "
$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

1.2 Matrices, vecteurs et élimination de Gauss-Jordan

1.2.1 Introduction aux matrices et aux vecteurs

Quand les chinois des temps anciens devaient résoudre un système de la forme

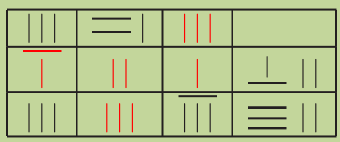
$$\left| \begin{array}{rcrcrcr} 3x & + & 21y & - & 3z & = & 0 \\ -6x & - & 2y & - & z & = & 62 \\ 2x & - & 3y & + & 8z & = & 32 \end{array} \right|$$

Ils prenaient tous les nombres présents dans le système et les plaçaient dans un tableau :

3	21	-3	0
-6	-2	-1	62
2	-3	8	32

Toutes les informations concernant ce système, à part les noms de variables, sont incluses dans ce tableau.

En fait, les entrées étaient représentées par des traits comme cela est montré dans le tableau suivant. Les nombres positifs sont représentés en noir, les négatifs en rouge (ce qui apparait en gris sur les imprimantes noir et blanc). Les équations étaient résolues pas des manipulations "à la main" sur les traits. La solution par ce procédé est laissée en exercice.



Aujourd'hui, un tableau de la forme

$$\begin{bmatrix} 3 & 21 & -3 & 0 \\ -6 & -2 & -1 & 62 \\ 2 & -3 & 8 & 32 \end{bmatrix}$$

est appelé **une matrice**. Il s'agit d'une matrice 3×4 car elle a 3 lignes et 4 colonnes. Les nombres qui la constituent sont ses coefficients.

$$\begin{array}{rc}
 & \text{les quatre colonnes de la matrice} \\
 & \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \end{array} \\
\begin{array}{r} \rightarrow \\ \text{les trois lignes de la matrice} \ \rightarrow \\ \rightarrow \end{array} & \begin{bmatrix} 3 & 21 & -3 & 0 \\ -6 & -2 & -1 & 62 \\ 2 & -3 & 8 & 32 \end{bmatrix}
\end{array}$$

Il faut noter que la première colonne correspond à la première inconnue du système, tandis que la première ligne correspond à la première équation.

On numérote les coefficients d'une matrice 3×4 générale notée A avec des doubles indices comme dans ce qui suit :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

Le premier indice correspond à la ligne et le deuxième à la colonne. Le coefficient a_{ij} est celui qui se trouve à la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne.

Dans l'exemple qui a été présenté, la **taille** de la matrice est 3×4 . Il peut y avoir des matrices de toutes les tailles possibles, comme par exemple,

$1 \times 3 = 1$ ligne et 3 colonnes,
$3 \times 1 = 3$ lignes et une colonne,
$3 \times 3 = 3$ lignes et 3 colonnes,
$20645 \times 953889 = 20645$ lignes et 953889 colonnes, etc...

1.2.2 Matrices et vecteurs: définitions et exemples

Définition 1.2.1. [Matrices]

- *Une* **matrice de taille** $n \times m$ *est un tableau rectangulaire de nombres avec n lignes et m colonnes. On note $A = [a_{i,j}]$, une telle matrice, l'élément $a_{i,j}$ étant placé sur la $i^{\text{ième}}$ ligne et la j-ième colonne.*

- *On dit que deux matrices $A = [a_{ij}]$ et $B = [b_{ij}]$ sont* **égales** *si et seulement si elles ont la même taille et si pour chaque ligne i et chaque colonne j, les coefficients correspondants sont égaux, i.e. $a_{ij} = b_{ij}$.*

- *Si une matrice $A = [a_{ij}]$ a autant de lignes que de colonnes (A est une matrice $n \times n$, $n \in \mathbb{N}$), on dit que A est une matrice* **carrée***, et les coefficients $a_{11}, a_{22}, ..., a_{nn}$ constituent la diagonale de A.*

- *Une matrice carrée de taille $n \times n$, $A = [a_{ij}]$, est dite* **diagonale** *si et seulement si les coefficients au dessus et en dessous de la diagonale sont nuls, c'est-à-dire $a_{ij} = 0$ quand $i \neq j$.*

- *Une matrice carrée de taille $n \times n$, $A = [a_{ij}]$, est dite* **triangulaire supérieure** *si et seulement si les coefficients au dessous de la diagonale sont nuls, c'est-à-dire $a_{ij} = 0$ quand $i > j$.*

- *Une matrice carrée de taille $n \times n$, $A = [a_{ij}]$, est dite* **triangulaire inférieure** *si et seulement si les coefficients au dessus de la diagonale sont nuls, c'est-à-dire $a_{ij} = 0$ quand $i < j$.*

- *Une matrice dont tous les coefficients sont nuls est appelée la* **matrice nulle** *et est notée 0.*

- *Une matrice carrée de taille $n \times n$ qui est diagonale, et dont tous les coefficients diagonaux sont égaux à 1, est appelée la* **matrice identité** *et est souvent notée I_n.*

Par exemple, considérons les matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} 5 & 0 & 0 \\ 4 & 0 & 0 \\ 3 & 2 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 5 & 7 & -5 & -7 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & -5 & 4 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

La matrice A est 2×3, B, C, D, E et F sont carrées. C est diagonale, C, D et F sont triangulaires supérieures, C et E sont triangulaire inférieures.

Définition 1.2.2. [Vecteurs]

- *Une matrice qui n'a qu'une seule colonne, (i.e. de taille $n \times 1$), est appelée* **un vecteur colonne**, *ou simplement un* **vecteur**. *Les coefficients sont dans ce cas les* **composantes** *du vecteur.*

- *L'ensemble de tous les vecteurs qui ont n composantes est noté $\mathbb{R}^n$.*

- *Une matrice qui n'a qu'une seule ligne (i.e. de taille $1 \times n$) est appelée* **vecteur ligne**.

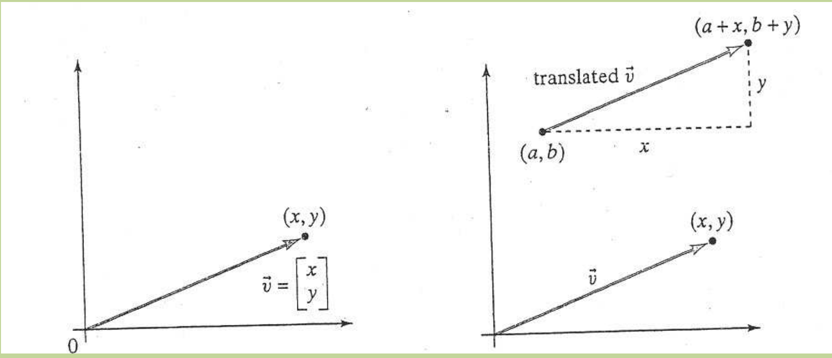


Figure 1.4: Représentation géométrique de vecteurs

- *Dans toute la suite, chaque fois que l'on parlera de* **vecteur**, *il s'agira de vecteur colonne.*

Un exemple de vecteur est donné par

$$\begin{bmatrix} 1 \\ 2 \\ 9 \\ 1 \end{bmatrix} \in \mathbb{R}^4$$

et

$$\begin{bmatrix} 1 & 5 & 5 & 8 & -9 \end{bmatrix}$$

est un vecteur ligne à 5 composantes.

Il faut noter que dans une matrice A de taille $n \times m$, les m colonnes de A sont des vecteurs de $\mathbb{R}^n$. Dans les cours de mathématique ou de physique on a introduit les vecteurs avec un point de vue géométrique. Par exemple dans le plan cartésien,

$$\overrightarrow{v} = \begin{bmatrix} x \\ y \end{bmatrix}$$

est la flèche qui relie l'origine au point (x, y). Cela peut-être aussi la flèche qui relie un point (a, b) au point $(a + x, b + y)$.

Quand on considère un ensemble infini de vecteurs, la représentation par flèches est impraticable. Dans ce cas, il est commode de représenter le vecteur

$$\overrightarrow{v} = \begin{bmatrix} x \\ y \end{bmatrix}$$

simplement par le point (x, y).

Par exemple l'ensemble des vecteurs

$$\overrightarrow{v} = \begin{bmatrix} x \\ x+1 \end{bmatrix}$$

$x \in \mathbb{R}$, peut être représenté par le droite $y = x + 1$.

Dans la suite on s'intéresse à des vecteurs numériques, une liste de nombres, que l'on peut représenter à l'aide d'une colonne. Dans le monde numérique, les informations sont stockées dans des chaines de nombres, i.e. des vecteurs. Par exemple un passage de musique de 10 secondes est stocké dans un vecteur qui a 440 000 composantes. Une photo satellite prise pour les prévisions météo est transmise à la terre par une chaine de nombres.

1.2.3 Lien entre les systèmes linéaires et les matrices

Avant de donner une méthode générale on traite en détail deux exemples.

Example 1.2.1. *On considère le système*

$$\left| \begin{array}{ccccccc} 2x & + & 8y & + & 4z & = & 2 \\ 2x & + & 5y & + & z & = & 5 \\ 4x & + & 10y & - & z & = & 1 \end{array} \right|$$

On considère la **matrice des coefficients** *du système*

$$\begin{bmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{bmatrix}$$

On peut ausi considérer la matrice dite **la matrice augmentée** *du système*

$$\begin{bmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{bmatrix}$$

Pour bien distinguer les coefficients du système, on met une ligne en pointillé dans la matrice augmenté comme suit :

$$\left[\begin{array}{ccc:c} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{array} \right]$$

Pour résoudre le système, il est plus efficace de faire les éliminations sur la matrice augmentée que dans le système lui-même. D'un point de vue conceptuel, les deux approches sont équivalentes, mais travailler avec la matrice augmentée demande moins d'écriture et est plus simple à lire.

- *Diviser une équation par un scalaire correspond à diviser une ligne par un scalaire*
- *Ajouter à une équation un multiple d'une autre équation correspond à ajouter à une ligne un multiple d'une autre ligne.*

On le fait conjointement sur le système et sur la matrice

$$(\text{L1}) \leftarrow (\text{L1}) \div 2$$

$$\left|\begin{array}{rcrcrcr} 2x & + & 8y & + & 4z & = & 2 \\ 2x & + & 5y & + & z & = & 5 \\ 4x & + & 10y & - & z & = & 1 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 1 \\ 2x & + & 5y & + & z & = & 5 \\ 4x & + & 10y & - & z & = & 1 \end{array}\right|$$

$$\left[\begin{array}{ccccc} 2 & 8 & 4 & \vdots & 2 \\ 2 & 5 & 1 & \vdots & 5 \\ 4 & 10 & -1 & \vdots & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccccc} 1 & 4 & 2 & \vdots & 1 \\ 2 & 5 & 1 & \vdots & 5 \\ 4 & 10 & -1 & \vdots & 1 \end{array}\right]$$

$$(\text{L2}) \leftarrow (\text{L2}) - 2(\text{L1}), \quad (\text{L3}) \leftarrow (\text{L3}) - 4(\text{L1})$$

$$\left|\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 1 \\ 2x & + & 5y & + & z & = & 5 \\ 4x & + & 10y & - & z & = & 1 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 1 \\ & - & 3y & - & 3z & = & 3 \\ & - & 6y & - & 9z & = & -3 \end{array}\right|$$

$$\left[\begin{array}{ccccc} 1 & 4 & 2 & \vdots & 1 \\ 2 & 5 & 1 & \vdots & 5 \\ 4 & 10 & -1 & \vdots & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccccc} 1 & 4 & 2 & \vdots & 1 \\ 0 & -3 & -3 & \vdots & 3 \\ 0 & -6 & -9 & \vdots & -3 \end{array}\right]$$

$$(\text{L2}) \leftarrow (\text{L2}) \div (-3)$$

$$\left|\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 1 \\ & - & 3y & - & 3z & = & 3 \\ & - & 6y & - & 9z & = & -3 \end{array}\right| \rightarrow \left|\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 1 \\ & & y & + & z & = & -1 \\ & - & 6y & - & 9z & = & -3 \end{array}\right|$$

$$\begin{bmatrix} 1 & 4 & 2 & \vdots & 1 \\ 0 & -3 & -3 & \vdots & 3 \\ 0 & -6 & -9 & \vdots & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 & 2 & \vdots & 1 \\ 0 & 1 & 1 & \vdots & -1 \\ 0 & -6 & -9 & \vdots & -3 \end{bmatrix}$$

$$(\mathrm{L1}) \leftarrow (\mathrm{L1}) - 4(\mathrm{L2}), \quad (\mathrm{L3}) \leftarrow (\mathrm{L3}) + 6(\mathrm{L2})$$

$$\left| \begin{array}{rrrrrrr} x & + & 4y & + & 2z & = & 1 \\ & & y & + & z & = & -1 \\ & - & 6y & - & 9z & = & -3 \end{array} \right| \rightarrow \left| \begin{array}{rrrrrrr} x & & & - & 2z & = & 5 \\ & & y & + & z & = & -1 \\ & & & - & 3z & = & -9 \end{array} \right|$$

$$\begin{bmatrix} 1 & 4 & 2 & \vdots & 1 \\ 0 & 1 & 1 & \vdots & -1 \\ 0 & -6 & -9 & \vdots & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -3 & -9 \end{bmatrix}$$

$$(\mathrm{L3}) \leftarrow (\mathrm{L3}) \div (-3)$$

$$\left| \begin{array}{rrrrrrr} x & & & - & 2z & = & 5 \\ & & y & + & z & = & -1 \\ & & & - & 3z & = & -9 \end{array} \right| \rightarrow \left| \begin{array}{rrrrrrr} x & & & - & 2z & = & 5 \\ & & y & + & z & = & -1 \\ & & & & z & = & 3 \end{array} \right|$$

$$\begin{bmatrix} 1 & 0 & -2 & \vdots & 5 \\ 0 & 1 & 1 & \vdots & -1 \\ 0 & 0 & -3 & \vdots & -9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 & \vdots & 5 \\ 0 & 1 & 1 & \vdots & -1 \\ 0 & 0 & 1 & \vdots & 3 \end{bmatrix}$$

$$(\mathrm{L1}) \leftarrow (\mathrm{L1}) + 2(\mathrm{L3}), (\mathrm{L2}) \leftarrow (\mathrm{L2}) - (\mathrm{L3})$$

$$\left| \begin{array}{rrrrrrr} x & & & - & 2z & = & 5 \\ & & y & + & z & = & -1 \\ & & & & z & = & 3 \end{array} \right| \rightarrow \left| \begin{array}{rrrrr} x & & & = & 11 \\ & y & & = & -4 \\ & & z & = & 3 \end{array} \right|$$

$$\begin{bmatrix} 1 & 0 & -2 & \vdots & 5 \\ 0 & 1 & 1 & \vdots & -1 \\ 0 & 0 & 1 & \vdots & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & \vdots & 11 \\ 0 & 1 & 0 & \vdots & -4 \\ 0 & 0 & 1 & \vdots & 3 \end{bmatrix}$$

Donc la solution peut-être représentée par le vecteur

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ -4 \\ 3 \end{bmatrix}$$

Dans l'exemple 1.2.1, la méthode fonctionne sans problème, dans le sens où on peut éliminer les coefficients hors de la diagonale et rendre ceux de la diagonale égale à 1. Ce processus fonctionne tant que l'on n'a pas de zéro sur la diagonale.

Ce n'est pas le cas du système traité dans l'exemple 1.2.2

Example 1.2.2.

$$\left|\begin{array}{ccccccccccc} & & & & x_3 & - & x_4 & - & x_5 & = & 4 \\ 2x_1 & + & 4x_2 & + & 2x_3 & + & 4x_4 & + & 2x_5 & = & 4 \\ 2x_1 & + & 4x_2 & + & 3x_3 & + & 3x_4 & + & 3x_5 & = & 4 \\ 3x_1 & + & 6x_2 & + & 6x_3 & + & 3x_4 & + & 6x_5 & = & 6 \end{array}\right|$$

qui a pour matrice augmentée :

$$B = \left[\begin{array}{cccccccc} 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\ 2 & 4 & 2 & 4 & 2 & \vdots & 4 \\ 2 & 4 & 3 & 3 & 3 & \vdots & 4 \\ 3 & 6 & 6 & 3 & 6 & \vdots & 6 \end{array}\right]$$

On veut mettre cette matrice dans une forme diagonale.

Pour suivre le déroulement du calcul, on introduit un curseur. Initialement, le curseur est placé en haut de la première colonne non nulle de la matrice.

$$\left[\begin{array}{cccccccc} \nearrow 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\ 2 & 4 & 2 & 4 & 2 & \vdots & 4 \\ 2 & 4 & 3 & 3 & 3 & \vdots & 4 \\ 3 & 6 & 6 & 3 & 6 & \vdots & 6 \end{array}\right]$$

Le but est de rendre la valeur du curseur égale à 1. On le fait en deux étapes comme suit.

Etape 1. *Si la valeur du curseur est nulle, cherche une ligne (vers le bas) où le coefficient de la colonne du curseur est non nulle et échange cette ligne avec la ligne du curseur. S'il n'y en a pas, bouge le curseur vers la droite et recommence.*

$$
\begin{array}{c}
\left[\begin{array}{ccccccc}
\nearrow 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\
2 & 4 & 2 & 4 & 2 & \vdots & 4 \\
2 & 4 & 3 & 3 & 3 & \vdots & 4 \\
3 & 6 & 6 & 3 & 6 & \vdots & 6
\end{array}\right]
\begin{array}{l}
\leftarrow \\
\leftarrow \\
\\
\\
\end{array}
\end{array}
$$

$$\downarrow$$

$$
\left[\begin{array}{ccccccc}
\nearrow 2 & 4 & 2 & 4 & 2 & \vdots & 4 \\
0 & 0 & 1 & -1 & -1 & \vdots & 4 \\
2 & 4 & 3 & 3 & 3 & \vdots & 4 \\
3 & 6 & 6 & 3 & 6 & \vdots & 6
\end{array}\right]
$$

Cela revient à faire l'opération suivante sur le système sans changer les solutions,

$$
\left|\begin{array}{ccccccccccc}
 & & & & x_3 & - & x_4 & - & x_5 & = & 4 \\
2x_1 & + & 4x_2 & + & 2x_3 & + & 4x_4 & + & 2x_5 & = & 4 \\
2x_1 & + & 4x_2 & + & 3x_3 & + & 3x_4 & + & 3x_5 & = & 4 \\
3x_1 & + & 6x_2 & + & 6x_3 & + & 3x_4 & + & 6x_5 & = & 6
\end{array}\right|
\begin{array}{l}
\leftarrow \\
\leftarrow \\
\\
\\
\end{array}
$$

$$\downarrow$$

$$
\left|\begin{array}{ccccccccccc}
2x_1 & + & 4x_2 & + & 2x_3 & + & 4x_4 & + & 2x_5 & = & 4 \\
 & & & & x_3 & - & x_4 & - & x_5 & = & 4 \\
2x_1 & + & 4x_2 & + & 3x_3 & + & 3x_4 & + & 3x_5 & = & 4 \\
3x_1 & + & 6x_2 & + & 6x_3 & + & 3x_4 & + & 6x_5 & = & 6
\end{array}\right|
$$

Etape 2. Divise la ligne du curseur par la valeur du curseur.

$$
\left[\begin{array}{ccccccc}
\nearrow 2 & 4 & 2 & 4 & 2 & \vdots & 4 \\
0 & 0 & 1 & -1 & -1 & \vdots & 4 \\
2 & 4 & 3 & 3 & 3 & \vdots & 4 \\
3 & 6 & 6 & 3 & 6 & \vdots & 6
\end{array}\right]
$$

$$\downarrow$$

$$\left[\begin{array}{ccccc:c} \nearrow 1 & 2 & 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & -1 & -1 & 4 \\ 2 & 4 & 3 & 3 & 3 & 4 \\ 3 & 6 & 6 & 3 & 6 & 6 \end{array}\right]$$

Cela revient à faire l'opération suivante sur le système sans changer les solutions,

$$\left|\begin{array}{ccccccccccc} 2x_1 & + & 4x_2 & + & 2x_3 & + & 4x_4 & + & 2x_5 & = & 4 \\ & & & & x_3 & - & x_4 & - & x_5 & = & 4 \\ 2x_1 & + & 4x_2 & + & 3x_3 & + & 3x_4 & + & 3x_5 & = & 4 \\ 3x_1 & + & 6x_2 & + & 6x_3 & + & 3x_4 & + & 6x_5 & = & 6 \end{array}\right|$$

$$\downarrow$$

$$\left|\begin{array}{ccccccccccc} x_1 & + & 2x_2 & + & x_3 & + & 2x_4 & + & x_5 & = & 2 \\ & & & & x_3 & - & x_4 & - & x_5 & = & 4 \\ 2x_1 & + & 4x_2 & + & 3x_3 & + & 3x_4 & + & 3x_5 & = & 4 \\ 3x_1 & + & 6x_2 & + & 6x_3 & + & 3x_4 & + & 6x_5 & = & 6 \end{array}\right|$$

Etape 3. Elimine tous les coefficients de la colonne du curseur qui ne sont pas sur la ligne du curseur en soustrayant à chaque ligne des multiples adaptés de la ligne du curseur.

$$\left[\begin{array}{ccccc:c} \nearrow 1 & 2 & 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & -1 & -1 & 4 \\ 2 & 4 & 3 & 3 & 3 & 4 \\ 3 & 6 & 6 & 3 & 6 & 6 \end{array}\right]$$

$$\downarrow$$

$$\left[\begin{array}{ccccc:c} \nearrow 1 & 2 & 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & -1 & -1 & 4 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 3 & -3 & 3 & 0 \end{array}\right]$$

Maintenant que l'on s'est occupé de la première colonne (qui correspond à la première variable), on bouge le curseur sur une nouvelle position, suivant *la diagonale et retoune à l'étape 1.*

$$\left[\begin{array}{cccccc:c} 1 & & 2 & 1 & 2 & 1 & 2 \\ 0 & \nearrow & 0 & 1 & -1 & -1 & 4 \\ 0 & & 0 & 1 & -1 & 1 & 0 \\ 0 & & 0 & 3 & -3 & 3 & 0 \end{array}\right]$$

Etape 4. Bouge le curseur une ligne plus bas et une colonne vers la droite. Retourne à l'étape 1.

On voit que la valeur du curseur est égal à 0*. Donc, on cherche un élément non nul sur la colonne du curseur (en dessous de lui), ce qui ne marche pas dans ce cas. Donc on passe à la colonne suivante :*

$$\left[\begin{array}{cccccc:c} 1 & 2 & & 1 & 2 & 1 & 2 \\ 0 & 0 & \nearrow & 1 & -1 & -1 & 4 \\ 0 & 0 & & 1 & -1 & 1 & 0 \\ 0 & 0 & & 3 & -3 & 3 & 0 \end{array}\right]$$

Le coefficient du curseur étant égal à 1, passe directement à l'étape 3:

$$\left[\begin{array}{cccccc:c} 1 & 2 & & 1 & 2 & 1 & 2 \\ 0 & 0 & \nearrow & 1 & -1 & -1 & 4 \\ 0 & 0 & & 1 & -1 & 1 & 0 \\ 0 & 0 & & 3 & -3 & 3 & 0 \end{array}\right]$$

$$\downarrow \quad \textit{Étape 3}$$

$$\left[\begin{array}{cccccc:c} 1 & 2 & & 0 & 3 & 2 & -2 \\ 0 & 0 & \nearrow & 1 & -1 & -1 & 4 \\ 0 & 0 & & 0 & 0 & 2 & -4 \\ 0 & 0 & & 0 & 0 & 6 & -12 \end{array}\right]$$

$$\downarrow \quad \textit{Étape 4}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & \vdots & -2 \\ 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\ 0 & 0 & 0 & \nearrow 0 & 2 & \vdots & -4 \\ 0 & 0 & 0 & 0 & 6 & \vdots & -12 \end{bmatrix}$$

$$\downarrow \quad \textit{Étape 1}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & \vdots & -2 \\ 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\ 0 & 0 & 0 & 0 & \nearrow 2 & \vdots & -4 \\ 0 & 0 & 0 & 0 & 6 & \vdots & -12 \end{bmatrix}$$

$$\downarrow \quad \textit{Étape 2}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & \vdots & -2 \\ 0 & 0 & 1 & -1 & -1 & \vdots & 4 \\ 0 & 0 & 0 & 0 & \nearrow 1 & \vdots & -2 \\ 0 & 0 & 0 & 0 & 6 & \vdots & -12 \end{bmatrix}$$

$$\downarrow \quad \textit{Étape 3}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & \vdots & 2 \\ 0 & 0 & 1 & -1 & 0 & \vdots & 2 \\ 0 & 0 & 0 & 0 & \nearrow 1 & \vdots & -2 \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 \end{bmatrix}$$

$$\downarrow \quad \textit{Étape 4}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & \vdots & 2 \\ 0 & 0 & 1 & -1 & 0 & \vdots & 2 \\ 0 & 0 & 0 & 0 & 1 & \vdots & -2 \\ 0 & 0 & 0 & 0 & \nearrow 0 & \vdots & 0 \end{bmatrix}$$

Lorsque l'on essaie d'appliquer l'étape 1 à cette matrice, on voit qu'il n'y a rien de plus à faire, car le curseur reste sur un 0 *: le processus de réduction est arrivé à sa fin. On pose*

$$E = \left[\begin{array}{ccccc:c} 1 & 2 & 0 & 3 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

On dit que la matrice E est la **forme réduite échelonnée par ligne de la matrice** B*. On écrit*

$$E = \mathrm{Frel}(B).$$

Définition 1.2.3. [Forme réduite échelonnée par ligne]. *Une matrice est sous la forme* **réduite échelonnée par ligne** *si elle satisfait les conditions suivantes :*

(a) Si une ligne a un coefficient non nul, alors son premier coefficient non nul (en partant de la gauche) appelé **pivot***, est égal à 1*

(b) Dans la colonne d'un pivot, tous les autres coefficients sont égaux à 0*.*

(c) Si une ligne a un pivot, alors toutes les autres lignes au dessus ont un pivot à gauche de celui-ci.

Une matrice sous forme réduite échelonnée par ligne peut avoir une ligne de 0 (comme dans l'exemple précédent), mais d'après (c), il ne peut s'agir que des dernières lignes de la matrice.

On a encadré les pivots dans la matrice $E = \mathrm{Frel}(B)$ de l'exemple précédent :

$$E = \left[\begin{array}{ccccc:c} \boxed{1} & 2 & 0 & 3 & 0 & 2 \\ 0 & 0 & \boxed{1} & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

qui représente le système

$$\left|\begin{array}{ccccccccccc} \boxed{x_1} & + & 2x_2 & & & + & 3x_4 & & & = & 2 \\ & & & & \boxed{x_3} & - & x_4 & & & = & 2 \\ & & & & & & & & \boxed{x_5} & = & -2 \end{array}\right|$$

On voit un "escalier" qui apparait en dessous des inconnues principales, identifiées comme étant (x_1, x_3, x_5).

On écrit le système sous la forme

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

On peut choisir librement les valeurs des variables libres (celles qui ne correspondent pas un pivot), $x_2 = s$, $x_4 = t$, où $s \in \mathbb{R}$, $t \in \mathbb{R}$. Le système a une infinité de solutions

$$\begin{array}{rcl} x_1 & = & 2 - 2s - 3t \\ x_2 & = & s \\ x_3 & = & 2 + t \\ x_4 & = & t \\ x_5 & = & -2 \end{array}$$

que l'on représente à l'aide du vecteur

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 - 2s - 3t \\ s \\ 2 + t \\ t \\ -2 \end{bmatrix}$$

On met aussi la solution sous la forme

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ -2 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

Par exemple lorsque $s = t = 0$, on a la solution particulière

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ -2 \end{bmatrix}$$

1.2.4 Méthode générale de résolution des systèmes linéaires

On commence par rappeler les étapes du calcul d'une forme réduite échelonnée par ligne.

Algorithme 1. [Algorithme de Gauss-Jordan]

Mets le curseur sur le coefficient le plus haut de la première colonne.

Etape 1. *Si la valeur du curseur est nulle, cherche une ligne (vers le bas) où le coefficient de la colonne du curseur est non nulle et échange cette ligne avec la ligne du curseur. S'il n'y en a pas, bouge le curseur vers la droite et recommence.*

Etape 2. *Divise la ligne du curseur par la valeur du curseur.*

Etape 3. *Elimine tous les coefficients de la colonne du curseur qui ne sont pas sur la ligne du curseur en soustrayant à chaque ligne des multiples adaptés de la ligne du curseur.*

Etape 4. *Bouge le curseur une ligne plus bas et une colonne vers la droite. Retourne à l'étape 1.*

A la fin de l'algorithme on obtient une forme réduite échelonnée par ligne de la matrice d'entrée.

On peut maintenant décrire une méthode générale de résolution des systèmes linéaires.

Définition 1.2.4. *Soit un système de n équations linéaires avec m inconnues*

$$\left|\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ \cdots & & \cdots & & \cdots & = & \cdots \\ a_{n1}x_1 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}\right|$$

La **matrice des coefficients**, *notée $A = [a_{i,j}]$, de ce système est de taille $n \times m$. La* **matrice augmentée du système**, *notée B, a n lignes et $m+1$ colonnes. Ses m premières colonnes sont les mêmes que celles de A, sa dernière colonne est le vecteur dont la $i^{\text{ième}}$ composante est b_i.*

La **méthode de résolution de Gauss-Jordan** d'un système linèaire consiste à prendre en entrée la matrice augmentée du système et à lui appliquer l'algorithme de Gauss Jordan. La matrice obtenue en sortie correspond à un nouveau système linéaire très simple qui a les mêmes solutions que le système linéaire de départ.

Vu que l'algorithme de Gauss-Jordan traite les colonnes de la gauche vers la droite on a le résultat suivant.

Proposition 1.2.1. *Soit A une matrice à n lignes et m colonnes, $m' < m$ et A' la matrice obtenue en prenant dans A les m' colonnes de gauche. Alors*

Figure 1.5: Portrait de Carl Friedrich Gauss (1777-1855)

$\mathrm{Frel}(A')$ *est la matrice obtenue en prenant dans* $\mathrm{Frel}(A)$ *les* m' *colonnes de gauche.*

En particulier si A est la matrice des coefficients d'un système linéaire et B est sa matrice augmentée, $\mathrm{Frel}(A)$ est la matrice obtenue en enlevant la dernière colonne de $\mathrm{Frel}(B)$.

1.3 Solution des systèmes linéaires, opérations sur les matrices

Dans cette section on se donne deux objectifs :

- Discuter le nombre de solutions d'un système linéaire et définir la notion de rang d'une matrice
- Présenter des règles de calcul sur les matrices et les vecteurs

1.3.1 Nombre de solutions d'un système linéaire, définition du rang

Exercice 1.3.1. *On considère trois systèmes dont les formes réduites échelonnées par ligne des matrices augmentées sont*

$$\mathbf{a.}\begin{bmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 1 \\ 0 & 0 & 0 & \vdots & 0 \end{bmatrix} \quad \mathbf{b.}\begin{bmatrix} 1 & 2 & 0 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 2 \\ 0 & 0 & 0 & \vdots & 0 \end{bmatrix} \quad \mathbf{c.}\begin{bmatrix} 1 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 0 & \vdots & 2 \\ 0 & 0 & 1 & \vdots & 3 \end{bmatrix}$$

Quel est le nombre de solutions dans chaque cas ?

Solution.

(a) La troisième ligne représente l'équation $0 = 1$, donc il n'y a pas de solution. On dit que le système est **inconsistant**.

(b) La matrice représente le système

$$\left| \begin{array}{ccccc} x_1 & + & 2x_2 & & = & 1 \\ & & & x_3 & = & 2 \end{array} \right| \quad \text{ou} \quad \left| \begin{array}{l} x_1 = 1 - 2x_2 \\ x_3 = 2 \end{array} \right|$$

la variable x_2 est **libre**, on peut lui attribuer la valeur arbitraire t, donc ce système admet une **infinité de solutions**

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1-2t \\ t \\ 2 \end{bmatrix}, t \in \mathbb{R}.$$

(c) Il n'y a pas de ligne représentant l'équation $0 = 1$, et il n'y a pas de variable libre, donc ce système admet **une et une seule solution** $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

□

Définition 1.3.1. *On dit qu'un système est* **consistant** *si il a au moins une solution, il est dit* **inconsistant** *si il n'a aucune solution.*

Théorème 1. [Nombre de solutions d'un système linéaire] *On note A la matrice des coefficients du système et B sa matrice augmentée.*

- *Un système linéaire est* **inconsistant** *si et seulement si* $\mathrm{Frel}(B)$ *contient la ligne* $[0 \cdots 0 \vdots 1]$, *qui représente l'équation* $0 = 1$.
- *Si un système linéaire est consistant, alors il a*
  - *soit* **exactement une solution**, *quand* $\mathrm{Frel}(A)$ *a un pivot par colonne,*
  - *soit* **une infinité de solution**, *quand une colonne de* $\mathrm{Frel}(A)$ *ne contient pas de pivot. Les variables correspondant à une colonne sans pivot sont appelées* **libres**.

Définition 1.3.2. [Rang] *Le* **rang** *d'une matrice A est le nombre de pivots dans la matrice* $\mathrm{Frel}(A)$, *où* $\mathrm{Frel}(A)$ *est la forme réduite échelonnée par ligne de la matrice A.*

Soit

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

alors

$$\mathrm{Rang}(A) = 2 \quad \text{car} \quad \mathrm{Frel}(A) = \begin{bmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Il faut noter que l'on a défini le rang d'une matrice et non celui d'un système ! Pour l'étude des systèmes on considère parfois la matrice des coefficients et parfois la matrice augmentée.

Définition 1.3.3. *Soit un système de n équations linéaires avec m inconnues avec pour matrice des coefficients A, de taille $n \times m$. On note $(x_1, ..., x_m)$ les variables. On dit que x_k est une* **variable avec pivot** *si la $k^{\text{ième}}$ colonne de* $\mathrm{Frel}(A)$ *contient un pivot.*

Théorème 2. *(a) On a toujours* $\mathrm{Rang}(A) \leq n$ *et* $\mathrm{Rang}(A) \leq m$.

(b) Si $\mathrm{Rang}(A) = n$*, alors le système est consistant*

(c) Si $\mathrm{Rang}(A) = m$*, alors le système a au plus une solution*

(d) Si $\mathrm{Rang}(A) < m$*, alors le système a soit une infinité de solutions soit aucune.*

Démonstration. (a) Par définition de la forme réduite échelonnée par ligne, il ne peut y avoir plus de pivot qu'il y a de lignes et de colonnes. Donc $\mathrm{Rang}(A) \leq n$ et $\mathrm{Rang}(A) \leq m$.

(b) On suppose que $\mathrm{Rang}(A) = n$. Alors chaque ligne de $\mathrm{Frel}(A)$ contient un pivot. Donc $\mathrm{Frel}(B)$, où B est la matrice augmentée, ne peut pas avoir la ligne $[0 \cdots 0 \vdots 1]$, en utilisant la proposition 1.2.1. Par conséquent, le système est consistant.

(c) Si $\mathrm{Rang}(A) = m$, on distingue plusieurs cas.

- Si $n = m$ on sait par (b) que le système est consistant. Il y a un pivot par colonne et on n'a donc aucune ligne $[0 \cdots 0 \vdots 1]$ dans la forme réduite de la matrice augmentée. Donc le système a une solution unique.

- $n > m$ Il y a deux cas possibles

  * On a une ligne (la $m + 1^{\text{ième}}$) $[0 \cdots 0 \vdots 1]$ dans $\mathrm{Frel}(B)$, où B est la matrice augmentée. Dans ce cas le système est inconsistant,

$*$ Les ligne i, pour $i > m$, sont de la forme $[0 \cdots 0 \vdots 0]$, on a

$$\mathrm{Frel}(B) = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 & \vdots & a_1 \\ 0 & 1 & 0 & . & . & . & 0 & \vdots & a_2 \\ 0 & 0 & 1 & . & . & . & 0 & \vdots & a_3 \\ . & . & . & . & . & . & . & \vdots & . \\ . & . & . & . & . & . & . & \vdots & . \\ . & . & . & . & . & . & . & \vdots & . \\ 0 & 0 & 0 & . & . & . & 1 & \vdots & a_m \\ 0 & 0 & 0 & . & . & . & 0 & \vdots & 0 \\ . & . & . & . & . & . & . & \vdots & . \\ 0 & 0 & 0 & . & . & . & 0 & \vdots & 0 \end{bmatrix}$$

et le système admet une solution unique.

– Notez qu'on ne peut pas avoir $n < m$, car dans ce cas $\mathrm{Rang}(A) \leq n < m$, ce qui contredit $\mathrm{Rang}(A) = m$.

(d) Si $\mathrm{Rang}(A) < m$, alors soit le système est inconsistant (lors qu'une des lignes de $\mathrm{Frel}(B)$, où B est la matrice augmentée, représente $0 = 1$), soit il y a $m - \mathrm{Rang}(A) > 0$ variables libres, et dans ce cas, le système a une infinité de solutions.

□

Problème 1.3.1. *Soit un système linéaire qui a moins d'équations que d'inconnues. Combien de solutions ce système peut-il avoir ?*

Théorème 3 (Système avec moins d'équations que d'inconnues)**.** *Un système qui a moins d'équations que d'inconnues a soit*

- *pas de solutions,*
- *une infinité de solutions.*

Démonstration. On suppose que $n < m$, et soit A la matrice des coefficients du système. Alors, on sait d'après (a) que $\mathrm{Rang}(A) \leq n < m$. Donc soit le système est inconsistant soit il y a $m - \mathrm{Rang}(A) > 0$ variables libres et on est dans le cas de (d) dans le Théorème 2. □

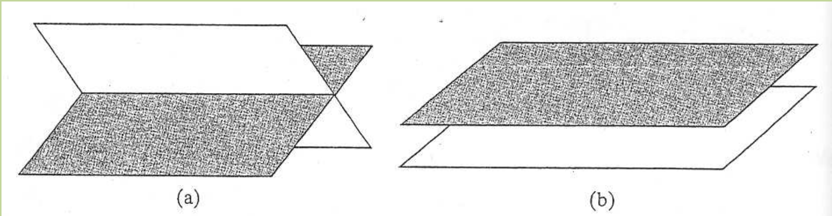


Figure 1.6: (a) deux plans qui s'intersectent en une droite, (b) deux plans parallèles

A titre d'illustration, considérons dans la Figure 1.6 un système de deux équations à trois inconnues. L'ensemble des solutions est soit une droite (intersection de deux plans non parallèles) soit l'ensemble vide (cas de deux plans parallèles)

Problème 1.3.2. *Considèrons un système linéaire de n équations à n inconnues. Quand est-ce que ce système admet une unique solution ?*

Théorème 4. *[Système de n équations avec n inconnues]*

Un système linéaire de n équations avec n inconnues admet une unique solution si et seulement si $\mathrm{Rang}(A) = n$. *Dans ce cas,*

$$\mathrm{Frel}(A) = \mathrm{I}_n = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 1 \end{bmatrix}$$

c'est-à-dire la matrice diagonale qui n'a que des 1 sur la diagonale (et des zéros partout ailleurs).

Démonstration. On suppose d'abord que $\mathrm{Rang}(A) = n$. Donc par définition, $\mathrm{Frel}(A)$ a n pivots égaux à 1, ce qui implique que $\mathrm{Frel}(A) = \mathrm{I}_n$. Alors B, la matrice augmentée du système admet pour forme réduite échelonnée par ligne une matrice de la forme

$$\left[\begin{array}{cccccccc|c} 1 & 0 & 0 & . & . & . & 0 & \vdots & a_1 \\ 0 & 1 & 0 & . & . & . & 0 & \vdots & a_2 \\ 0 & 0 & 1 & . & . & . & 0 & \vdots & a_3 \\ . & . & . & . & . & . & . & \vdots & . \\ . & . & . & . & . & . & . & \vdots & . \\ . & . & . & . & . & . & . & \vdots & . \\ 0 & 0 & 0 & . & . & . & 1 & \vdots & a_n \end{array}\right],$$

ce qui fait que le système admet une solution unique.

Si à présent $\text{Rang}(A) < n$, alors soit le système est inconsistant soit il est consistant et $\text{Frel}(A)$ a des colonnes sans pivot, donc des variables libres. Donc il a donc soit aucune solution soit une infinité de solutions. Dans ce cas, le système n'a jamais de solution unique.

Conclusion: lorsque $n = m$, le système a une solution unique si et seulement si $\text{Rang}(A) = n$. □

1.3.2 Calcul Matriciel

Définition 1.3.4 (Somme de matrices)**.** *Soient $A = [a_{ij}]$ et $B = [b_{ij}]$ deux matrices de taille $n \times m$. La matrice $A + B$ est la matrice de taille $n \times m$ dont le coefficient de la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne est égal à $a_{ij} + b_{ij}$.*

Autrement dit

$$A + B = [a_{ij} + b_{ij}].$$

Autrement dit :

$$\begin{bmatrix} a_{11} & . & . & . & a_{1m} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{n1} & . & . & . & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & . & . & . & b_{1m} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ b_{n1} & . & . & . & b_{nm} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11} + b_{11} & . & . & . & a_{1m} + b_{1m} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{n1} + b_{n1} & . & . & . & a_{nm} + b_{nm} \end{bmatrix}$$

Définition 1.3.5 (Produit d'une matrice par un scalaire)**.** *Soient $A = [a_{ij}]$ une matrices de taille $n \times m$ et k un nombre réel, aussi nommé scalaire dans ce contexte. La matrice kA est la matrice de taille $n \times m$ dont le coefficient de la $i^{ième}$ ligne et la $j^{ième}$ colonne est égal à ka_{ij}.*

Autrement dit

$$kA = [ka_{ij}],$$

$$k \begin{bmatrix} a_{11} & . & . & . & . & a_{1m} \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ a_{n1} & . & . & . & . & a_{nm} \end{bmatrix} = \begin{bmatrix} ka_{11} & . & . & . & . & ka_{1m} \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ ka_{n1} & . & . & . & . & ka_{nm} \end{bmatrix}$$

Example 1.3.1.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 3 & 1 \\ 5 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 8 & 5 & 4 \\ 9 & 8 & 5 \end{bmatrix}$$

$$3 \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ -3 & 9 \end{bmatrix}$$

On a pour but de définir le produit de matrice, ce qui n'est pas immédiat. On commence par le produit d'une matrice de taille $n \times m$ par un vecteur à m composantes.

Définition 1.3.6 (Produit matrice vecteur)**.** *Soit A une matrice de taille $n \times m$ dont les vecteurs colonnes sont les m vecteurs de $\mathbb{R}^n$ (i.e. ayant n composantes) notés $\overrightarrow{v}_1, ..., \overrightarrow{v}_m$, et soit $\overrightarrow{x} \in \mathbb{R}^m$ dont les composantes sont notées $x_1, ..., x_m$. Alors le produit $A\overrightarrow{x}$ est le vecteur de $\mathbb{R}^n$ défini par*

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & . & . & . & . & \overrightarrow{v}_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_m \end{bmatrix} = x_1 \overrightarrow{v}_1 + x_2 \overrightarrow{v}_2 + ... + x_m \overrightarrow{v}_m$$

Il faut noter que le produit est bien défini seulement si le nombre de colonnes de la matrice A coincide avec le nombre de composantes du vecteur $\overrightarrow{x}$.

Example 1.3.2.

$$\begin{bmatrix}1 & 0 & -1\\ 1 & 2 & 3\end{bmatrix}\begin{bmatrix}3\\ 1\\ 2\end{bmatrix} = 3\begin{bmatrix}1\\ 1\end{bmatrix} + 1\begin{bmatrix}0\\ 2\end{bmatrix} + 2\begin{bmatrix}-1\\ 3\end{bmatrix} = \begin{bmatrix}3\\ 3\end{bmatrix} + \begin{bmatrix}0\\ 2\end{bmatrix} + \begin{bmatrix}-2\\ 6\end{bmatrix} = \begin{bmatrix}1\\ 11\end{bmatrix}$$

Le produit

$$A\overrightarrow{x} = \begin{bmatrix}1 & 0 & -1\\ 1 & 2 & 3\end{bmatrix}\begin{bmatrix}3\\ 1\end{bmatrix}$$

n'est pas défini car le nombre de colonnes de la matrice n'est pas égal au nombre de composante du vecteur $\overrightarrow{x}$.

On considère la matrice identité 3×3,

$$\mathrm{I}_3 = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Soit $\overrightarrow{x} \in \mathbb{R}^3$. *Déterminer le produit* $\mathrm{I}_3\overrightarrow{x}$.

Par définition,

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = x_1\begin{bmatrix}1\\ 0\\ 0\end{bmatrix} + x_2\begin{bmatrix}0\\ 1\\ 0\end{bmatrix} + x_3\begin{bmatrix}0\\ 0\\ 1\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \overrightarrow{x}$$

Donc pour chaque vecteur $\overrightarrow{x} \in \mathbb{R}^3$, $\mathrm{I}_3\overrightarrow{x} = \overrightarrow{x}$. *C'est la raison de la terminologie "matrice identité".*

Notre définition du produit matriciel fait intervenir l'expression

$$x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m,$$

où les x_i sont des "scalaires" et les $\overrightarrow{v}_i$ des vecteurs de $\mathbb{R}^n$. de telles expressions sont très fréquentes en algèbre, et portent un nom.

Définition 1.3.7 (Combinaisons linéaires)**.** *On dit qu'un vecteur* $\overrightarrow{b} \in \mathbb{R}^n$ *est une combinaison linéaire des vecteurs* $\overrightarrow{v}_1, \overrightarrow{v}_2, ... \overrightarrow{v}_m$, *tous dans* $\mathbb{R}^n$, *si il existe des scalaires (i.e. des nombres réels)* $x_1, ..., x_m$ *tels que*

$$\overrightarrow{b} = x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m.$$

En particulier, le produit $A\overrightarrow{x}$ est la combinaison linéaire des vecteurs colonnes de la matrice A avec pour scalaires correspondants les coefficients de $\overrightarrow{x}$,

$$A\overrightarrow{x} = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & . & . & . & . & \overrightarrow{v}_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_m \end{bmatrix} = x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m$$

Dans le travail théorique, il est souvent utile de définir le produit $A\overrightarrow{x}$ comme la combinaison linéaire $x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m$. De même, on pourra être conduit à considérer la combinaison linéaire $x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m$ en introduisant la matrice

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & . & . & . & . & \overrightarrow{v}_m \end{bmatrix} \quad \text{et le vecteur} \quad \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_m \end{bmatrix}$$

et écrire $x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m = A\overrightarrow{x}$

Le produit matrice vecteur va être très utile pour l'étude des systèmes linéaires, comme le montre l'exemple suivant.

Example 1.3.3. *Soit le système*

$$\begin{vmatrix} 3x_1 & + & x_2 & = & 7 \\ x_1 & + & 2x_2 & = & 4 \end{vmatrix}$$

qui admet pour matrice augmentée

$$\begin{bmatrix} 3 & 1 & \vdots & 7 \\ 1 & 2 & \vdots & 4 \end{bmatrix}$$

La solution de ce système peut être vue comme l'intersection de deux droites dans le plan x_1, x_2, comme cela est illustré dans la Figure 1.7.

On peut également écrire le système sous la forme vectorielle

$$\begin{bmatrix} 3x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix},$$

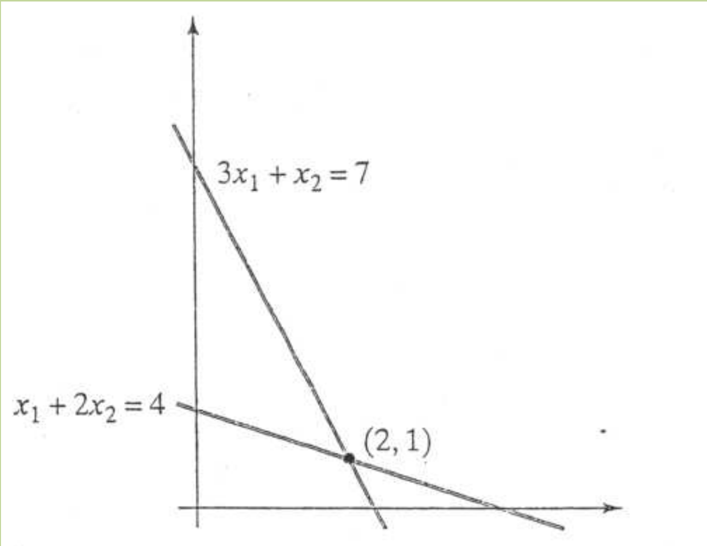


Figure 1.7: Représentation de la solution du système

ce qui peut encore s'écrire en terme de combinaisons linéaires sous la forme

$$x_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix},$$

et la solution peut s'interpréter comme définir le vecteur $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ *à l'aide de deux vecteurs paralèles aux vecteurs* $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ *et* $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *en formant un parallèlogramme, comme le montre la Figure 1.8.*

On peut écrire, en utilisant la définition du produit d'un vecteur par une matrice,

$$x_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

ce qui fait que notre système linéaire peut s'écrire sous la forme

$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix},$$

soit encore en posant

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad \overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \overrightarrow{b} = \begin{bmatrix} 7 \\ 4 \end{bmatrix},$$

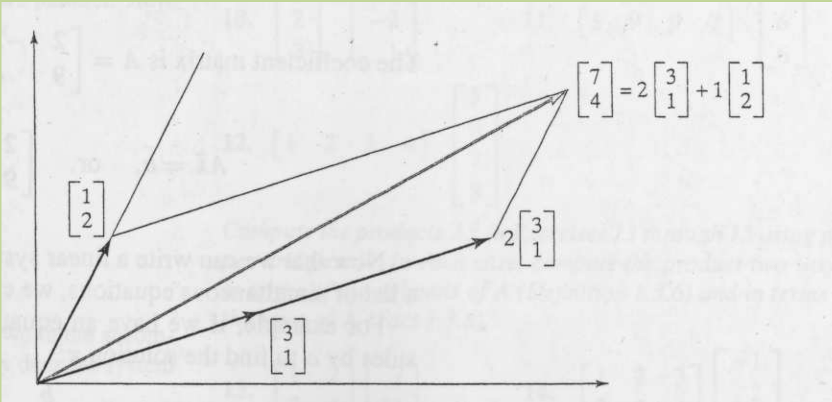


Figure 1.8: Représentation de la solution par un parallèlogramme

$$A\overrightarrow{x} = \overrightarrow{b}.$$

Et sa matrice augmentée se met sous la forme

$$[A \vdots \overrightarrow{b}] = \begin{bmatrix} 3 & 1 & \vdots & 7 \\ 1 & 2 & \vdots & 4 \end{bmatrix}$$

Définition 1.3.8 (Forme matricielle d'un système linéaire)**.** *Soit un système linéaire ayant pour matrice augmentée* $[A \vdots \overrightarrow{b}]$*. Alors ce système peut aussi s'écrire sous forme matricielle*

$$A\overrightarrow{x} = \overrightarrow{b}.$$

En effet, la $i^{\text{ième}}$ composante du vecteur $A\overrightarrow{x}$ est $a_{i1}x_1 + ...a_{im}x_m$, en utilisant les notations habituelles. Donc la $i^{\text{ième}}$ composante de l'équation $A\overrightarrow{x} = \overrightarrow{b}$ est $a_{i1}x_1 + ...a_{im}x_m = b_i$, ce qui correspond bien à la $i^{\text{ième}}$ équation du système qui a $[A \vdots \overrightarrow{b}]$ pour matrice augmentée.

Example 1.3.4. *Le système*

$$\left| \begin{array}{ccccc} 2x_1 & - & 3x_2 & + & 5x_3 = & 7 \\ 9x_1 & + & 4x_2 & - & 6x_3 = & 8 \end{array} \right|$$

a pour matrice des coefficients et second membre

$$A = \begin{bmatrix} 2 & -3 & 5 \\ 9 & 4 & -6 \end{bmatrix}, \quad \overrightarrow{b} = \begin{bmatrix} 7 \\ 8 \end{bmatrix},$$

et se met sous la forme $A\vec{x} = \vec{b}$. *Sa matrice augmentée est*

$$B = [A \vdots \vec{b}] = \begin{bmatrix} 2 & -3 & 5 & \vdots & 7 \\ 9 & 4 & -6 & \vdots & 8 \end{bmatrix}$$

Remarquons que l'on a travaillé avec les colonnes de la matrice dans la définition du produit matrice vecteur. On peut aussi travailler avec les lignes. Pour cela on commence par définir le produit d'un vecteur ligne par un vecteur colonne.

Définition 1.3.9 (Produit d'un vecteur ligne par un vecteur colonne)**.** *Soient* $\underset{\rightarrow}{\ell}$ *un vecteur ligne à* m *composantes et* $\vec{x} \in \mathbb{R}^m$, *de composantes* $\ell_1, ..., \ell_m$ *et* $x_1, ..., x_m$ *respectivement. Alors* $\underset{\rightarrow}{\ell} \cdot \vec{x} \in \mathbb{R}$ *est le nombre défini par*

$$\underset{\rightarrow}{\ell} \cdot \vec{x} = \begin{bmatrix} \ell_1 & \ell_2 & . & . & . & \ell_m \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_m \end{bmatrix} = \ell_1 x_1 + \ell_2 x_2 + ... + \ell_m x_m.$$

Example 1.3.5. *Revenons à l'exemple 1.3.2, ou on a calculé le produit* $A\vec{x}$ *où*

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 3 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

et

$$A\vec{x} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 2 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} -2 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix} =$$
$$\begin{bmatrix} 3 \cdot 1 + 1 \cdot 0 + 2 \cdot (-1) \\ 3 \cdot 1 + 1 \cdot 2 + 2 \cdot 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$$

On considère les vecteurs lignes de la matrice A,

$$\underset{\rightarrow}{\ell}_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}, \quad \underset{\rightarrow}{\ell}_2 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

Alors la première composante du vecteur $A\vec{x}$ *est le produit* $\underset{\rightarrow}{\ell}_1 \cdot \vec{x}$,

$$\underset{\rightarrow}{\ell}_1 \cdot \vec{x} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = 1 \cdot 3 + 0 \cdot 1 + (-1) \cdot 2 = 1$$

tandis que la deuxième composante est le produit

$$\underset{\rightarrow}{\ell}_2 \cdot \overrightarrow{x} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 + 3 \cdot 2 = 11$$

Proposition 1.3.1 (Le produit matrice vecteur à l'aide des lignes)**.** *Soit A une matrice de taille $n \times m$ avec $\underset{\rightarrow}{\ell}_1, ..., \underset{\rightarrow}{\ell}_n$ pour vecteurs lignes et soit $\overrightarrow{x} \in \mathbb{R}^m$. Alors on a*

$$A\overrightarrow{x} = \begin{bmatrix} \underset{\rightarrow}{\ell}_1 \cdot \overrightarrow{x} \\ \underset{\rightarrow}{\ell}_2 \cdot \overrightarrow{x} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\rightarrow}{\ell}_n \cdot \overrightarrow{x} \end{bmatrix}.$$

Démonstration. On note a_{ij} les coefficients de la matrice A et on désigne par $\overrightarrow{v}_1, ..., \overrightarrow{v}_m$ ses vecteurs colonnes, $\underset{\rightarrow}{\ell}_1, ..., \underset{\rightarrow}{\ell}_n$ ses vecteurs lignes. On remarque tout d'abord que pour j entre 1 et m, les composantes du vecteur $\overrightarrow{v}_j$ sont $a_{1j}, ..., a_{nj}$ et pour i entre 1 et n, les composantes du vecteur $\underset{\rightarrow}{\ell}_i$ sont $a_{i1}, a_{i2} ..., a_{im}$. En particulier la $i^{\text{ième}}$ composante du vecteur $\overrightarrow{v}_j$ est a_{ij}.

Par définition, on a

$$i^{\text{ième}}\text{composante de } A\overrightarrow{x} = i^{\text{ième}}\text{composante de } (x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + ... + x_m\overrightarrow{v}_m)$$

$$= x_1(i^{\text{ième}}\text{composante de } \overrightarrow{v}_1] + ... + x_m(i^{\text{ième}}\text{composante de } \overrightarrow{v}_m] =$$

$$x_1 a_{i1} + x_2 a_{i2} + ... + x_m a_{im} = \underset{\rightarrow}{\ell}_i \cdot \overrightarrow{x},$$

d'où le résultat annoncé. □

Example 1.3.6.

$$A\overrightarrow{x} = \begin{bmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 2 \cdot (-8) + 2 \cdot 4 \\ 1 \cdot 4 + 2 \cdot (-8) + 3 \cdot 4 \\ 4 \cdot 4 + 5 \cdot (-8) + 6 \cdot 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Il faut remarquer le fait important que le produit de A par $\overrightarrow{x}$ peut être nul, même lorsque ni A ni $\overrightarrow{x}$ le sont (ce qui n'arrive jamais avec des nombres réels, où le produit de deux réels non nuls n'est jamais nul). En terme de système linéaire, il existe des systèmes linéaires homogènes ayant des solutions non toutes nulles.

Définition 1.3.10. *Un système linéaire de n équations en m variables est* **homogène** *quand son second membre est $\overrightarrow{0} \in \mathbb{R}^n$.*

Proposition 1.3.2. *a) Un système homogène est toujours consistant.*

b) Un système homogène de n équations en m variables avec pour matrice des coefficients A à n lignes et m colonnes a une infinité de solutions si et seulement si $\text{Rang}(A) < m$ *, autrement dit s'il existe des variables libres.*

Démonstration. a) $\overrightarrow{0} \in \mathbb{R}^m$ vérifie $A\overrightarrow{0} = \overrightarrow{0}$ (on note de la même manière les vecteurs dont toutes les composantes sont nulles $\overrightarrow{0}$, dans $\mathbb{R}^n$ et dans $\mathbb{R}^m$).

b) Puisque le système est consistant il a une infinité de solutions si et seulement si $\text{Rang}(A) < m$ d'après le Théorème 1. □

Proposition 1.3.3 (Règles algébriques pour le calcul du produit matrice vecteur)**.** *Soit A une matrice de taille $n \times m$, $\overrightarrow{x} \in \mathbb{R}^m$ et $\overrightarrow{y} \in \mathbb{R}^m$ deux vecteurs, $k \in \mathbb{R}$ un scalaire. Alors on a*

(a) $A(\overrightarrow{x} + \overrightarrow{y}) = A\overrightarrow{x} + A\overrightarrow{y}$,

(b) $A(k\overrightarrow{x}) = k(A\overrightarrow{x})$

Démonstration. On désigne par $\underrightarrow{\ell}_i$ la $i^{\text{ième}}$ ligne de A. On a alors

$$i^{\text{ième}}\text{composante de } A(\overrightarrow{x} + \overrightarrow{y}) = \underrightarrow{\ell}_i \cdot (\overrightarrow{x} + \overrightarrow{y}) = \underrightarrow{\ell}_i \cdot \overrightarrow{x} + \underrightarrow{\ell}_i \cdot \overrightarrow{y} =$$

$$i^{\text{ième}}\text{composante de } A\overrightarrow{x} + i^{\text{ième}}\text{composante de } A\overrightarrow{y} =$$

$$i^{\text{ième}}\text{composante de } (A\overrightarrow{x} + A\overrightarrow{y}),$$

d'où le résultat. Le point (b) est laissé en exercice. □

Chapter 2

Applications linéaires

2.1 Introduction aux applications linéaires

Imaginez que vous êtes membre d'un bateau de gardes maritimes en méditeranée en train de rechercher des trafiquants (dangereux !). Périodiquement, vous envoyez par radio votre position au port de Marseille, position définie par

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

où x_1 est la longitude est, x_2 la latitude nord. Vous pensez que vos émissions sont captées par les trafiquants. Aussi, vous les codez par de nouvelles variables

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

en utilisant le code suivant

$$\begin{array}{rcrcr} y_1 & = & x_1 & + & 3x_2, \\ y_2 & = & 2x_1 & + & 5x_2. \end{array}$$

Par exemple, si votre position est 5° E, 42° N, ou bien encore

$$\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 42 \end{bmatrix},$$

et votre position codée est

$$\overrightarrow{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 131 \\ 220 \end{bmatrix},$$

Le codage peut être encore écrit sous la forme

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 \\ 2x_1 + 5x_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

soit encore

$$\boxed{\overrightarrow{y} = A\overrightarrow{x}}$$

où

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}.$$

Une "*application*" de la forme

$$\overrightarrow{x} \longrightarrow \overrightarrow{y} = A\overrightarrow{x}$$

est appelée une **application linéaire**. On rapellera le concept général "d'application" un peu plus loin.

Ce chapitre a pour objet l'étude détaillée de ce concept fondamental.

Lorsque le bateau atteint une nouvelle position, l'officier de service au port reçoit le message codé

$$\overrightarrow{b} = \begin{bmatrix} 133 \\ 223 \end{bmatrix}.$$

Il doit déterminer la position du bateau, et pour cela il doit résoudre le système

$$A\overrightarrow{x} = \overrightarrow{b},$$

c'est-à-dire

$$\left| \begin{array}{rcrcr} x_1 & + & 3x_2 & = & 133 \\ 2x_1 & + & 5x_2 & = & 223 \end{array} \right|,$$

qui a pour solution

$$\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 43 \end{bmatrix}.$$

Au bout du $20^{\text{ième}}$ système à résoudre, l'officier de service au port est un peu fatigué. Il se demande s'il ne peut pas résoudre le système général

$$\left| \begin{array}{rcrcr} x_1 & + & 3x_2 & = & y_1 \\ 2x_1 & + & 5x_2 & = & y_2 \end{array} \right|,$$

où y_1 et y_2 peuvent prendre n'importe quelle valeur. En clair, il voudrait coder l'application

$$\overrightarrow{y} \longrightarrow \overrightarrow{x},$$

qui est la "réciproque" ou l'"inverse" de l'application

$$\overrightarrow{x} \longrightarrow \overrightarrow{y} = A\overrightarrow{x}.$$

Pour résoudre cette question, il n'y a pas grand chose de nouveau. On utilise la méthode d'élimination de Gauss-Jordan

$$\left|\begin{array}{ccccccc} x_1 & + & 3x_2 & = & y_1 & & \\ 2x_1 & + & 5x_2 & = & & & y_2 \end{array}\right|$$

$$(\text{L2}) \leftarrow (\text{L2}) - 2(\text{L1})$$

$$\left|\begin{array}{ccccccc} x_1 & + & 3x_2 & = & y_1 & & \\ & - & x_2 & = & -2y_1 & + & y_2 \end{array}\right|$$

$$(\text{L2}) \leftarrow (\text{L2}) \div (-1)$$

$$\left|\begin{array}{ccccccc} x_1 & + & 3x_2 & = & y_1 & & \\ & & x_2 & = & 2y_1 & - & y_2 \end{array}\right|$$

$$(\text{L1}) \leftarrow (\text{L1}) - 3(\text{L2})$$

$$\left|\begin{array}{ccccccc} x_1 & & & = & -5y_1 & + & 3y_2 \\ & & x_2 & = & 2y_1 & - & y_2 \end{array}\right|$$

Donc le code est donné par les formules

$$\begin{array}{ccccc} x_1 & = & -5y_1 & + & 3y_2, \\ x_2 & = & 2y_1 & - & y_2, \end{array}$$

soit encore

$$\overrightarrow{x} = B\overrightarrow{y}, \quad B = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}.$$

On peut le revoir au niveau des matrices: on met la matrice de la transformation qu'on étudie à gauche et la matrice identité à droite et on utilise des transformations élémentaires sur les lignes pour faire apparaître la matrice identité à gauche: on a trouvé l'application linéaire réciproque, ou inverse.

$$\left[\begin{array}{cccccc} 1 & 3 & \vdots & 1 & 0 \\ 2 & 5 & \vdots & 0 & 1 \end{array}\right]$$

$$(\text{L2}) \leftarrow (\text{L2}) - 2(\text{L1})$$

$$\left[\begin{array}{cc:cc} 1 & 3 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{array}\right]$$

$$(\text{L2}) \leftarrow (\text{L2}) \div (-1)$$

$$\left[\begin{array}{cc:cc} 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & -1 \end{array}\right]$$

$$(\text{L1}) \leftarrow (\text{L1}) - 3(\text{L2}),$$

$$\left[\begin{array}{cc:cc} 1 & 0 & -5 & 3 \\ 0 & 1 & 2 & -1 \end{array}\right]$$

Par conséquent "l'application" $\overrightarrow{x} = B\overrightarrow{y}$ est la réciproque ou inverse de l'application $\overrightarrow{y} = A\overrightarrow{x}$. On dit que la matrice B est la matrice inverse de la matrice A et on pose $B = A^{-1}$.

Toutes les applications linéaires n'ont pas forcément de réciproque. Supposons qu'un officier un peu étourdi sur le bateau, choisisse le code

$$\begin{array}{rcrcr} y_1 & = & x_1 & + & 2x_2 \\ y_2 & = & 2x_1 & + & 4x_2 \end{array}, \quad \text{qui admet pour matrice} \quad A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix},$$

et lorsque l'officier de garde au port veux décoder la position, avec par exemple

$$\overrightarrow{b} = \begin{bmatrix} 89 \\ 178 \end{bmatrix},$$

il va découvrir avec un peu d'embarras que le système correspondant

$$\left| \begin{array}{rcrcr} x_1 & + & 2x_2 & = & 89 \\ 2x_1 & + & 4x_2 & = & 178 \end{array} \right|$$

admet une infinité de solutions,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 89 - 2t \\ t \end{bmatrix}, \quad t \in \mathbb{R}.$$

Puisque ce système n'admet pas une unique solution, il est impossible de déduire la position du bateau avec ce code. L'application linéaire sous-jacente n'admet pas de réciproque, et la matrice A n'est pas inversible. Ce code ne peut pas servir à communiquer la position du bateau !

On commence par rappeler le concept d'**application**.

Définition 2.1.1 (Applications)**.** *Soient X et Y deux ensembles. On dit qu'une relation T entre X et Y est une* **application** *de graphe $G_T \subset X \times Y$, lorsque tout élément de X possède exactement une image dans Y. On dit que X est l'ensemble de départ de T, Y son ensemble d'arrivée. On parle parfois de x comme l'entrée, et de y comme la sortie.*

Par le passé, vous avez étudié des applications pour lesquelles les entrées et les sorties sont des nombres réels, comme par exemple

$$y = x^2, \quad f(x) = e^x.$$

De telles applications sont souvent appelées fonctions.

Vous avez pu aussi croiser des applications pour lesquelles l'entrée ou la sortie est vectorielle, comme

$$y = x_1^2 + x_2^2 + x_3^2,$$

qui définit une application $\mathbb{R}^3 \to \mathbb{R}$, l'entrée étant le vecteur $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, et la sortie le nombre $y = x_1^2 + x_2^2 + x_3^2$.

La formule

$$\overrightarrow{r} = \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix},$$

formule définit une application de $\mathbb{R}$ dans $\mathbb{R}^3$ avec t en entrée et $\overrightarrow{r}$ en sortie.

Définition 2.1.2 (Applications linéaires)**.** *Une application $T : \mathbb{R}^m \to \mathbb{R}^n$ est une application linéaire si il existe une matrice A de taille $n \times m$, telle que pour tout $\overrightarrow{x} \in \mathbb{R}^m$.*

$$T(\overrightarrow{x}) = A\overrightarrow{x}.$$

Notez que le **nombre de colonnes** *de A est le nombre de composantes de l'***entrée** *et que le* **nombre de lignes** *de A est le nombre de composantes de la* **sortie***.*

Il est important de noter que les applications linéaires sont des cas particulier d'applications. Les entrées et les sorties sont des vecteurs. Si on note $\overrightarrow{y}$ le vecteur sortie $T(\overrightarrow{x})$, alors on peut écrire

$$\overrightarrow{y} = A\overrightarrow{x}.$$

Écrivons cette relation composante par composante.

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & . & . & . & . & a_{1m} \\ a_{21} & a_{22} & . & . & . & . & a_{2m} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & . & . & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ . \\ x_m \end{bmatrix},$$

ou bien encore

$$\begin{array}{ll} y_1 = & a_{11}x_1 + a_{12}x_2 + ... + a_{1m}x_m \\ y_2 = & a_{21}x_1 + a_{22}x_2 + ... + a_{2m}x_m \\ . & . \qquad\qquad\qquad\qquad . \\ . & . \qquad\qquad\qquad\qquad . \\ . & . \qquad\qquad\qquad\qquad . \\ y_n = & a_{n1}x_1 + a_{n2}x_2 + ... + a_{nm}x_m. \end{array}$$

Les variables de sortie y_i (il y en a autant que de lignes dans la matrice) dépendent linéairement des variables d'entrée x_j (il y en a autant que de colonnes de la matrice).

Example 2.1.1. *L'application linéaire qui va de* $\mathbb{R}^4$ *dans* $\mathbb{R}^3$*,*

$$\begin{array}{lllllllll} y_1 & = & 7x_1 & + & 3x_2 & - & 9x_3 & + & 8x_4, \\ y_2 & = & 6x_1 & + & 2x_2 & - & 8x_3 & + & 7x_4, \\ y_3 & = & 8x_1 & + & 4x_2 & & & + & 7x_4, \end{array}$$

admet pour matrice

$$\begin{bmatrix} 7 & 3 & -9 & 8 \\ 6 & 2 & -8 & 7 \\ 8 & 4 & 0 & 7 \end{bmatrix}.$$

Example 2.1.2 (Matrice identité)**.** *La transformation identité sur* $\mathbb{R}^n$*,*

$$\begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R}^n, \\ \overrightarrow{x} \longrightarrow \overrightarrow{x}, \end{cases}$$

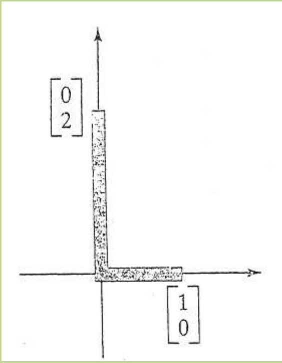


Figure 2.1: La lettre L

définie par les équations,

$$\begin{array}{rcl} y_1 & = & x_1 \\ y_2 & = & \quad x_2 \\ \cdot & & \qquad\quad \cdot \\ \cdot & & \qquad\qquad \cdot \\ \cdot & & \qquad\qquad\quad \cdot \\ \cdot & & \qquad\qquad\qquad \cdot \\ y_n & = & \qquad\qquad\qquad\qquad x_n \end{array}$$

admet pour matrice la matrice carré I_n*, de taille* $n \times n$ *(qui est diagonale et qui n'a que des* 1 *sur la diagonale). On rappelle que le système à* n *équations et* n *inconnues,* $A\overrightarrow{x} = \overrightarrow{b}$*, a une et une seule solution si et seulement si* $\mathrm{Frel}(A) = \mathrm{I}_n$*.*

Example 2.1.3. *Considérons la lettre L (voir Figure 2.1) constituée par les vecteurs* $\begin{bmatrix}1\\0\end{bmatrix}$ *et* $\begin{bmatrix}0\\2\end{bmatrix}$*. Décrire l'effet sur la lettre L de la transformation linéaire définie par*

$$T(\overrightarrow{x}) = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\overrightarrow{x},$$

et décrire cette transformation avec des mots simples.

Solution. On a

$$T\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} \quad \text{et} \quad T\begin{bmatrix}0\\2\end{bmatrix} = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}0\\2\end{bmatrix} = \begin{bmatrix}-2\\0\end{bmatrix},$$

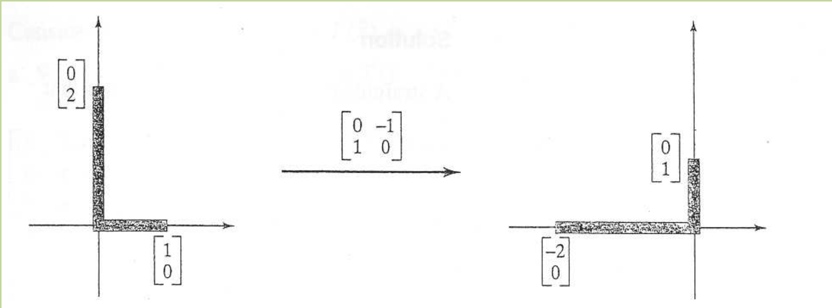


Figure 2.2: Image de L par l'application T

comme indiqué sur la figure ci-dessous. On voit que la lettre L a effectué une rotation de 90° dans le sens contraire des aiguilles d'une montre (ou sens trigonométrique).

On étudie l'effet de cette application linéaire sur un vecteur quelconque $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$:

$$T(\overrightarrow{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

On observe que les vecteurs $\overrightarrow{x}$ et $T\overrightarrow{x}$ ont la même longueur,

$$\sqrt{x_1^2 + x_2^2} = \sqrt{(-x_2)^2 + x_1^2},$$

et que de plus, ils sont orthogonaux puisque leur produit scalaire est égal à 0,

$$\overrightarrow{x} \cdot T(\overrightarrow{x}) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} = -x_1 x_2 + x_2 x_1 = 0.$$

En observant les signes des composantes, on remarque que lorsque que le vecteur $\overrightarrow{x}$ est dans le premier quadrant ($x_1 \geq 0$, $x_2 \geq 0$), alors $T(\overrightarrow{x})$ est dans le deuxième quadrant ($x_1 \leq 0$, $x_2 \geq 0$). On en déduit que $T(\overrightarrow{x})$ se déduit de $\overrightarrow{x}$ par une rotation de 90° dans le sens contraire des aiguilles d'une montre.

$\square$

Exercice 2.1.1. *On considère l'application linéaire de* $\mathbb{R}^3 \to \mathbb{R}^3$, $T(\overrightarrow{x}) = A\overrightarrow{x}$ *pour laquelle la matrice* A *est la matrice* 3×3

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

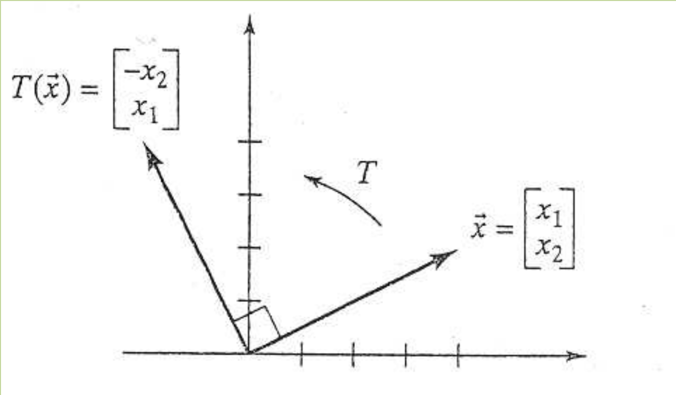


Figure 2.3: Rotation de 90° dans le sens contraire des aiguilles d'une monde

Déterminer les vecteurs

$$A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad et \quad A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Un calcul direct montre que

$$A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \quad A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}.$$

Il faut surtout noter que $A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ est le premier vecteur colonne de la matrice A et que $A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ est le troisième vecteur colonne de la matrice A. On généralise cette remarque dans le résultat suivant.

Proposition 2.1.1 (colonnes de la matrice d'une application linéaire.). *Soit*

$T : \mathbb{R}^m \to \mathbb{R}^n$ une application linéaire, A sa matrice. On pose

$$\overrightarrow{e}_j = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \leftarrow j^{\text{ème}} ligne \\ \\ \\ \end{array} \in \mathbb{R}^m.$$

Alors le $j^{\text{ème}}$ vecteur colonne de la matrice A est le vecteur $T(\overrightarrow{e}_j) \in \mathbb{R}^n$.

En d'autres termes, la matrice A peut être décrite comme suit :

$$A = \begin{bmatrix} | & | & & | \\ T(\overrightarrow{e}_1) & T(\overrightarrow{e}_2) & \cdots & T(\overrightarrow{e}_m) \\ | & | & & | \end{bmatrix}$$

Pour justifier ce résultat, on écrit A sous la forme

$$A = \begin{bmatrix} | & | & & | \\ \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m \\ | & | & & | \end{bmatrix}.$$

Alors on a

$$A\overrightarrow{e}_j = \begin{bmatrix} | & | & & | \\ \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} =$$

$$0\overrightarrow{v}_1 + 0\overrightarrow{v}_2 + ... + 0\overrightarrow{v}_{j-1} + 1\overrightarrow{v}_j + ... + 0\overrightarrow{v}_m = \overrightarrow{v}_j,$$

par définition du produit $A\overrightarrow{e}_j$, d'où le résultat.

La liste des vecteurs $\overrightarrow{e}_1, \overrightarrow{e}_1, ..., \overrightarrow{e}_m$ est souvent appelée **la base canonique de $\mathbb{R}^m$**. On expliquera plus tard dans le cours la notion abstraite de "base" dans $\mathbb{R}^m$. Le terme "base" se justifie par le fait qu'un vecteur $\overrightarrow{x} \in \mathbb{R}^m$ de composantes $x_1, x_2, \cdots, x_m$ peut s'écrire de manière unique sous forme d'une combinaison linéaire des $\overrightarrow{e}_i$. En effet, on vérifie facilement que

$$\overrightarrow{x} = x_1 \overrightarrow{e}_1 + x_2 \overrightarrow{e}_2 + \cdots + x_m \overrightarrow{e}_m.$$

Dans le cas particulier de la "dimension" 3, la base canonique de $\mathbb{R}^3$ est souvent notée $\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}$.

Problème 2.1.1. *On considère une application linéaire*

$$T : \left\{ \begin{array}{l} \mathbb{R}^m \longrightarrow \mathbb{R}^n, \\ \overrightarrow{x} \longrightarrow \overrightarrow{y} = A\overrightarrow{x}, \end{array} \right.$$

(a) Quelle est la relation entre $T(\overrightarrow{v})$, $T(\overrightarrow{w})$ et $T(\overrightarrow{v} + \overrightarrow{w})$, où $\overrightarrow{v}$ et $\overrightarrow{w}$ sont deux vecteurs quelconques de $\mathbb{R}^m$?

(b) Quelle est la relation entre $T(\overrightarrow{v})$ et $T(k\overrightarrow{v})$, où $\overrightarrow{v} \in \mathbb{R}^m$ est un vecteur quelconque, et $k \in \mathbb{R}$ un scalaire quelconque ?

Solution. (a) En utilisant les règles de calculs décrites dans le chapitre 1, on a

$$T(\overrightarrow{v} + \overrightarrow{w}) = A(\overrightarrow{v} + \overrightarrow{w}) = A\overrightarrow{v} + A\overrightarrow{w} = T(\overrightarrow{v}) + T(\overrightarrow{w}).$$

En d'autre termes, "l'image de la somme est la somme des image“s.

(b) Toujours avec les mêmes règles de calcul,

$$T(k\overrightarrow{v}) = A(k\overrightarrow{v}) = kA(\overrightarrow{v}) = kT(\overrightarrow{v}).$$

En d'autres termes, l'image du produit d'un vecteur par un scalaire est égal au produit du même scalaire par l'image du vecteur considéré. □

La Figure 2.4 illustre cet exemple dans le cas de la rotation d'angle 90° dans le sens trigonométrique (sens contraire des aiguilles d'une montre).

Dans le problme 2.1.1, on a vu qu'une application linéaire satisfait

$$\begin{array}{l} \forall \overrightarrow{v} \in \mathbb{R}^m, \quad \forall \overrightarrow{w} \in \mathbb{R}^m, \quad \forall k \in \mathbb{R}, \\ T(\overrightarrow{v} + \overrightarrow{w}) = T(\overrightarrow{v}) + T(\overrightarrow{w}), \quad T(k\overrightarrow{v}) = kT(\overrightarrow{v}). \end{array}$$

Il se pose la question de savoir si la réciproque est vraie, à savoir : une application qui satisfait ces propriétés est-elle bien une application linéaire ?

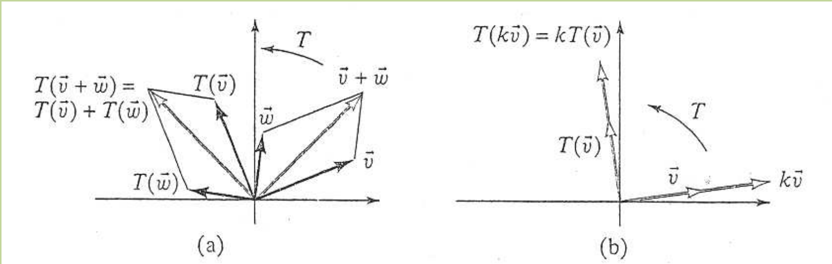


Figure 2.4:

Théorème 5 (Caractérisation des applications linéaires)**.** *Soit* $T : \mathbb{R}^m \to \mathbb{R}^n$ *une application. L'application* T *est linéaire (i.e. il existe une matrice* $n \times m$ A *telle que pour tout* $\overrightarrow{x} \in \mathbb{R}^m$*,* $T(\overrightarrow{x}) = A\ \overrightarrow{x}$*) si et seulement si*

(a) $\forall \overrightarrow{v} \in \mathbb{R}^m$, $\forall \overrightarrow{w} \in \mathbb{R}^m$ *on a* $T(\overrightarrow{v} + \overrightarrow{w}) = T(\overrightarrow{v}) + T(\overrightarrow{w})$

(b) $\forall \overrightarrow{v} \in \mathbb{R}^m$, $\forall k \in \mathbb{R}$ *on a* $T(k\overrightarrow{v}) = kT(\overrightarrow{v})$*.*

Démonstration. Dans l'exemple ??, on a vu qu'une application linéaire satisfait simultanément à (a) et (b). Pour montrer la réciproque, considérons une application $T : \mathbb{R}^m \to \mathbb{R}^n$ qui satisfait (a) et (b) et montrons qu'elle est linéaire. Pour cela, nous devons prouver qu'il existe une matrice A de taille $n \times m$ telle que

$$\forall \overrightarrow{x} \in \mathbb{R}^m, \quad T(\overrightarrow{x}) = A\overrightarrow{x}.$$

Soit $\overrightarrow{e}_1, \overrightarrow{e}_2, \cdots, \overrightarrow{e}_m$ la base canonique de $\mathbb{R}^m$ déjà introduite. On a alors

$$T(\overrightarrow{x}) = T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\right) = T(x_1\overrightarrow{e}_1 + x_2\overrightarrow{e}_2 + \cdots + x_m\overrightarrow{e}_m).$$

Or on a

$$T(\overrightarrow{x}) = T(x_1\overrightarrow{e}_1 + x_2\overrightarrow{e}_2 + \cdots + x_m\overrightarrow{e}_m) =$$

$$T(x_1\overrightarrow{e}_1) + T(x_2\overrightarrow{e}_2) + \cdots + T(x_m\overrightarrow{e}_m)$$

d'après la propriété **a**. En utilisant ensuite la propriété **b**, on obtient

$$T(\overrightarrow{x}) = x_1T(\overrightarrow{e}_1) + x_2T(\overrightarrow{e}_2) + \cdots + x_mT(\overrightarrow{e}_m).$$

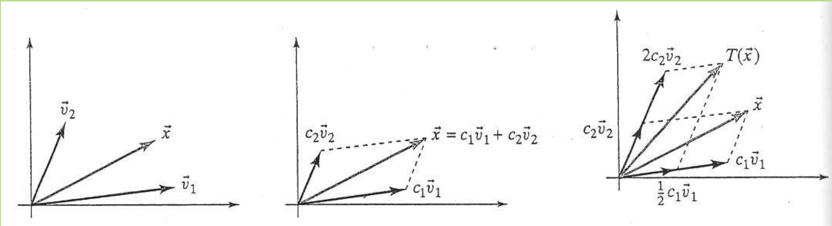


Figure 2.5: (a) (b) (c)

Enfin en utilisant la définition du produit d'une matrice par un vecteur, on voit d'après ce qui précède que

$$T(\overrightarrow{x}) = \begin{bmatrix} | & | & & | \\ T(\overrightarrow{e}_1) & T(\overrightarrow{e}_2) & \cdots & T(\overrightarrow{e}_m) \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = A\overrightarrow{x},$$

ce qui achève la démonstration.

Autrement dit, on construit A en prenant pour vecteurs colonnes les images des vecteurs $\overrightarrow{e}_i$ par l'application T. $\square$

Exercice 2.1.2. *On considère une application linéaire $T : \mathbb{R}^2 \to \mathbb{R}^2$ qui vérifie $T(\overrightarrow{v}_1) = (1/2)\overrightarrow{v}_1$ et $T(\overrightarrow{v}_2) = 2\overrightarrow{v}_2$, où les vecteurs $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ sont comme sur la figure 2.5. (a).*

Décrire le vecteur $T(\overrightarrow{x})$ pour un vecteur $\overrightarrow{x}$ quelconque.

Solution. En utilisant un paralèllogramme, on peut décomposer $\overrightarrow{x}$ comme étant une combinaison linéaire des vecteurs $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ (voir figure 2.5. (b)),

$$\overrightarrow{x} = c_1\overrightarrow{v}_1 + c_2\overrightarrow{v}_2.$$

La linéarité de T conduit à

$$T(\overrightarrow{x}) = T(c_1\overrightarrow{v}_1 + c_2\overrightarrow{v}_2) = c_1T(\overrightarrow{v}_1) + c_2T(\overrightarrow{v}_2) = \frac{1}{2}c_1\overrightarrow{v}_1 + 2c_2\overrightarrow{v}_2.$$

Le vecteur $c_1\overrightarrow{v}_1$ est diminué de moitié, $c_2\overrightarrow{v}_2$ est doublé (voir figure 2.5. (c)).

Imagine que $\overrightarrow{x}$ se trouve sur une feuille de caoutchouc. La transformation T dilate cette feuille d'un facteur 2 dans le sens du vecteur $\overrightarrow{v}_2$ et la contracte d'un facteur 1/2 dans la direction de $\overrightarrow{v}_1$. $\square$

Solution.

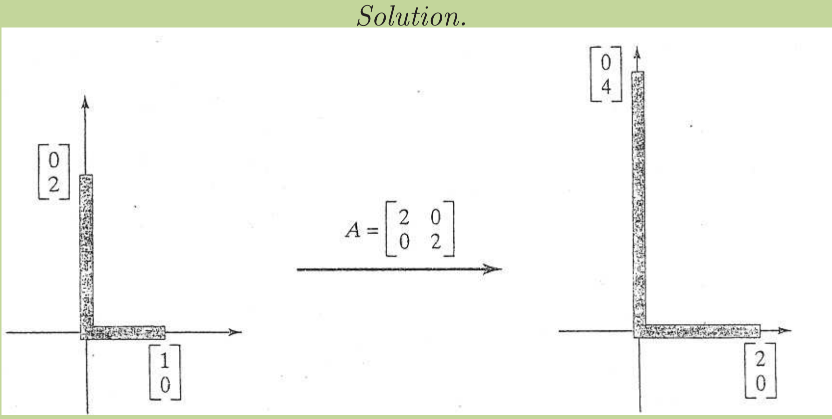


Figure 2.6: Matrice A

2.2 Les application linéaires en géométrie

Dans l'exemple 2.1.3, on a vu que la matrice $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$ représente une rotation de 90° dans le sens trigonométrique. De nombreuses matrices 2×2 représentent également des transformations géométriques simples.

Cette section a pour objet l'étude de quelques transformations géométriques, essentiellement dans le plan.

Problème 2.2.1. *On considère les matrices*

$$A = \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}, \quad B = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \quad C = \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix},$$

$$D = \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}, \quad E = \begin{bmatrix}1 & 0.5\\0 & 1\end{bmatrix}, \quad F = \begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}.$$

Décrire les effets de chacune de ces matrices sur notre lettre L *(en rappelant que son pied et le vecteur* $\begin{bmatrix}1\\0\end{bmatrix}$ *et son "dos" le vecteur* $\begin{bmatrix}0\\2\end{bmatrix}$*).*

Dans la Figure 2.6, notre lettre L est dilatée par un facteur 2. On appelle cette application **l'homothétie vectorielle** de rapport 2.

Dans la Figure 2.7, notre lettre L est aplatie, ou ratatinée, sur l'axe horizontal. On appelle cette transformation la **projection sur l'axe horizontal**.

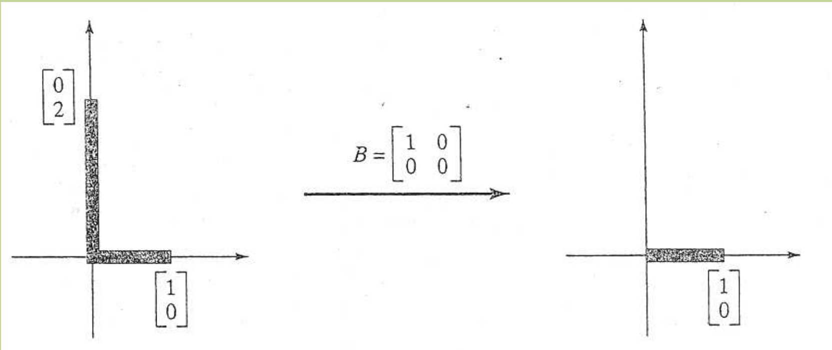


Figure 2.7: Matrice B

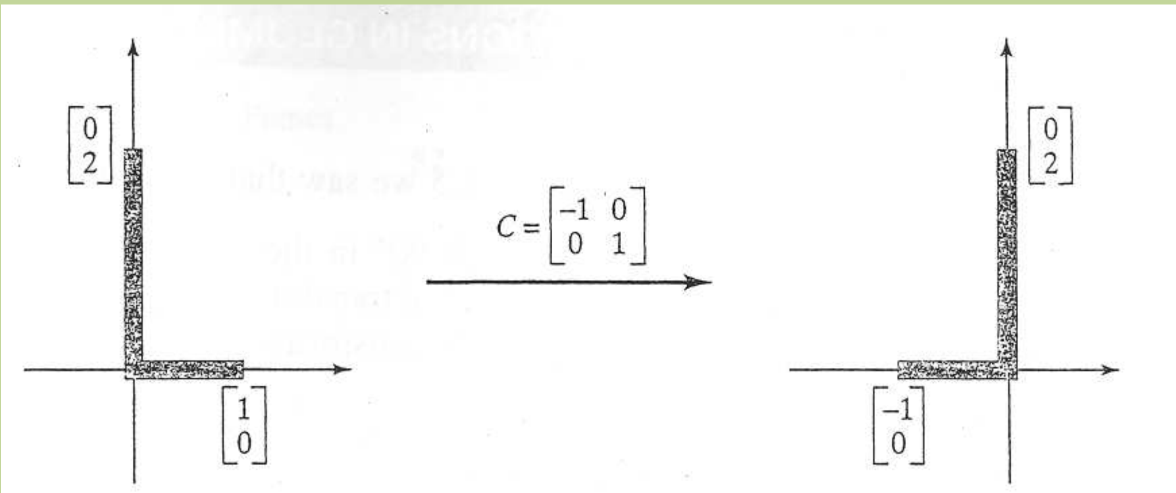


Figure 2.8: Matrice C

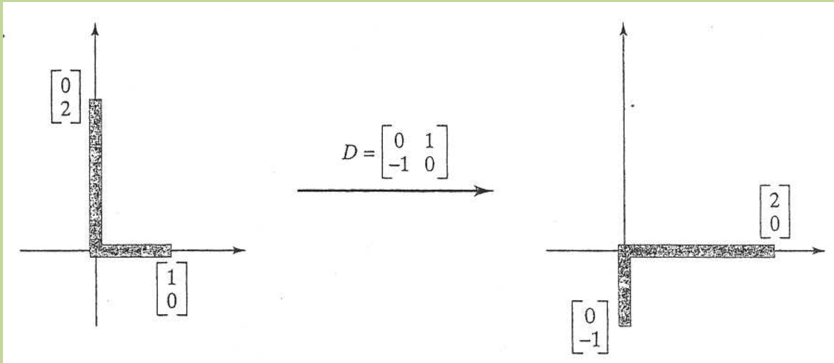


Figure 2.9: Matrice D

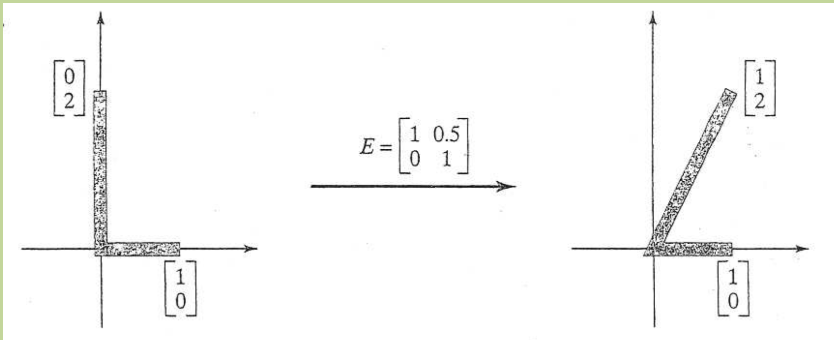


Figure 2.10: Matrice E

Dans la Figure 2.8, notre lettre L s'est retournée sur elle même par rapport à l'axe vertical. On appelle cette transformation la **symétrie par rapport à l'axe vertical**.

Dans la Figure 2.9, notre L a tourné de 90°, dans le sens des aiguilles d'une montre. Ce résultat est l'opposé de ce que nous avons eu dans l'exemple 2.1.3.

Dans la Figure 2.10, le pied du L n'a pas bougé, tandisque que le dos s'est penché horizontalement vers la droite, notre lettre devient *L* (penché). Cette transformation est appelée en Mathématiques une **transvection** horizontale.

Dans la Figure 2.11, d'une part la lettre L a tourné de 45° dans le sens trigonométrique, mais en plus a été dilatée d'un rapport de $\sqrt{2}$. Il sagit d'une **rotation composée par une homothétie vectorielle, ou**

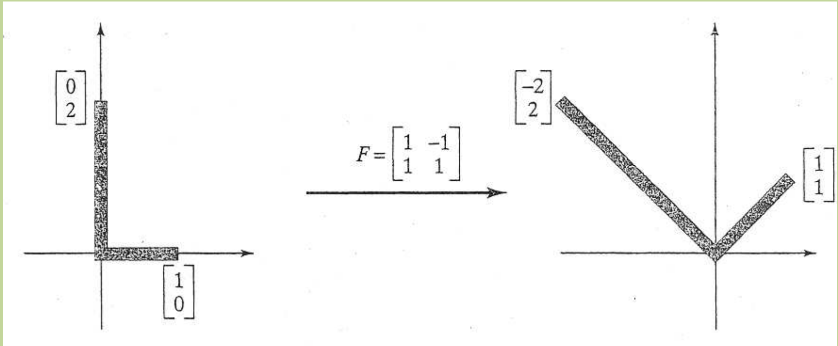


Figure 2.11: Matrice F

similitude. □

2.2.1 Homothéties vectorielles

Résumé 2.2.1. *Pour toute constante positive k, la matrice $H_k = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ définit une* **homothétie vectorielle**, *avec*

$$H_k \overrightarrow{x} = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \overrightarrow{x} = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} kx_1 \\ kx_2 \end{bmatrix} = k\overrightarrow{x}.$$

Lorsque $k > 1$, il s'agit d'une **dilatation**, *lorsque $k < 1$, on parle de* **contraction**.

On remarque que $H_1 = \mathrm{I}_2$, et que d'une manière générale, $H_k = k\mathrm{I}_2$.

On peut définir bien entendu une homothétie vectorielle pour $k < 0$. Dans ce cas, elle est la composée de l'homothétie de rapport positif $-k$ et de la symétrie centrale $\overrightarrow{x} \to -\overrightarrow{x}$.

2.2.2 Projections.

Considérons une droite L (pour "line" en anglais) dans le plan qui passe par l'origine. Soit $\overrightarrow{x}$ un vecteur quelconque. Il peut être décomposé de manière unique sous la forme

$$\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp},$$

où $\overrightarrow{x}^{||}$ est parallèle à la droite L et $\overrightarrow{x}^{\perp}$ est orthogonal à la droite L (voir figure 2.12 (a) ci-dessous)

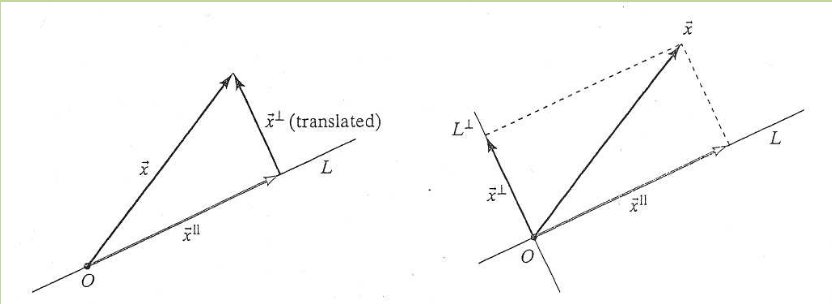


Figure 2.12: (a) (b)

L'application

$$\begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \\ \overrightarrow{x} \longrightarrow \overrightarrow{x}^{||}, \end{cases}$$

est appelée la projection de $\overrightarrow{x}$ sur la droite L, et notée

$$\mathrm{proj}_L(\overrightarrow{x}) = \overrightarrow{x}^{||}.$$

On peut voir $\mathrm{proj}_L(\overrightarrow{x})$ comme étant l'ombre de $\overrightarrow{x}$ projetée sur L que l'on éclaire directement.

Soit $L^{\perp}$ la droite orthogonale à L. On remarque que $\overrightarrow{x}^{\perp}$ est parallèle à $L^{\perp}$. Par conséquent, on peut interpréter $\overrightarrow{x}^{\perp}$ comme étant le projeté de $\overrightarrow{x}$ sur la droite $L^{\perp}$ (voir figure 2. 12 (b)).

On peut utiliser le produit scalaire pour déterminer une formule analytique qui décrit la projection. Pour cela, on considère un vecteur unitaire $\overrightarrow{u}$ parallèle à la droite L. Comme $\overrightarrow{x}^{||}$ est parallèle à L, il existe un scalaire $k \in \mathbb{R}$ tel que

$$\overrightarrow{x}^{||} = k\overrightarrow{u}.$$

Pour détereminer k, on remarque ensuite que

$$\overrightarrow{x}^{\perp} = \overrightarrow{x} - \overrightarrow{x}^{||} = \overrightarrow{x} - k\overrightarrow{u},$$

et que ce vecteur est perpendiculaire à la droite L. Par conséquent,

$$(\overrightarrow{x} - k\overrightarrow{u}) \cdot \overrightarrow{u} = 0,$$

d'où l'on déduit, puisque $\overrightarrow{u}$ est unitaire, $(\overrightarrow{u} \cdot \overrightarrow{u} = ||\overrightarrow{u}||^2 = 1)$

$$0 = \overrightarrow{x} \cdot \overrightarrow{u} - k(\overrightarrow{u} \cdot \overrightarrow{u}) = \overrightarrow{x} \cdot \overrightarrow{u} - k,$$

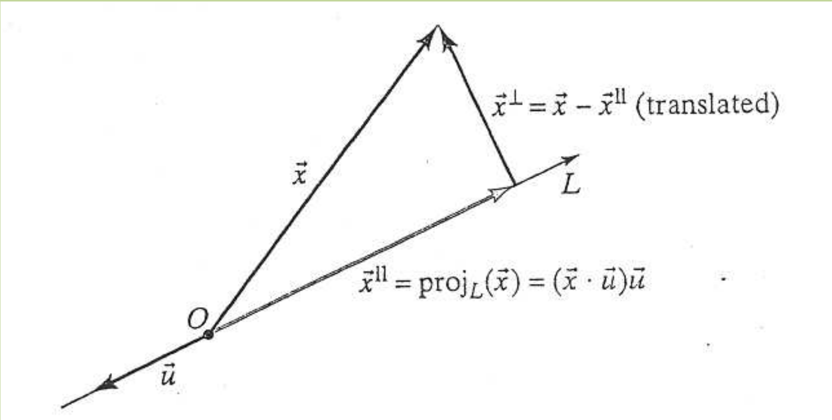


Figure 2.13: Description de la projection

c'est-à-dire

$$\boxed{\overrightarrow{x} \cdot \overrightarrow{u} = k.}$$

On en déduit finalement que

$$\boxed{\mathrm{proj}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} = k\overrightarrow{u} = (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u},}$$

ce qui est illustré par la figure 2. 13

Il se pose la question de savoir si cette projection est bien une application linéaire et quelle est sa matrice.

Écrivons

$$\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \overrightarrow{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Alors, on a

$$\mathrm{proj}_L(\overrightarrow{x}) = (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} = \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} =$$
$$(x_1u_1 + x_2u_2) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1^2x_1 + u_1u_2x_2 \\ u_1u_2x_1 + u_2^2x_2 \end{bmatrix} =$$
$$\begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix} \overrightarrow{x},$$

ce qui montre la linéarité de la projection. On notera que la matrice de proj_L est symétrique, c'est-à-dire que les termes qui sont au dessus et en dessous de la diagonale sont égaux.

Problème 2.2.2. *Trouver la matrice A de la projection sur la droite D qui a pour vecteur directeur le vecteur* $\overrightarrow{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

Solution. Il nous suffit juste de connaitre un vecteur directeur unitaire de la droite D, lequel est défini par

$$\overrightarrow{u} = \frac{\overrightarrow{v}}{||\overrightarrow{v}||} = \frac{1}{5}\overrightarrow{v} = \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}.$$

Il suffit à présent d'utiliser la formule précédente, avec $u_1 = 0.8$ et $u_2 = 0.6$, ce qui donne

$$A = \begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.48 \\ 0.48 & 0.36 \end{bmatrix}.$$

□

Résumé 2.2.2 (Projections.). *Soit L une droite dans le plan qui passe par l'origine. Chaque vecteur du plan $\overrightarrow{x}$ admet une unique décomposition*

$$\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp},$$

où $\overrightarrow{x}^{||}$ est parallèle à L et où $\overrightarrow{x}^{\perp}$ est orthogonal à L.

L'application

$$\begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \\ \overrightarrow{x} \longrightarrow T(\overrightarrow{x}) = \overrightarrow{x}^{||}, \end{cases}$$

est la projection sur la droite L, souvent notée proj_L.

Etant donné $\overrightarrow{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ *un vecteur directeur unitaire de la droite L, alors on a pour chaque vecteur* $\overrightarrow{x}$

$$\mathrm{proj}_L(\overrightarrow{x}) = (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

Enfin l'application $\overrightarrow{x} \to \mathrm{proj}_L(\overrightarrow{x})$ *est linéaire et admet pour matrice la matrice*

$$\begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix}.$$

Remarque 2.2.1. *La notion de projection que nous avons défini ici, est en fait une notion de* **projection orthogonale**. *Ce concept peut se généraliser à une classe plus vaste de projecteurs. Dans la suite, sauf mention du contraire, lorsque l'on parlera de projection, il sagira toujours d'une projection orthogonale comme elle l'a été définie plus haut.*

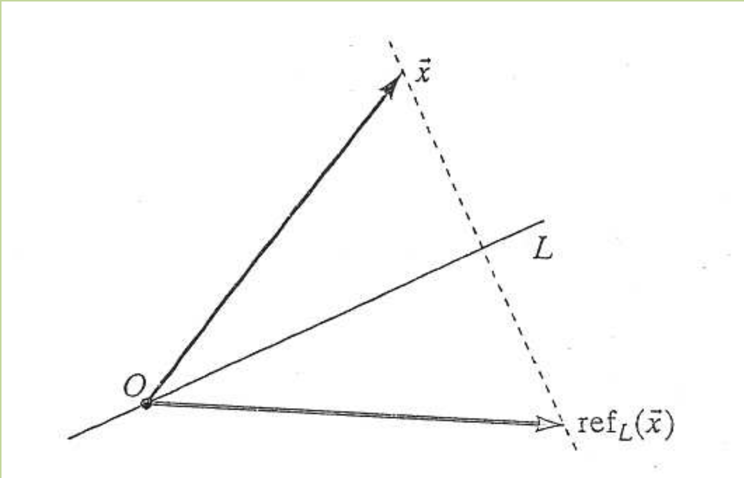


Figure 2.14: Symétrie par rapport à une droite

2.2.3 Symétries

On considère de nouveau une droite L (pour "line" en anglais) dans le plan qui passe par l'origine. Soit $\overrightarrow{x}$ un vecteur du plan. Le vecteur symétrique de $\overrightarrow{x}$ par rapport à L est représenté sur la figure ci-dessous, et on le note $\mathrm{sym}_L(\overrightarrow{x})$ (noté sur la figure 2.14 $\mathrm{ref}_L(\overrightarrow{x})$, pour "reflection" en anglais)

On fait "tourner" le vecteur $\overrightarrow{x}$ autour de la droite L.

Dans les cours d'analyse, vous avez certainement vu des exemples de symétrie par rapport à l'axe horizontal et l'axe vertical (en comparant les graphes de $y = f(x)$, $y = -f(x)$, $y = f(-x)$ par exemple).

On utilise à nouveau la décomposition $\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp}$ pour obtenir une fomule qui caractérise $\mathrm{sym}_L(\overrightarrow{x})$. On observe graphiquement que

$$\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} - \overrightarrow{x}^{\perp}.$$

On peut exprimer $\mathrm{sym}_L(\overrightarrow{x})$ à l'aide seulement de $\overrightarrow{x}^{\perp}$ ou bien encore de $\overrightarrow{x}^{||}$ comme suit :

$$\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} - \overrightarrow{x}^{\perp} = (\overrightarrow{x} - \overrightarrow{x}^{\perp}) - \overrightarrow{x}^{\perp} = \overrightarrow{x} - 2\overrightarrow{x}^{\perp},$$

et aussi

$$\begin{aligned} \mathrm{sym}_L(\overrightarrow{x}) &= \overrightarrow{x}^{||} - (\overrightarrow{x} - \overrightarrow{x}^{||}) &&= 2\overrightarrow{x}^{||} - \overrightarrow{x} \\ &= 2\mathrm{proj}_L(\overrightarrow{x}) - \overrightarrow{x} &&= 2(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} - \overrightarrow{x}. \end{aligned}$$

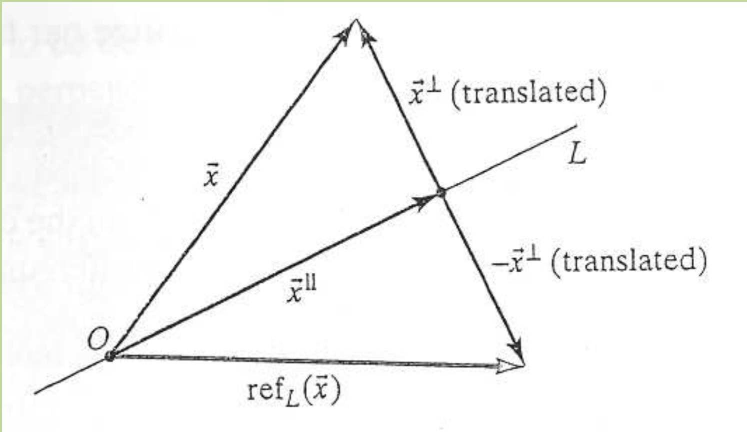


Figure 2.15:

Cette dernière formule garantit la linéarité de l'opérateur sym_L et nous permet d'en trouver sa matrice qui est

$$2\begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix} - \mathrm{I}_2 = \begin{bmatrix} 2u_1^2-1 & 2u_1u_2 \\ 2u_1u_2 & 2u_2^2-1 \end{bmatrix}.$$

En utilisant le fait que $\overrightarrow{u}$ soit unitaire, on remarque que l'égalité $u_1^2+u_2^2=1$, conduit à

$$\begin{array}{rcl} a \;= 2u_1^2-1 & = & 2u_1^2-(u_1^2+u_2^2) = u_1^2-u_2^2, \\ 2u_2^2-1 & = & 2u_2^2-(u_1^2+u_2^2) = u_2^2-u_1^2 = -a. \end{array}$$

Il en résulte que la matrice sym_L est de la forme

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix},$$

avec $a = u_1^2 - u_2^2, b = 2u_1u_2$ où u_1 et u_2 sont les coordonnées d'un vecteur unitaire de la droite L.

On remarque également que

$$\begin{array}{rcl} \underline{a^2+b^2} & = & (u_1^2-u_2^2)^2+4u_1^2u_2^2 = u_1^4-2u_1^2u_2^2+u_1^4+4u_1^2u_2^2 \\ & = & u_1^4+2u_1^2u_2^2+u_1^4 = (u_1^2+u_2^2)^2 \qquad\qquad = \underline{1}, \end{array}$$

toujours parce que $\overrightarrow{u}$ est unitaire.

On peut montrer le fait suivant (voir les exercices). Soit T une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qui admet une matrice de la forme

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix} \quad \text{avec} \quad a^2+b^2=1,$$

alors il existe une droite L telle que

$$T = \mathrm{sym}_L.$$

Résumé 2.2.3 (Symétries). *Soit L une droite dans le plan qui passe par l'origine. On décompose un vecteur du plan $\overrightarrow{x}$ sous la forme*

$$\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp},$$

où $\overrightarrow{x}^{||}$ est parallèle à L et où $\overrightarrow{x}^{\perp}$ est orthogonal à L. L'application linéaire $T(\overrightarrow{x}) = \overrightarrow{x}^{||} - \overrightarrow{x}^{\perp}$ est la symétrie par rapport à la droite L et est souvent notée sym_L *:*

$$\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} - \overrightarrow{x}^{\perp}.$$

La formule suivante relie $\mathrm{proj}_L(\overrightarrow{x})$ *et* $\mathrm{sym}_L(\overrightarrow{x})$ *:*

$$\mathrm{sym}_L(\overrightarrow{x}) = 2\mathrm{proj}_L(\overrightarrow{x}) - \overrightarrow{x} = 2\,(\overrightarrow{x}\cdot\overrightarrow{u})\overrightarrow{u} - \overrightarrow{x}.$$

La matrice de la symétrie est de la forme

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix} \quad \textit{avec} \quad a^2+b^2=1,$$

avec $a = u_1^2 - u_2^2, b = 2u_1u_2$ où u_1 et u_2 sont les coordonnées d'un vecteur unitaire de la droite L.

Remarque 2.2.2. *Ici, on a défini uniquement des symétries orthogonales. Ce concept peut se généraliser à des symétries non orthogonales. Dans ce cours, on ne considère que des symétries orthogonales.*

La figure 2.16. permet de comprendre la formule

$$\mathrm{sym}_L(\overrightarrow{x}) = 2\mathrm{proj}_L(\overrightarrow{x}) - \overrightarrow{x}.$$

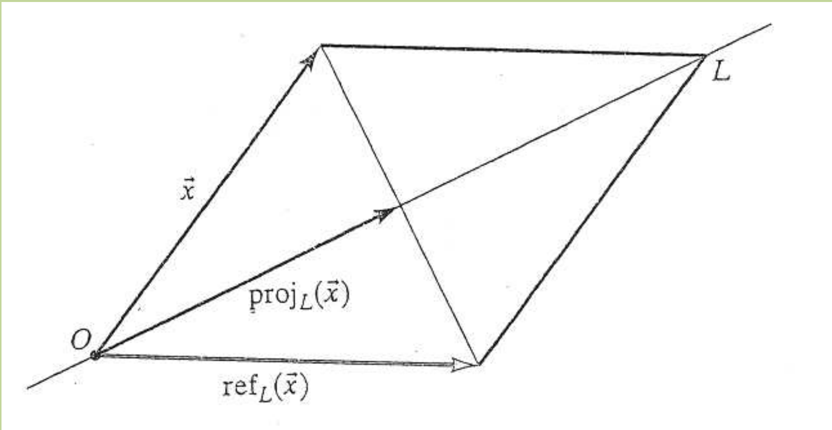


Figure 2.16: Symétrie avec un losange

2.2.4 Projections et Symétrie dans l'espace

Bien que le chapitre soit consacré au cas de la dimension 2, on va étudier rapidement les symétries et les projections dans le cas de la dimension 3, dans la mesure où la théorie est la même.

Soit L une droite dans l'espace qui passe par l'origine. On peut encore décomposer un vecteur $\overrightarrow{x}$ de l'espace de manière unique sous la forme

$$\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp},$$

où $\overrightarrow{x}^{||}$ est parallèle à L et où $\overrightarrow{x}^{\perp}$ est orthogonal à L. On définit encore

$$\mathrm{proj}_L(\overrightarrow{x}) = \overrightarrow{x}^{||},$$

et on a la formule

$$\mathrm{proj}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} = (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u},$$

où $\overrightarrow{u}$ est un vecteur directeur unitaire de L.

On définit la symétrie par rapport à L comme

$$\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x}^{||} - \overrightarrow{x}^{\perp},$$

et on a la formule

$$\mathrm{sym}_L(\overrightarrow{x}) = 2\mathrm{proj}_L(\overrightarrow{x}) - \overrightarrow{x} = 2(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} - \overrightarrow{x}$$

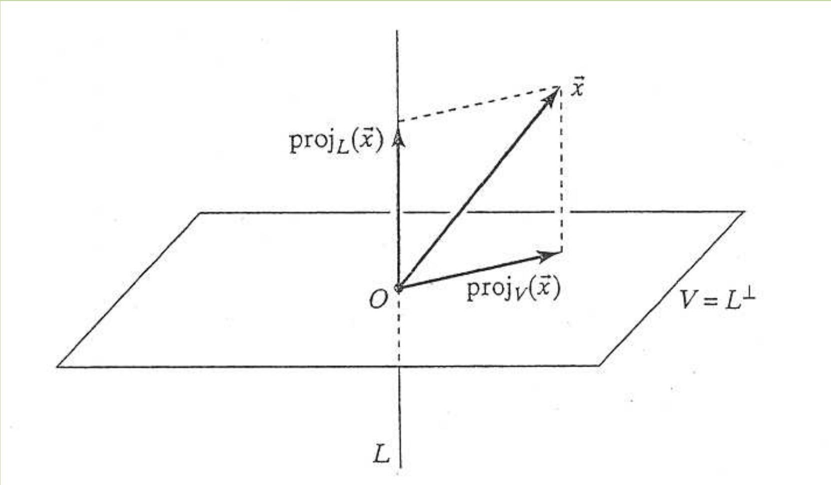


Figure 2.17: Projection dans l'espace

Soit $V = L^\perp$ le plan qui passe par l'origine et qui est perpendiculaire à la droite L. On note que $\overrightarrow{x}^\perp$ est parallèle au plan V, ce qui permet de définir la projection orthogonale sur le plan V comme étant

$$\mathrm{proj}_V(\overrightarrow{x}) = \overrightarrow{x}^\perp.$$

On peut déterminer analytiquement $\mathrm{proj}_V(\overrightarrow{x})$. En effet, on a toujours

$$\overrightarrow{x} = \mathrm{proj}_V(\overrightarrow{x}) + \mathrm{proj}_L(\overrightarrow{x}),$$

d'où

$$\mathrm{proj}_V(\overrightarrow{x}) = \overrightarrow{x} - \mathrm{proj}_L(\overrightarrow{x}) = \overrightarrow{x} - (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

De même pour la symétrie, on définit

$$\begin{aligned}\mathrm{sym}_L(\overrightarrow{x}) = \mathrm{proj}_L(\overrightarrow{x}) - \mathrm{proj}_V(\overrightarrow{x}) &= 2\mathrm{proj}_L(\overrightarrow{x}) - \overrightarrow{x}\\ &= 2(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} - \overrightarrow{x}.\end{aligned}$$

et on a

$$\mathrm{sym}_V(\overrightarrow{x}) = \mathrm{proj}_V(\overrightarrow{x}) - \mathrm{proj}_L(\overrightarrow{x}) = -\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x} - 2(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

Résumé 2.2.4 (Projections et symétries dans l'espace)**.** *On a les formules, en fonction du vecteur unitaire $\overrightarrow{u}$ de la droite L*

Projection sur la droite L

$$\mathrm{proj}_L(\overrightarrow{x}) = (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

Symétrie par rapport à L

$$\mathrm{sym}_L(\overrightarrow{x}) = 2\,(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} - \overrightarrow{x}$$

Projection sur le plan V

$$\mathrm{proj}_V(\overrightarrow{x}) = \overrightarrow{x} - (\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

Symétrie par rapport à V

$$\mathrm{sym}_V(\overrightarrow{x}) = -\mathrm{sym}_L(\overrightarrow{x}) = \overrightarrow{x} - 2\,(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u}.$$

Exercice 2.2.1. *Soit $V \subset \mathbb{R}^3$ le plan admettant pour équation cartésienne $2x_1 + x_2 - 2x_3 = 0$ et soit $\overrightarrow{x} = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}$. Déterminer $\mathrm{sym}_V(\overrightarrow{x})$.*

Solution. Le vecteur $\overrightarrow{v} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ est un vecteur orthogonal à V. Par conséquent,

$$\overrightarrow{u} = \frac{1}{||\overrightarrow{v}||}\overrightarrow{v} = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

est un vecteur unitaire perpendiculaire à V. On peut utiliser les formules déterminées précédement.

$$\begin{aligned}
\mathrm{sym}_V(\overrightarrow{x}) = \overrightarrow{x} - 2\,(\overrightarrow{x} \cdot \overrightarrow{u})\overrightarrow{u} &= \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} - \frac{2}{9}\left(\begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}\right)\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} - 4\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \\ -8 \end{bmatrix} \\
&= \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}
\end{aligned}$$

□

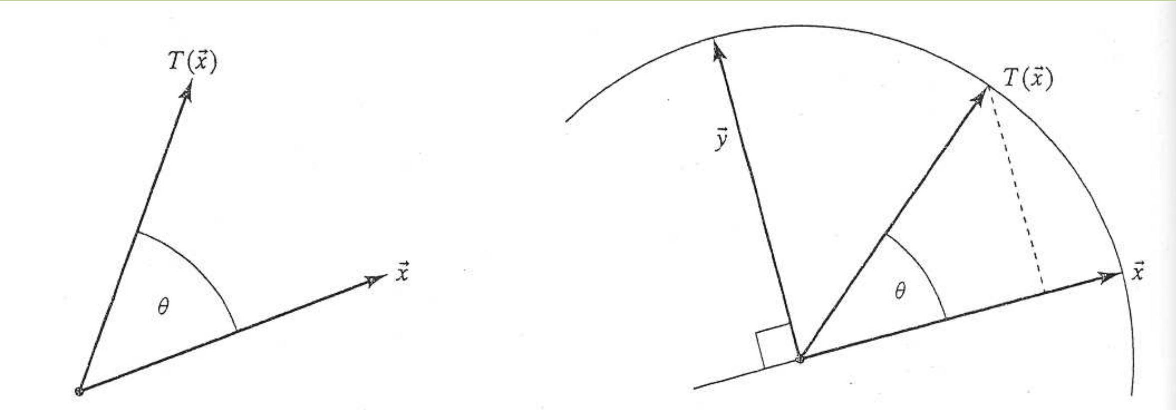


Figure 2.18: Rotation d'angle θ

2.2.5 Rotations.

On considère l'application $T : \mathbb{R}^2 \to \mathbb{R}^2$ qui fait tourner un vecteur $\overrightarrow{x}$ quelconque dans le sens trigonométrique d'un angle θ fixé, comme on le montre dans la figure ci-dessous.

Notons que dans l'exemple 2.1.3, on a étudié le cas $\theta = \pi/2$.

On rappelle la notion de coordonnées polaires.

Définition 2.2.1 (Coordonnées polaire.). *A un point $(x, y) \in \mathbb{R}^2$ on associe un couple (r, ρ) où r est un réel positif ou nul et ρ un angle entre 0 ret 2π tels que*

$$x = r\cos(\rho)$$

$$y = r\sin(\rho)$$

Notons que si $\overrightarrow{x}$ a pour coordonnées polaires (r, ρ), alors $\overrightarrow{t} = T(\overrightarrow{x})$ a pour coordonnée polaires $(r, \rho + \theta)$.

On rappelle les formules trigonométriques pour la somme des angles

$$\cos(\rho + \theta) = \cos(\rho)\cos(\theta) - \sin(\rho)\sin(\theta)$$

$$\sin(\rho + \theta) = \cos(\rho)\sin(\theta) + \sin(\rho)\cos(\theta)$$

Déterminons la matrice de T. On pose $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r\cos(\rho) \\ r\sin(\rho) \end{bmatrix}$, et on introduit le vecteur $\overrightarrow{y}$ déduit de $\overrightarrow{x}$ par une rotation d'angle θ. On a

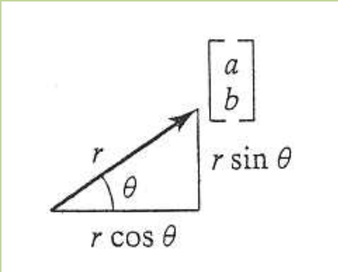


Figure 2.19: Coordonnées polaires

alors $\overrightarrow{y} = \begin{bmatrix} r\cos(\rho+\theta) \\ r\sin(\rho+\theta) \end{bmatrix}$. En utilisant les formules trigonométriques pour la somme des angles, on a

$$\begin{aligned} T(\overrightarrow{x}) = \begin{bmatrix} r\cos(\rho+\theta) \\ r\sin(\rho+\theta) \end{bmatrix} &= \begin{bmatrix} r\cos(\rho)\cos(\theta) - r\sin(\rho)\sin(\theta) \\ r\cos(\rho)\sin(\theta) + r\sin(\rho)\cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta)x_1 - \sin(\theta)x_2 \\ \sin(\theta)x_1 + \cos(\theta)x_2 \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

Le calcul qui précède montre que T est bien linéaire et admet pour matrice la matrice

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Résumé 2.2.5 (Rotation)**.** *La matrice d'une rotation dans* $\mathbb{R}^2$ *d'angle* θ *est la matrice*

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix},$$

qui est de la forme

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

avec $a^2 + b^2 = 1$.

Par exemple, la matrice de la rotation d'angle $\pi/6$ ($= 30°$) est la matrice

$$\begin{bmatrix} \cos(\pi/6) & -\sin(\pi/6) \\ \sin(\pi/6) & \cos(\pi/6) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$$

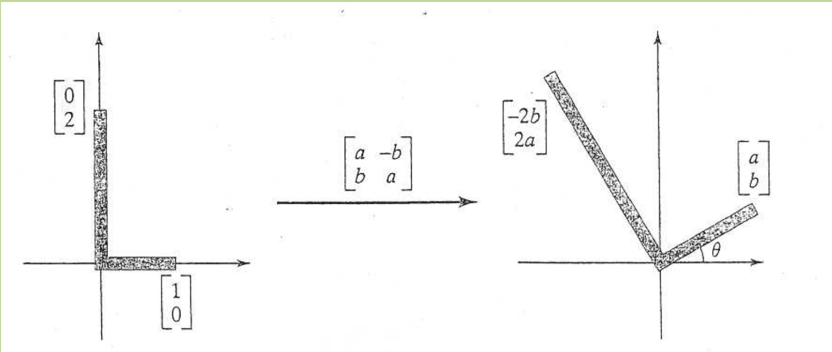


Figure 2.20: Similitude

2.2.6 Rotations composées avec des homothéties: similitudes

Exercice 2.2.2. *Soient a et b deux nombres réels quelconques. Comment l'application linéaire sur $\mathbb{R}^2$ définie par*

$$T(\overrightarrow{x}) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \overrightarrow{x},$$

agit-elle sur la lettre L *?*

Solution. La figure ci-dessous suggère que T est la composée d'une rotation avec une homothétie, aussi appelée similitude.

On raisonne en utilisant les coordonnées polaires : il sagit d'une rotation dont l'angle est l'angle de phase du vecteur $\begin{bmatrix} a \\ b \end{bmatrix}$, composée avec l'homothétie de rapport sa norme $r = \sqrt{a^2 + b^2}$.

On vérifie cette affirmation en écrivant $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}$,

On a alors

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos(\theta) & -r\sin(\theta) \\ r\sin(\theta) & r\cos(\theta) \end{bmatrix} = r \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

□

Résumé 2.2.6 (Rotation composée avec une homothétie: similitude)**.** *Une matrice de la forme* $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ *est la matrice d'une rotation composée avec une homothétie.*

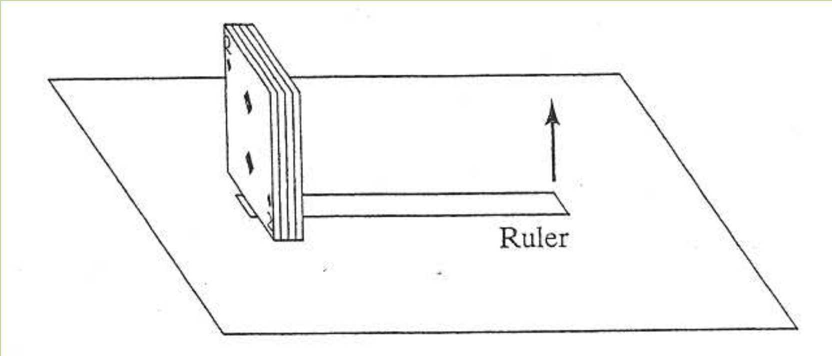


Figure 2.21: Un jeu de cartes sur une règle

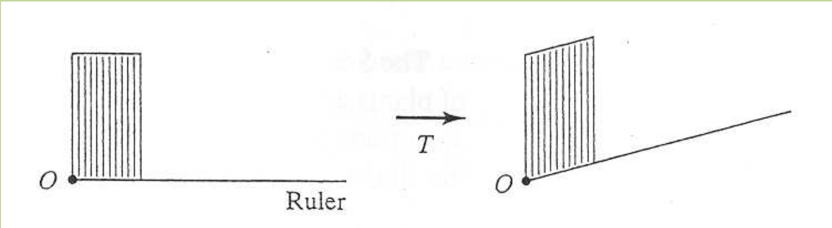


Figure 2.22: Déformation du jeu de cartes

On désigne par r et θ les coordonnées polaires du vecteur $\begin{bmatrix} a \\ b \end{bmatrix}$, alors l'angle de la rotation est de mesure θ et r est le rapport de l'homothétie.

2.2.7 Transvections

On introduit les transvections avec une expérience qui utilise un jeu de carte et une règle.

Les cartes sont maintenues en position verticale, le bord de la règle sur lequel repose le 2 de carreau est fixe, et on soulève l'autre bout de la règle.

La figure 2.22. montre l'expérience vue sur la tranche du jeu de cartes.

Ce type de transformation est appelée *transvection verticale*. L'expérience est dans l'espace, mais si on ne considère que la vue sur la tranche, on est conduit à étudier une transformation dans le plan.

On trace un vecteur $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ sur la tranche du jeu de cartes et on

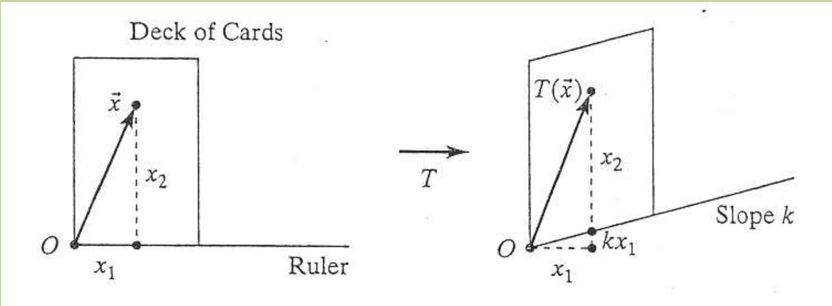


Figure 2.23: Déformation de la tranche du jeu de cartes

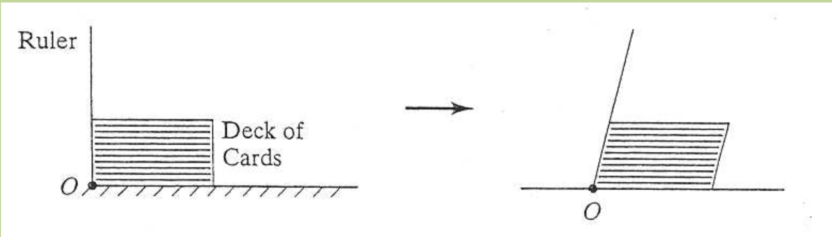


Figure 2.24: Transvection horizontale

cherche une formule analytique pour déterminer $T(\overrightarrow{x})$, en se guidant avec la figure ci dessous.

Dans ce qui suit, k désigne la pente de la règle après l'expérience.

On a

$$T(\overrightarrow{x}) = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ kx_1 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \overrightarrow{x}.$$

On en déduit que la matrice de T est une matrice de la forme $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ où k est une constante réelle.

On peut aussi définir les transvections horizontales, comme sur la figure 2.24

On laisse en exercice la vérification que la matrice d'une transvection horizontale est de la forme $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$.

Les transvections obliques sont également très importants dans les applications, mais on ne les étudie pas pour le moment.

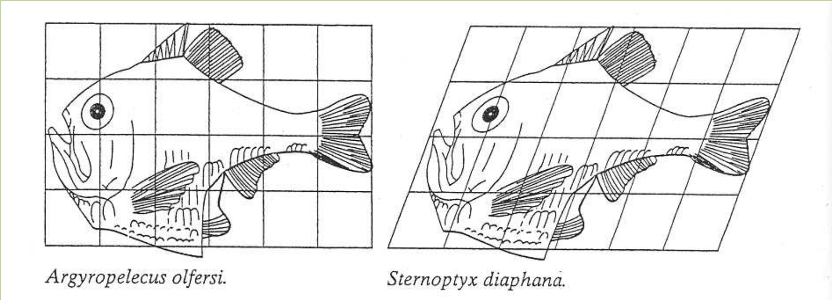


Figure 2.25: Transvection horizontale sur des poissons

Résumé 2.2.7 (Transvection verticale et horizontale)**.** *La matrice d'une transvection verticale est de la forme* $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$*, et la matrice d'une transvection horizontale est de la forme* $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$*, où* $k \in \mathbb{R}$ *est u constante arbitraire.*

Des chercheurs du collège écossais d'Arcy, ont montré comment on peut déduire la forme d'une espèce de celle d'une espèce proche, que ce soit pour les animaux ou les plantes, en utilisant des applications linéaires aussi bien de des transformations non linéaires.

Dans la figure 2.25. on utilise un transvection horizontale pour transformer la forme d'une espèce de poisson en une autre.

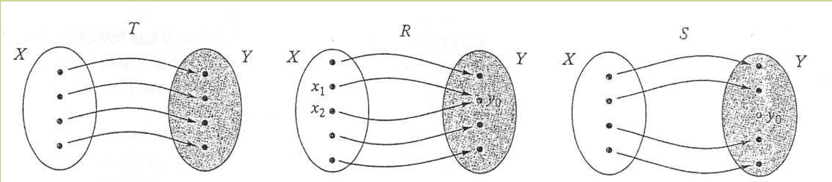


Figure 2.26: Exemples d'applications

2.3 Réciproque d'une application linéaire

On commence par rappeler le concept d'application inversible.

Définition 2.3.1 (Fonctions inversibles)**.** *Une application $T : X \to Y$ est dite inversible si, pour tout $y \in Y$, l'équation $T(x) = y$ admet une unique solution $x \in X$.*

Dans la figure 2.26.

- T est inversible ,
- R n'est pas inversible car l'équation $R(x) = y_0 \in Y$ admet deux solutions $x_1 \in X$ et $x_2 \in X$,
- S n'est pas inversible car il n'y a aucun $x \in X$ qui satisfasse $S(x) = y_0$.

Définition 2.3.2 (Réciproque, ou inverse)**.** *Soit $T : X \to Y$ inversible. Alors sa* **réciproque** *(ou inverse) $T^{-1} : Y \to X$ est définie par*

$$T^{-1}(y) = (\textit{l'unique } x \in X \textit{tel que } T(x) = y).$$

Autrement dit,

$$T^{-1}(y) = x \quad \textit{est équivalent à} \quad T(x) = y.$$

On notera que T^{-1} est une application. On a toujours

$$\forall\, x \in X, \quad T^{-1}(T(x)) = x, \quad (T^{-1} \circ T = \mathrm{Id}_X)$$

tandis que

$$\forall\, y \in Y, \quad T(T^{-1}(y)) = y, \quad (T \circ T^{-1} = \mathrm{Id}_Y)$$

Donc T^{-1} est inversible et on a l'identité

$$(T^{-1})^{-1} = T.$$

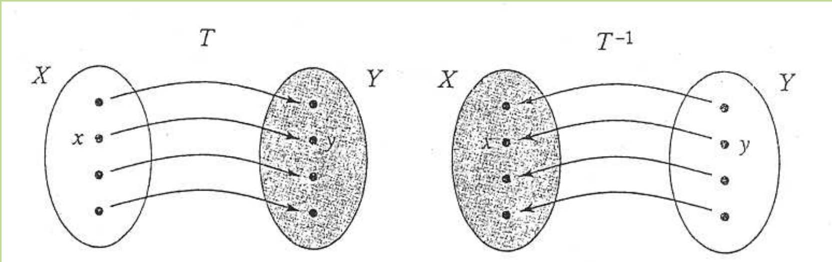


Figure 2.27: Une fonction T et sa réciproque T^{-1}

Lorsque l'application est définie par une expression analytique, on peut trouver sa réciproque en résolvant directement l'équation. Par exemple, l'inverse de

$$y = \frac{x^3 - 1}{5},$$

(bien définie sur $\mathbb{R}$) est donnée par l'expression

$$x = \sqrt[3]{5y + 1}.$$

On considère l'application linéaire T de $\mathbb{R}^m$ dans $\mathbb{R}^n$ définie par

$$\overrightarrow{y} = T(\overrightarrow{x}) = A\overrightarrow{x},$$

où A est une matrice de taille $n \times m$. Par définition, cette application est inversible si et seulement si pour chaque $\overrightarrow{y} \in \mathbb{R}^n$, le système linéaire

$$A\overrightarrow{x} = \overrightarrow{y}$$

admet une solution unique $\overrightarrow{x} \in \mathbb{R}^m$.

Problème 2.3.1. *Quelles sont les matrices correspondant à une application linéaire inversible ?*

Solution. En utilisant les techniques de la section 1.3, on va déterminer pour quelles matrices A c'est le cas.

Rappelons qu'une matrice à n lignes et m colonnes représente une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$. On distinguera successivement les cas

$$n < m,$$
$$n = m,$$
$$n > m.$$

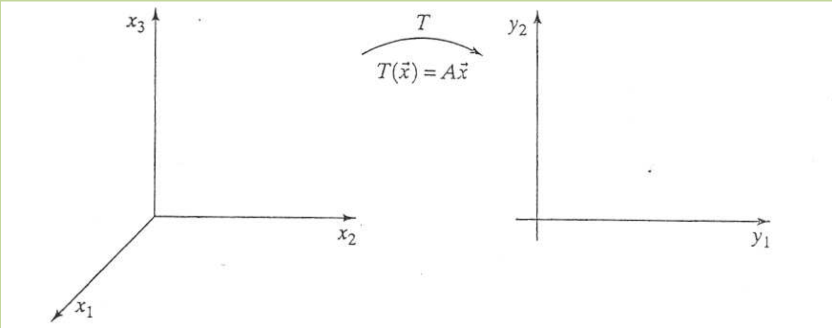


Figure 2.28: Cas $m = 3$, $n = 2$

Cas 1 : $n < m$

Le système $A\overrightarrow{x} = \overrightarrow{y}$ a moins d'équations (n) que d'inconnues (m). Donc étant donné $\overrightarrow{y} \in \mathbb{R}^n$, le système $A\overrightarrow{x} = \overrightarrow{y}$ admet soit une infinité de solution, soit est inconsistant d'après le théorème 2 d. L'application $\overrightarrow{y} = A\overrightarrow{x}$ n'est pas inversible.

A titre d'exemple, on considère le cas $m = 3$, $n = 2$, *i.e.* Intuitivement, on doit s'attendre à ce qu'il y ait des "écrasements" si on veut transformer $\mathbb{R}^3$ linéairement en $\mathbb{R}^2$. En d'autres termes, il faut s'attendre à ce que plusieurs points de $\mathbb{R}^3$ aient la même image dans $\mathbb{R}^2$. Le raisonnement précédent montre que c'est exactement ce qui arrive, et l'équation

$$T(\overrightarrow{x}) = \overrightarrow{0}$$

admet une infinité de solutions (voir exercice 2.1.42 par exemple).

Il est surprenant cependant de s'entendre dire qu'il existe des applications inversibles, **non linéaires**, de $\mathbb{R}^3$ dans $\mathbb{R}^2$.

Cas 2 : $n = m$

On est dans le cas où le système $A\overrightarrow{x} = \overrightarrow{y}$ a autant d'équations que d'inconnues. On sait par le théorème 4 que ce système admet une unique solution si et seulement si

$$\mathrm{Frel}(A) = \mathrm{I}_n = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 1 \end{bmatrix}.$$

On en déduit que l'application $\overrightarrow{y} = A\overrightarrow{x}$ est inversible si et seulement si $\mathrm{Frel}(A) = \mathrm{I}_n$, ou bien de manière équivalente si $\mathrm{Rang}A = n$.

Cas 3 : $n > m$

Dans ce cas, nous allons montrer que l'application linéaire $\overrightarrow{y} = A\overrightarrow{x}$ n'est pas inversible.

On rapelle que l'on a toujours $\mathrm{Rang}A \leq m$.

On commence par le cas $\mathrm{Rang}A < m$. Alors on sait par le théorème 2 d. que quelque soit $\overrightarrow{y} \in \mathbb{R}^n$, le système $A\overrightarrow{x} = \overrightarrow{y}$ est soit inconsistant, soit admet une infinité de solutions. Dans ce cas, l'application linéaire n'est $\overrightarrow{y} = A\overrightarrow{x}$ n'est pas inversible.

On étudie le cas $\mathrm{Rang}A = m$. On sait par le théorème 2 (c) que, quelque soit $\overrightarrow{y} \in \mathbb{R}^n$, le système $A\overrightarrow{x} = \overrightarrow{y}$ admet au plus une solution. Pour $\overrightarrow{0} \in \mathbb{R}^n$, le système est consistant et a donc a une et une seule solution.

Dans ce cas, la matrice réduite échelonée par ligne de la matrice augmentée du système $A\overrightarrow{x} = \overrightarrow{0}$ est de la forme :

$$\left[\begin{array}{ccccccc|c}
1 & 0 & 0 & . & . & . & 0 & 0 \\
0 & 1 & 0 & . & . & . & 0 & 0 \\
0 & 0 & 1 & . & . & . & 0 & 0 \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
0 & 0 & 0 & . & . & . & 1 & 0 \\
0 & 0 & 0 & . & . & . & 0 & 0 \\
. & . & . & . & . & . & . & . \\
0 & 0 & 0 & . & . & . & 0 & 0
\end{array}\right]$$

On prend

$$\overrightarrow{e}_{m+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}.$$

On inverse les transformations élémentaires qui ont permis de passer de A

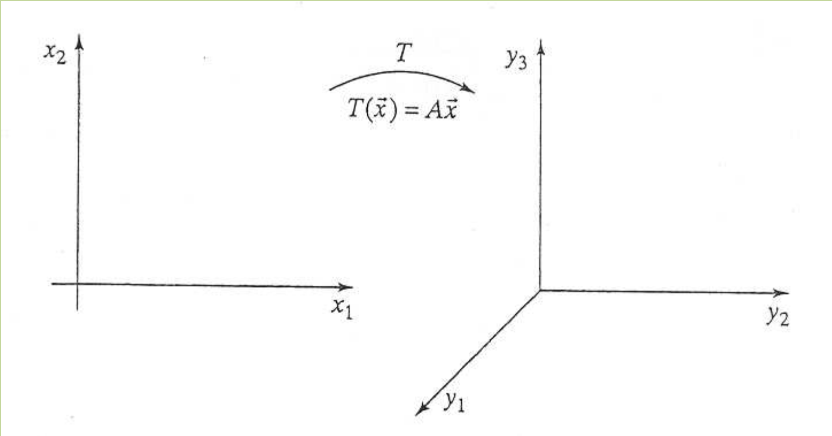


Figure 2.29: Cas $n = 3$, $m = 2$

à Frel(A) sur le système Frel$A\overrightarrow{x} = \overrightarrow{e}_{m+1}$, qui est visiblement inconsistant, et on obtient un système $A\overrightarrow{x} = \overrightarrow{b}$ qui est inconsistant également.

Donc, dans tous les cas, l'application linéaire $\overrightarrow{y} = A\overrightarrow{x}$ n'est pas inversible quand $m < n$.

A titre d'exemple, considérons le cas $n = 3$, $m = 2$.

On comprend bien intuitivement que l'on ne peut pas remplir tout $\mathbb{R}^3$ en plongeant linéairement $\mathbb{R}^2$ dans $\mathbb{R}^3$. Il y aura beaucoup de vecteurs $\overrightarrow{y} \in \mathbb{R}^3$ pour lesquels le système $T(\overrightarrow{x}) = \overrightarrow{y}$ sera inconsistant. □

Résumé 2.3.1 (Application linéaires inversibles)**.**

- *L'application linéaire T de $\mathbb{R}^m$ dans $\mathbb{R}^n$ définie par*

$$\overrightarrow{y} = T(\overrightarrow{x}) = A\overrightarrow{x},$$

où A est une matrice de taille $n \times m$, n'est **jamais inversible** *si $n \neq m$.*

- *L'application linéaire T de $\mathbb{R}^n$ dans $\mathbb{R}^n$ définie par*

$$\overrightarrow{y} = T(\overrightarrow{x}) = A\overrightarrow{x},$$

où A est une **matrice carrée** *de taille $n \times n$, est* **inversible** *si et seulement si* Frel(A) $=$ I$_n$*, c'est à dire si et seulement si* Rang$A = n$*.*

On considère les applications de $\mathbb{R}^2$ dans $\mathbb{R}^2$ étudiées dans la section 2.2. A titre d'exercice, on peut observer que

- Les rotations sont inversibles,
- les symétries sont inversibles,
- les transvections (cisaillement) sont inversibles,
- Les projection othogonales sur une droite **ne sont pas** inversibles.

Définition 2.3.3. *Une matrice A est dite* **inversible** *si l'application $\overrightarrow{y} = A\overrightarrow{x}$ est inversible. La matrice de l'application réciproque est notée A^{-1}.*

On a donc montré précédemment :

Proposition 2.3.1 (Inverse et forme réduite échelonnée par ligne)**.** *Une matrice A de taille $n \times m$ est inversible si et seulement si*

(a) A est une matrice carrée, i.e $n = m$,

(b) $\mathrm{Frel}(A) = \mathrm{I}_n$.

Example 2.3.1. *On considère la matrice $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, et on se demande si elle est inversible. Pour cela, on lui applique l'algorithme de Gauss-Jordan.*

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \longrightarrow \begin{matrix} \\ -4(\mathrm{L}_1) \\ -7(\mathrm{L}_1) \end{matrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix} \longrightarrow$$

$$\div(-3) \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -12 \end{bmatrix} \longrightarrow \begin{matrix} -2(\mathrm{L}_2) \\ \\ 6(\mathrm{L}_2) \end{matrix} \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \mathrm{Frel}(A).$$

La matrice A n'est pas inversible car $\mathrm{Frel}(A) \neq \mathrm{I}_3$.

On a donc montr;é :

Proposition 2.3.2 (Inversion et systèmes linéaires)**.** *Soit A une matrice de taille $n \times n$.*

a) Si il existe $\overrightarrow{b}_0 \in \mathbb{R}^n$ tel que le système $A\overrightarrow{x} = \overrightarrow{b}_0$ admet une solution unique, la matrice est inversible. Autrement dit, si A n'est pas inversible, alors pour tout $\overrightarrow{b} \in \mathbb{R}^n$ le système $A\overrightarrow{x} = \overrightarrow{b}$ a soit une infinité de solutions, soit aucune.

b) On considère le cas particulier d'un système homogène. Le système $A\overrightarrow{x} = \overrightarrow{0}$ admet toujours $\overrightarrow{x} = \overrightarrow{0}$ comme solution. Si $\overrightarrow{0}$ est l'unique solution de ce système, A est inversible.

Etant donnée une matrice A inversible, on se demande comment déterminer A^{-1} d'un point de vue pratique. On utilise l'algorithme de Gauss-Jordan comme on le montre sur l'exemple suivant.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix},$$

qui définit l'application linéaire

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + x_3 \\ 2x_1 + 3x_2 + 2x_3 \\ 3x_1 + 8x_2 + 2x_3 \end{bmatrix}.$$

Pour trouver la réciproque de cette application, on résoud le système suivant avec x_1, x_2, x_3 pour inconnues,

$$\left|\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & = & y_1 & & \\ 2x_1 & + & 3x_2 & + & 2x_3 & = & & y_2 & \\ 3x_1 & + & 8x_2 & + & 2x_3 & = & & & y_3 \end{array}\right| \longrightarrow$$

$$\begin{array}{r} \\ -2(\mathrm{L}_1) \\ -3(\mathrm{L}_1) \end{array} \left|\begin{array}{cccccccccc} x_1 & + & x_2 & + & x_3 & = & y_1 & & & \\ & & x_2 & & & = & -2y_1 & + & y_2 & \\ & & 5x_2 & - & x_3 & = & -3y_1 & & & + \; y_3 \end{array}\right| \longrightarrow$$

$$\begin{array}{r} -(\mathrm{L}_2) \\ \\ -5(\mathrm{L}_2) \end{array} \left|\begin{array}{cccccccccc} x_1 & & & + & x_3 & = & 3y_1 & - & y_2 & \\ & & x_2 & & & = & -2y_1 & + & y_2 & \\ & & & - & x_3 & = & 7y_1 & - & 5y_2 & + \; y_3 \end{array}\right| \longrightarrow$$

$$\begin{array}{r} \\ \\ \div(-1) \end{array} \left|\begin{array}{cccccccccc} x_1 & & & + & x_3 & = & 3y_1 & - & y_2 & \\ & & x_2 & & & = & -2y_1 & + & y_2 & \\ & & & & x_3 & = & -7y_1 & + & 5y_2 & - \; y_3 \end{array}\right| \longrightarrow$$

$$\begin{array}{r} -(\mathrm{L}_3) \\ \\ \\ \end{array} \left|\begin{array}{cccccccc} x_1 & & & = & 10y_1 & - & 6y_2 & + \; y_3 \\ & x_2 & & = & -2y_1 & + & y_2 & \\ & & x_3 & = & -7y_1 & + & 5y_2 & - \; y_3 \end{array}\right|,$$

ce qui fait que la matrice inverse de la matrice A est la matrice

$$A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{bmatrix}.$$

On peut aussi décrire les calculs précédents sous la forme matricielle suivante,

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array}\right] \longrightarrow$$

$$\begin{array}{c} \\ -2(\mathrm{L}_1) \\ -3(\mathrm{L}_1) \end{array} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 5 & -1 & -3 & 0 & 1 \end{array}\right] \longrightarrow$$

$$\begin{array}{c} -(\mathrm{L}_2) \\ \\ -5(\mathrm{L}_2) \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -1 & 7 & -5 & 1 \end{array}\right] \longrightarrow$$

$$\begin{array}{c} \\ \\ \div(-1) \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array}\right] \longrightarrow$$

$$\begin{array}{c} -(\mathrm{l}_3) \\ \\ \\ \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array}\right]$$

On est passé par des opérations élémentaires sur les lignes de

$$\left[\begin{array}{c|c} A & \mathrm{I}_3 \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array}\right]$$

à

$$\left[\begin{array}{c|c} \mathrm{I}_3 & B \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array}\right]$$

avec
$$B = A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{bmatrix}.$$

Algorithme 2. *Pour trouver l'inverse d'une matrice A de taille $n \times n$, on forme la matrice $C = [A \vdots \mathrm{I}_n]$ de taille $n \times (2n)$ et on calcule* $\mathrm{Frel}(C)$.

- *Si* $\mathrm{Frel}(C)$ *est de la forme* $[\mathrm{I}_n \vdots B]$, *alors A est inversible et $A^{-1} = B$.*
- *Si* $\mathrm{Frel}(C)$ *a une autre forme i.e. si sa partie de gauche n'est pas égale à I_n), alors A n'est pas inversible.*
- *Notez que dans tous les cas, la moitié gauche de* $\mathrm{Frel}(C)$ *est égal à* $\mathrm{Frel}(A)$, *par la proposition 1.2.1.*

L'inverse d'une matrice 2×2 est facile à calculer (voir exercices)

Proposition 2.3.3 (Inverse et déterminant d'une matrice 2×2)**.** *a) La matrice* 2×2
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
est inversible si et seulement si $ad - bc \neq 0$.

La quantité $ad - bc$ est appelée le **déterminant** *de la matrice A et est noté* $\mathrm{Det}(A)$,
$$\mathrm{Det}(A) = \mathrm{Det}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$
b) Si la matrice A est inversible, alors
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\mathrm{Det}(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

2.4 Produits matriciels et composition des applications

2.4.1 Introduction au produit de matrices

On a l'habitude de faire des **produits de nombres**.

Par exemple

$$2 \times 3 = 6$$

et on est habitué aux propriétés suivantes

- il n'y a pas de diviseur de 0: si un produit de deux nombres est nul c'est que l'un de ces deux nombres est nul
- le produit de deux nombres est commutatif:

  $$2 \times 3 = 3 \times 2$$

  et plus generalement pour tous nombres b et a

  $$a \times b = b \times a$$

On va généraliser le produit de nombre au **produit des tableaux de nombres**, c'est à-dire au produit de **matrices**.

Si

$$B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}, A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

sont deux matrices carrées de taille 2 (avec deux lignes et deux colonnes) on définit

$$B\ A = \begin{bmatrix} b_1\ a_1 + b_2\ a_3 & b_1\ a_2 + b_2\ a_4 \\ b_3\ a_1 + b_4\ a_3 & b_3\ a_2 + b_4\ a_4 \end{bmatrix}.$$

$B\ A$ est aussi une matrice de taille 2.

Par exemple, si

$$B = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}, A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix},$$

alors

$$B\ A = \begin{bmatrix} 6 \times 1 + 7 \times 3 & 6 \times 2 + 7 \times 5 \\ 8 \times 1 + 9 \times 3 & 8 \times 2 + 9 \times 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$$

Pour les débutants on dispose le calcul ainsi

$$\begin{array}{cccc} & & 1 & 2\\ & & 3 & 5\\ & & & \\ 6 & 7 & 27 & 47\\ 8 & 9 & 35 & 61\end{array}$$

Le produit des matrices a des propriétés étranges par rapport au produit de nombres

- il y a des diviseurs de 0: si un produit de deux matrices est nul (toutes les composantes sont nulles) il peut arriver qu'aucune des deux matrices ne soit nulle.

  Par exemple Si $B = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$ et $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$,

$$\begin{array}{cccc} & & 2 & 4\\ & & 1 & 2\\ & & & \\ 1 & -2 & 0 & 0\\ -2 & 4 & 0 & 0\end{array}$$

  autrement dit

$$B\ A = \begin{bmatrix} 1 \times 2 + -2 \times 1 & 1 \times 4 + -2 \times 2 \\ -2 \times 2 + 4 \times 1 & -2 \times 4 + 4 \times 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- le produit de deux matrices n'est pas toujours commutatif:

$$A\ B \neq B\ A.$$

  Par exemple si comme tout à l'heure $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ et $B = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$

$$\begin{array}{cccc} & & 1 & -2\\ & & -2 & 4\\ & & & \\ 2 & 4 & -6 & 12\\ 1 & 2 & -3 & 6\end{array}$$

  autrement dit

$$A\ B = \begin{bmatrix} 2 \times 1 + 4 \times -2 & 2 \times -2 + 4 \times 4 \\ 1 \times 1 + 2 \times -2 & 1 \times -2 + 2 \times 4 \end{bmatrix} = \begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix} \neq B\ A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

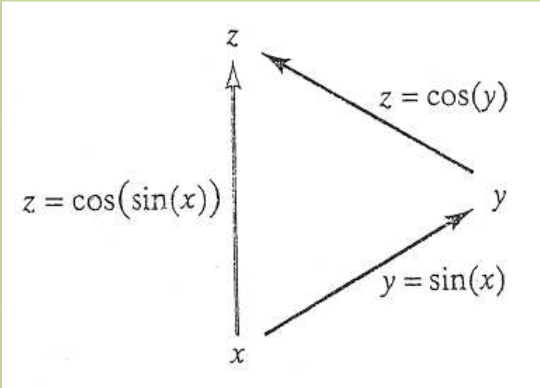


Figure 2.30: Composition des applications

2.4.2 Composition des applications et produit de matrices

C'est pour étudier la composition des applications linéaires que la multiplication des matrices va être très utile.

On commence par rappeler le concept de la composition de deux applications. La composition de $y = \sin(x) = f(x)$ avec la fonction $z = \cos(y) = g(y)$ est la fonction $z = \cos(\sin(x)) = (g \circ f)(x)$.

On peut composer de la même manière les applications linéaires. Retournons à l'exemple du début de la section 2.1. La position $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ du bateau est donnée par une position codée $\overrightarrow{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Le code est donné par l'application linéaire

$$\overrightarrow{y} = A\overrightarrow{x}, \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}.$$

On avait oublié un détail : la position du bateau est transmise à un central à Paris, et est codée à nouveau par l'application

$$\overrightarrow{z} = B\overrightarrow{y}, \quad B = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}.$$

La position du bateau reçue à Paris est donnée par la formule

$$\overrightarrow{z} = B(A\overrightarrow{x}),$$

comme étant la composition de $\overrightarrow{y} = A\overrightarrow{x}$ avec $\overrightarrow{z} = B\overrightarrow{y}$.

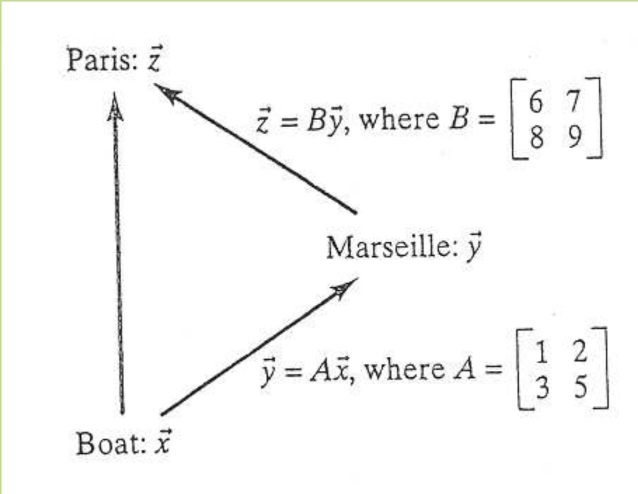


Figure 2.31: Composition d'applications linéaires

Problème 2.4.1. *Est-ce que l'application composée de deux applications linéaires de $\mathbb{R}^2$ and $\mathbb{R}^2$ est linéaire, et si oui quelle est sa matrice ?*

Solution. On écrit les formules composantes par composante,

$$(1)\ \begin{cases} z_1 = 6y_1 + 7y_2, \\ z_2 = 8y_1 + 9y_2, \end{cases} \qquad (2)\ \begin{cases} y_1 = x_1 + 2x_2, \\ y_2 = 3x_1 + 5x_2, \end{cases}$$

puis on substitue dans (1) les formules données pour les y_i dans (2), ce qui donne

$$\begin{aligned} z_1 &= 6(x_1 + 2x_2) + 7(3x_1 + 5x_2) &&= (6 \cdot 1 + 7 \cdot 3)x_1 + (6 \cdot 2 + 7 \cdot 5)x_2 \\ &= 27x_1 + 47x_2, \\ z_2 &= 8(x_1 + 2x_2) + 9(3x_1 + 5x_2) &&= (8 \cdot 1 + 9 \cdot 3)x_1 + (8 \cdot 2 + 9 \cdot 5)x_2 \\ &= 35x_1 + 61x_2, \end{aligned}$$

ce qui montre que la composée est bien linéaire et a pour matrice

$$BA = \begin{bmatrix} 6 \cdot 1 + 7 \cdot 3 & 6 \cdot 2 + 7 \cdot 5 \\ 8 \cdot 1 + 9 \cdot 3 & 8 \cdot 2 + 9 \cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}.$$

□

On considère maintenant le cas de matrices non néssairement 2×2.

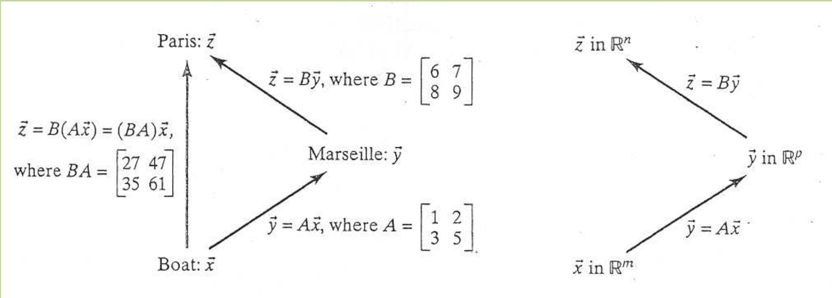


Figure 2.32: Vers le cas général

Proposition 2.4.1 (Composition d'applications linéaires)**.** *Soient B une matrice de taille $n \times p$ associée à l'application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$*

$$U(\overrightarrow{y}) = B\overrightarrow{y}$$

et A une matrice de taille $p \times m$ associée à l'application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^p$

$$T(\overrightarrow{x}) = A\overrightarrow{x}.$$

L'application composée

$$(U \circ T)(\overrightarrow{x}) = B(A\overrightarrow{x}),$$

de $\mathbb{R}^m$ dans $\mathbb{R}^n$, est linéaire.

Démonstration. On utilise la caractérisation des applications linéaires (section 2.1). On a :

$$\begin{aligned} (U \circ T)(\overrightarrow{v} + \overrightarrow{w}) &= B(A(\overrightarrow{v} + \overrightarrow{w})) && = B(A\overrightarrow{v} + A\overrightarrow{w}) \\ &= B(A\overrightarrow{v}) + B(A\overrightarrow{w}) && = (U \circ T)(\overrightarrow{v}) + (U \circ T)(\overrightarrow{w}) \\ (U \circ T)(k\overrightarrow{v}) &= B(A(k\overrightarrow{v})) && = B(kA\overrightarrow{v}) \\ &= kB(A\overrightarrow{v}) && = k(U \circ T)(\overrightarrow{v}). \end{aligned}$$

□

Définition 2.4.1 (Produit de matrices)**.** *Soient B une matrice de taille $n \times p$ associée à l'application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$*

$$U(\overrightarrow{y}) = B\overrightarrow{y}$$

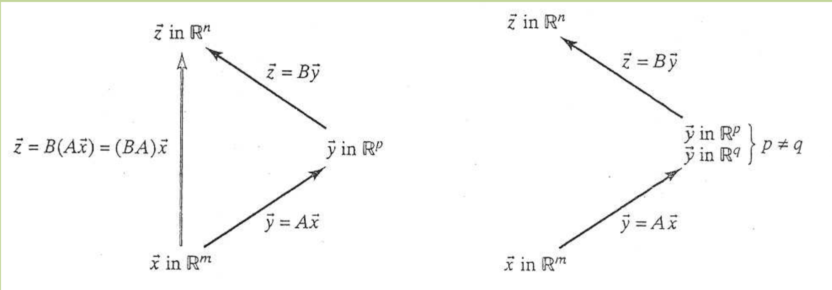


Figure 2.33: Compatibilité colonnes/lignes

et A une matrice de taille $p \times m$ associée à l'application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^p$

$$T(\overrightarrow{x}) = A\overrightarrow{x}.$$

le **produit** *BA, de taille $n \times m$ est défini comme étant la matrice de l'application linéaire composée*

$$(U \circ T)(\overrightarrow{x}) = B(A\overrightarrow{x})$$

qui a tout $\overrightarrow{x} \in \mathbb{R}^m$ associe $(U \circ T)(\overrightarrow{x}) \in \mathbb{R}^n$. Le produit BA est donc une matrice de taille $n \times m$ et on a

$$(U \circ T)(\overrightarrow{x}) = B(A\overrightarrow{x}) = (BA)\overrightarrow{x}.$$

Remarque 2.4.1. *Si le nombre de colonnes de B n'est pas égal au nombre de lignes de A, les applications $\overrightarrow{z} = B\overrightarrow{y}$ et $\overrightarrow{y} = A\overrightarrow{x}$ ne peuvent pas être composées car l'ensemble d'arrivée de $\overrightarrow{y} = A\overrightarrow{x}$ est différent de l'ensemble de départ de de $\overrightarrow{z} = B\overrightarrow{y}$. Autrement dit, la sortie de l'application $T(\overrightarrow{x}) = A\overrightarrow{x}$ n'est pas une entrée pour l'application $U(\overrightarrow{y}) = B\overrightarrow{y}$. Dans ce cas, la produit BA n'est pas défini.*

La définition qu'on vient de donner ne semble pas donner de moyens concrets pour calculer numériquement le produit de deux matrices. Pourtant ce moyen concret suit directement des définitions.

Soient B une matrice de taille $n \times p$ et A une matrice de taille $p \times m$. Étudions les colonnes de la matrice produit BA :

$$\begin{aligned}(i^{\text{ème}}\text{colonne de } BA) &= (BA)\overrightarrow{e}_i \\ &= B(A\overrightarrow{e}_i) \\ &= B(i^{\text{ème}}\text{colonne de } A).\end{aligned}$$

En notant $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ les colonnes de A, on a alors

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}$$

Proposition 2.4.2 (Colonnes d'une matrice produit)**.** *Soient B une matrice de taille $n \times p$ et A une matrice de taille $p \times m$. On note $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ les colonnes de A. alors le prduit BA est défini par*

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}$$

Pour déterminer BA il suffit d'effectuer la multiplication de B par chaque colonne de A et de recombiner en matrice l'ensemble des vecteurs ainsi déterminés.

On a vu dans la première section que la multiplication des matrices est une opération non-commutative, ce qui n'est pas une surprise. En effet, la composition des fonctions n'est pas une opération commutative.

Remarque 2.4.2 (La multiplication des matrices n'est pas commutative)**.** *Soient B une matrice de taille $n \times p$ et A une matrice de taille $p \times n$. Alors AB est une matrice de taille $p \times p$ et BA de taille $n \times n$. Dans le cas où $p = n$, on peut comparer les produits AB et BA.*

En général, $AB \neq BA$. Néanmoins, il arrive parfois que $AB = BA$; dans ce cas, on dit que les matrices **commutent**.

Il est utile d'avoir une formule analytique pour la composante ij du produit BA. On rappelle que

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}$$

le coefficient ij du produit BA est la $i^{\text{ième}}$ composante du vecteur $B\vec{v}_j$, qui est le produit de la $i^{\text{ième}}$ ligne de B par la $j^{\text{ième}}$ colonne de A.

Si on note $[BA]_{ij}$ le coefficient à la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne de la matrice produit BA, on a alors

$$[BA]_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ip}a_{pj} = \sum_{k=1}^{p} b_{ik}a_{kj}.$$

Proposition 2.4.3 (Les coefficients de la matrice produit)**.** *Soient B une matrice de taille $n \times p$ et A une matrice de taille $p \times m$. Le coefficient ij du produit BA est le produit de la $i^{\text{ième}}$ ligne de B par la $j^{\text{ième}}$ colonne de A.*

La matrice

$$BA = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots & \ldots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pj} & \cdots & a_{pm} \end{bmatrix}$$

est la matrice de taille $n \times m$ dont le coefficient à la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne est donné par la formule

$$[BA]_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ip}a_{pj} = \sum_{k=1}^{p} b_{ik}a_{kj}.$$

Example 2.4.1.

$$\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 6 \cdot 1 + 7 \cdot 3 & 6 \cdot 2 + 7 \cdot 5 \\ 8 \cdot 1 + 9 \cdot 3 & 8 \cdot 2 + 9 \cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}.$$

2.4.3 Calculs algébriques avec les matrices

Nous allons décrire dans ce qui suit les prinicipes du calcul algébrique des matrices.

Soit A une matrice carrée de taille $n \times n$, inversible. La matrice A multipliée par la matrice A^{-1} représente l'application identité.

Résumé 2.4.1 (Matrice inverse)**.** *Étant donnée A une matrice inversible carrée de taille $n \times n$, on a*

$$\boxed{AA^{-1} = A^{-1}A = \mathrm{I}_n.}$$

Composer l'application identité par une application linéaire des deux cotés, laisse invariante l'application linéaire considérée.

Résumé 2.4.2 (Multiplication par la matrice identité)**.** *Étant donnée A une matrice carrée de taille $n \times n$, on a*

$$A\,\mathrm{I}_n = \mathrm{I}_n A = A.$$

Problème 2.4.2. *Soit A une matrice $n \times p$, B une matrice $p \times q$, C une matrice $q \times m$. Quelle est la relation entre $(AB)C$ et $A(BC)$?*

Solution. Une manière de réfléchir à ce problème (même si ce n'est pas la plus élégante), consiste à écrire C à l'aide de ses vecteurs colonnes, $C = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix}$. On a alors

$$\begin{aligned} (AB)C &= (AB)\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix} \\ &= \begin{bmatrix} (AB)\overrightarrow{v}_1 & (AB)\overrightarrow{v}_2 & \cdots & (AB)\overrightarrow{v}_n \end{bmatrix}, \end{aligned}$$

tandis-que

$$\begin{aligned} A(BC) &= A\begin{bmatrix} B\overrightarrow{v}_1 & B\overrightarrow{v}_2 & \cdots & B\overrightarrow{v}_n \end{bmatrix} \\ &= \begin{bmatrix} A(B\overrightarrow{v}_1) & A(B\overrightarrow{v}_2) & \cdots & A(B\overrightarrow{v}_n) \end{bmatrix}, \end{aligned}$$

Puisque $(AB)\overrightarrow{v}_i = A(B\overrightarrow{v}_i)$ par définition du produit matriciel, on en déduit que $(AB)C = A(BC)$. □

Résumé 2.4.3 (Associativité du produit matriciel)**.** *On a toujours*

$$(AB)C = A(BC),$$

et on écrira ABC au lieu de $A(BC) = (AB)C$.

Une démonstration plus conceptuelle repose sur l'associativité de la composition des applications. Les deux application linéaires

$$T(\overrightarrow{x}) = ((AB)C)\overrightarrow{x}, \quad L(\overrightarrow{x}) = (A(BC))\overrightarrow{x}$$

sont identiques, car la définition de la multiplication des matrices montre que

$$T(\overrightarrow{x}) = ((AB)C)\overrightarrow{x} = (AB)(C\overrightarrow{x}) = A(B(C\overrightarrow{x})),$$

tandis-que

$$L(\overrightarrow{x}) = (A(BC))\overrightarrow{x} = A((BC)\overrightarrow{x})) = A(B(C\overrightarrow{x})).$$

Les ensembles de départ et d'arrivée respectifs des applications linéaires définies par les matrices A, B, C, BC, AB, $A(BC)$ et $(AB)C$ sont décrits dans la figure ci-dessous.

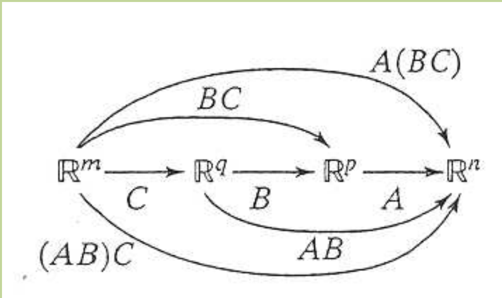


Figure 2.34: Associativité du produit matriciel

Problème 2.4.3. *Soient A et B deux matrices carrées de taille $n \times n$. On suppose A et B inversibles. Est-ce que le produit BA est encore inversible ?*

Solution. Pour trouver la réciproque de l'application linéaire

$$\overrightarrow{y} = BA\overrightarrow{x},$$

on va résoudre l'équation en $\overrightarrow{x}$ en deux temps. On commence par multiplier à gauche les deux membres de cette équations par B^{-1} :

$$B^{-1}\overrightarrow{y} = B^{-1}BA\overrightarrow{x} = \mathrm{I}_n A\overrightarrow{x} = A\overrightarrow{x}.$$

On multiplie ensuite à gauche par A^{-1}. Il vient

$$A^{-1}B^{-1}\overrightarrow{y} = A^{-1}A\overrightarrow{x} = \mathrm{I}_n\overrightarrow{x} = \overrightarrow{x}.$$

Ce calcul montre que l'application

$$\overrightarrow{y} = BA\overrightarrow{x}$$

est inversible et que son inverse est l'application

$$\overrightarrow{x} = A^{-1}B^{-1}\overrightarrow{y}.$$

□

Résumé 2.4.4 (Inverse d'un produit de matrices.)**.** *Soient A et B deux matrices carrées inversibles de taille n. Alors le produit AB est inversible et on a*

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

Attention à l'ordre des produits !

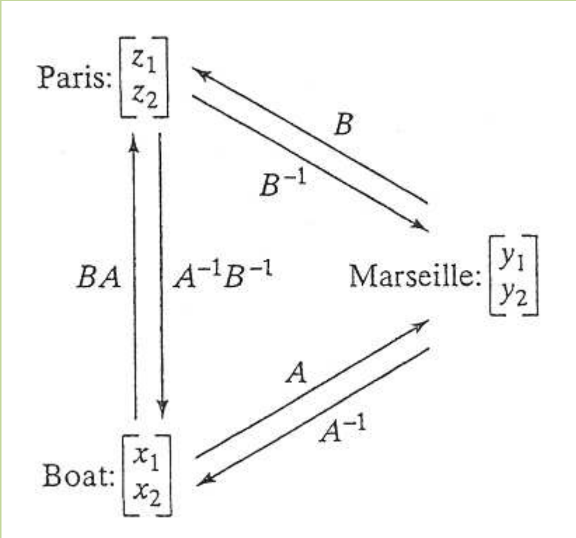


Figure 2.35: Inverse d'un produit de matrices

Pour vérifier ce résultat autrement, on effectue le calcul suivant, en utilisant l'associativité du produit,

$$(A^{-1}B^{-1})(BA) = A^{-1}(B^{-1}B)A = A^{-1}(\mathrm{I}_n)A = A^{-1}A = \mathrm{I}_n,$$

et tout marche très bien.

Pour mieux comprendre l'ordre des facteurs dans la formule $(BA)^{-1} = A^{-1}B^{-1}$, repensons à notre histoire de bateau marseillais.

Pour trouver la position effective $\overrightarrow{x}$ à partir du double codage $\overrightarrow{z}$, on commence par effectuer la transformation $\overrightarrow{y} = B^{-1}\overrightarrow{z}$ et *ensuite* la transformation $\overrightarrow{x} = A^{-1}\overrightarrow{y}$. Donc la réciproque de l'application $\overrightarrow{z} = BA\overrightarrow{x}$ est bien l'application $\overrightarrow{x} = A^{-1}B^{-1}\overrightarrow{z}$.

Proposition 2.4.4 (Un critère d'inversibilité)**.** *Soient A et B deux matrices carrées de taille n telles que*

$$BA = \mathrm{I}_n.$$

Alors

a) A et B sont toutes les deux inversibles

b) $A^{-1} = B$ et $B^{-1} = A$,

c) $AB = \mathrm{I}_n$.

La définition de l'inverse d'une application nous dit que lorsque $BA = \mathrm{I}_n$ **et** $AB = \mathrm{I}_n$, alors A et B sont inversibles et inverses l'une de l'autre. Le résultat ci-dessus dit que l'équation $BA = \mathrm{I}_n$ à elle seule suffit pour assurer que A et B soient inversibles et inverses l'une de l'autre (à condition que B et A soient carrées).

Démonstration. Pour montrer que A est inversible, il nous suffit de montrer que le système homogène $A\overrightarrow{x} = \overrightarrow{0}$ admet $\overrightarrow{0}$ comme unique solution , d'après la proposition 2.3.2. Multiplions à gauche l'équation $A\overrightarrow{x} = \overrightarrow{0}$ par B. On obtient $BA\overrightarrow{x} = B\overrightarrow{0} = \overrightarrow{0}$. Comme $BA = \mathrm{I}_n$, il en résulte que $\overrightarrow{x} = \overrightarrow{0}$. Donc A est inversible.

En multipliant à droite l'équation $BA = \mathrm{I}_n$ par A^{-1}, il vient

$$(BA)A^{-1} = \mathrm{I}_n A^{-1}, \quad \text{c'est-à-dire} \quad B = A^{-1}.$$

La matrice B étant l'inverse de A est aussi inversible, et $B^{-1} = (A^{-1})^{-1} = A$ (résulte des définitions). Enfin, $AB = AA^{-1} = \mathrm{I}_n$. □

À titre d'application, considérons le cas $2D$. Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ avec un déterminant non nul. On va vérifier que

$$B = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\mathrm{Det}(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = A^{-1}.$$

Pour cela il est suffisant de vérifier que $BA = \mathrm{I}_2$. On a

$$\begin{aligned} BA = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \frac{1}{ad - bc}\begin{bmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{bmatrix} \\ &= \mathrm{I}_2 \end{aligned}$$

Problème 2.4.4. *Soient A, B et C trois matrices carrées de taille n telles que $ABC = \mathrm{I}_n$. Montrer que B est inversible et exprimer B^{-1} en fonction A et C.*

Solution. On écrit $ABC = (AB)C = \mathrm{I}_n$. On en déduit que AB et C sont inverses l'une de l'autre, donc elles commutent. On a donc également $C(AB) = \mathrm{I}_n = (CA)B$. il en résulte que B est inversible et $B^{-1} = CA$. □

Proposition 2.4.5 (Distributivité du produit matriciel)**.** *Soient A et B deux matrices de taille $n \times p$, C et D de taille $p \times m$. On a alors*

$$A(C + D) = AC + AD \quad \text{et}$$
$$(A + B)C = AC + BC.$$

Ce point sera à vérifier dans l'exercice 63, section 2.4, et le point suivant dans l'exercice 64 :

Proposition 2.4.6. *Soit A une matrice de taille $n \times p$, B de taille $p \times m$, $k \in \mathbb{R}$. Alors*

$$(kA)B = A(kB) = k(AB).$$

2.4.4 Calcul matriciel par blocs

Il est parfois utile de subdiviser une grosse matrice en sous-matrices (blocs) accolées les une aux autres. Par exemple, on peut considérer la matrice 4×4

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 \end{bmatrix}$$

comme étant une matrice 2×2 ayant pour coefficients quatre matrices 2×2

$$A = \left[\begin{array}{cc:cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

où $A_{11} = \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 3 & 4 \\ 7 & 8 \end{bmatrix}$, etc...

Le découpage peut être avec des matrices de taille différentes, par exemple on peut avoir

$$A = \left[\begin{array}{ccc:c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 8 & 7 & 6 \\ \hline 5 & 4 & 3 & 2 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

On a une intéressante propriété du partitionnement des matrices par blocs :

Proposition 2.4.7 (Calcul matriciel par blocs)**.** *Le produit de matrices par bloc s'effectue de la même manière que le produit habituel, sauf que les coefficients sont remplacés par les blocs :*

$$BA = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1p} \\ B_{21} & B_{22} & \cdots & B_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ B_{i1} & B_{i2} & \cdots & B_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{np} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1j} & \cdots & A_{1m} \\ A_{21} & a_{22} & \cdots & A_{2j} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pj} & \cdots & A_{pm} \end{bmatrix}$$

est la matrice dont le bloc labélisé par les indices i et j est est donné par la formule

$$[BA]_{ij} = B_{i1}A_{1j} + B_{i2}A_{2j} + \cdots + B_{ip}A_{pj} = \sum_{k=1}^{p} B_{ik}A_{kj},$$

sous réserve que les tailles des blocs soient compatibles pour pouvoir définir les produits matriciels $B_{ik}A_{kj}$.

La démonstration de ce résultat est laissé en exercice.

Exercice 2.4.1.

$$A = \left[\begin{array}{cc:c} 0 & 1 & -1 \\ 1 & 0 & 1 \end{array}\right] \left[\begin{array}{cc:c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right]$$

$$= \left[\begin{array}{c:c} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}\begin{bmatrix} 7 & 8 \end{bmatrix} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ 6 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}\begin{bmatrix} 9 \end{bmatrix} \end{array}\right]$$

$$= \left[\begin{array}{cc:c} -3 & -3 & -3 \\ 8 & 10 & 12 \end{array}\right]$$

Calculez le produit de l'exemple 3 par la méthode standard et comparez les résultats.

Cet exemple est donné juste à titre d'illustration et n'a pas une grande portée. Le calcul par bloc est souvent très utile dans des raisonnements théoriques comme cela est illustré dans l'exemple suivant.

Problème 2.4.5. *Soit A la matrice par blocs*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

où A_{11} est une matrice carrée de taille n ($A_{11} \in M_n(\mathbb{R})$), A_{22} est matrice carrée de taille m ($A_{22} \in M_m(\mathbb{R})$), et enfin A_{12} est une matrice $n \times m$ ($A_{12} \in M_{n,m}(\mathbb{R})$).

a) Pour quelles matrices A_{11}, A_{22}, A_{12} la matrice A est-elle inversible.

b) Dans le cas où A est inversible, exprimer A^{-1} en fonction des matrice A_{11}, A_{22}, A_{12}.

Solution. On cherche $B \in M_{n+m}(\mathbb{R})$ telle que

$$BA = \mathrm{I}_{n+m} = \begin{bmatrix} \mathrm{I}_n & 0 \\ 0 & \mathrm{I}_m \end{bmatrix}.$$

On cherche la matrice B sous la forme par blocs

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

où $B_{11} \in M_n(\mathbb{R})$, $B_{22} \in M_m(\mathbb{R})$ et ainsi de suite. Le fait que B soit l'inverse de A conduit aux équations maricielles suivantes :

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \begin{bmatrix} \mathrm{I}_n & 0 \\ 0 & \mathrm{I}_m \end{bmatrix},$$

ce qui conduit en utilisant la règle de calculs par blocs, on obtient le système matriciel :

$$\left| \begin{array}{rcr} B_{11}A_{11} & = & \mathrm{I}_n \\ B_{11}A_{12} + B_{12}A_{22} & = & 0 \\ B_{21}A_{11} & = & 0 \\ B_{21}A_{12} + B_{22}A_{22} & = & \mathrm{I}_m \end{array} \right|.$$

La première équation implique que A_{11} doit être inversible et que l'on a $B_{11} = A_{11}^{-1}$.

En multipliant la troisième équation à droite par A_{11}^{-1} on obtient $B_{21} = 0$. La quatrième équation se simplifie alors en $B_{22}A_{22} = \mathrm{I}_m$. Il en résulte que A_{22} doit être inversible et que $B_{22} = A_{22}^{-1}$.

Finalement, la deuxième équation devient $A_{11}^{-1}A_{12} + B_{12}A_{22} = 0$, c'est-à-dire $A_{11}^{-1}A_{12} = -B_{12}A_{22}$, soit encore $B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$.

Conclusion

a) A est inversible si et seulement si A_{11} et A_{22} sont inversibles (il n'y a aucune condition particulière sur A_{12}).

b) Lorsque A est inversible, son inverse est donnée par la matrice écrite en bloc :

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}.$$

$\square$

Exercice 2.4.2. *Vérifier ce résultat avec la matrice A suivante*

$$A = \left[\begin{array}{cc:ccc} 1 & 1 & 1 & 2 & 3 \\ 1 & 2 & 4 & 5 & 6 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]^{-1} = \left[\begin{array}{rr:rrr} 2 & -1 & 2 & 1 & 0 \\ -1 & 1 & -3 & -3 & -3 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Chapter 3

Sous-Espaces de $\mathbb{R}^n$ et leur dimension

3.1 Image et noyau d'une application linéaire

3.1.1 L'image d'une application linéaire

Définition 3.1.1 (Image). *L'image d'une application est l'ensemble des valeurs prises par l'application dans l'ensemble d'arrivée. Soit f l'application*

$$f : \begin{cases} X \longrightarrow Y \\ x \longrightarrow f(x), \end{cases}$$

alors

$$\begin{aligned} \mathrm{Im}(f) &= \{f(x) \mid x \in X\}, \\ &= \{b \in Y \mid \exists x \in X;\ b = f(x)\}. \end{aligned}$$

Example 3.1.1. *Un groupe de journalistes X et un groupe de chefs d'état Y sont face à face dans une conférence de presse. Chaque journaliste jette une chaussure à la tête d'un chef d'état, et fait mouche à chaque fois. On considère l'application $y = f(x)$ qui va de X dans Y et qui associe à chaque journaliste x la cible y de sa chaussure. L'image de f est l'ensemble des chefs d'état ayant été touché par une chaussure.*

Example 3.1.2. *L'image de la fonction exponentielle $f(x) = e^x$ de $\mathbb{R}$ dans $\mathbb{R}$ est l'ensemble des réels strictement positifs, $\mathrm{Im}(f) = \mathbb{R}_+^\star$. En effet, $f(x) = e^x$ est positif quelque soit $x \in \mathbb{R}$. Soit maintenant $b > 0$. Alors, on a*

$$b = e^{\ln b} = f(\ln b).$$

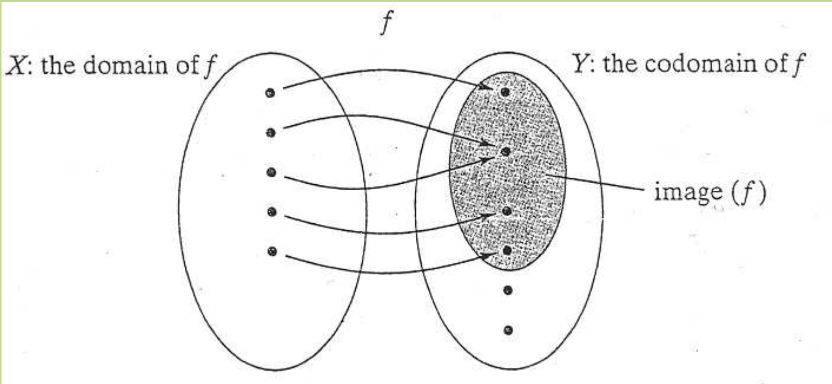


Figure 3.1: Image d'une fonction

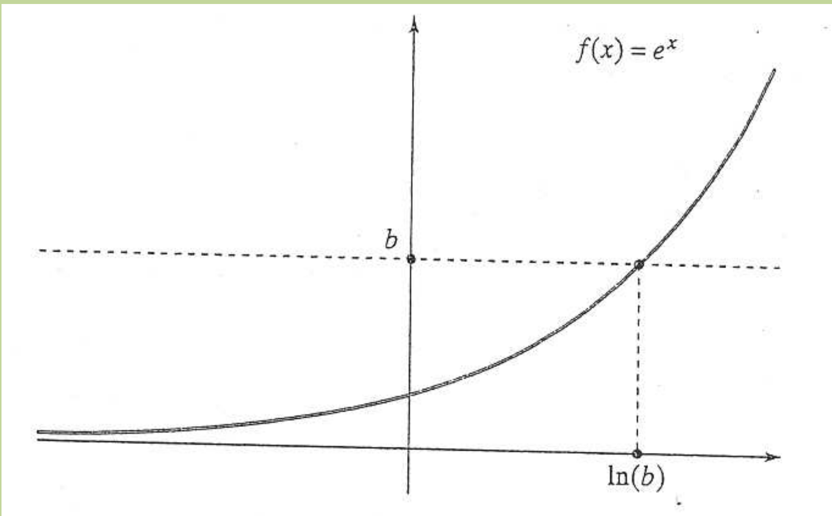


Figure 3.2: Image de la fonction exponentielle

Plus généralement, l'image de la fonction $y = f(x)$ de $\mathbb{R}$ dans $\mathbb{R}$ est l'ensemble des $b \in \mathbb{R}$, tels que la droite $y = b$ a une intersection non vide avec le graphe de f.

L'image de f est la projection othogonale du graphe de f sur l'axe vertical.

Example 3.1.3. *Soit*

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R}^2 \\ t \longrightarrow \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}. \end{cases}$$

L'image de f est le cercle unité centré à l'origine, cercle que l'on note S_1. En effet, pour $t \in \mathbb{R}$, $\cos(t)^2 + \sin(t)^2 = 1$, et donc $f(t) \in S_1$. Réciproquement, chaque vecteur $\overrightarrow{u}$ de norme 1 peut se mettre sous la forme

$$\overrightarrow{u} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} = f(t),$$

où t est l'angle polaire.

La fonction f de l'exemple 3.1.3 est une **paramétrisation** du cercle unité. Plus généralement, la paramétrisation d'une courbe C dans $\mathbb{R}^2$, est une fonction $g : \mathbb{R} \to \mathbb{R}^2$ qui a pour image la courbe C.

Example 3.1.4. *Soit*

$$f : \begin{cases} X \longrightarrow Y \\ x \longrightarrow f(x), \end{cases}$$

une application inversible. Alors $\mathrm{Im}(f) = Y$. En effet, on note que par définition, $\mathrm{Im}(f) \subset Y$. Soit $b \in Y$ quelconque. Comme f est inversible, il existe un unique $x \in X$ tel que $b = f(x)$, c'est-à-dire $b \in \mathrm{Im}(f)$. Cela étant vérifié $\forall\, b \in Y$, il en résulte que $Y \subset \mathrm{Im}(f)$, ce qui finalement conduit à $\mathrm{Im}(f) = Y$, comme annoncé (voir figure 3.4).

Example 3.1.5. *Soit $T : \mathbb{R}^3 \to \mathbb{R}^3$ la projection orthogonale sur le plan x_1, x_2, c'est-à-dire $T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$. Alors $\mathrm{Im}(T)$ est le plan x_1, x_2, constitué par les vecteurs de la forme $\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$, $x_i \in \mathbb{R}$.*

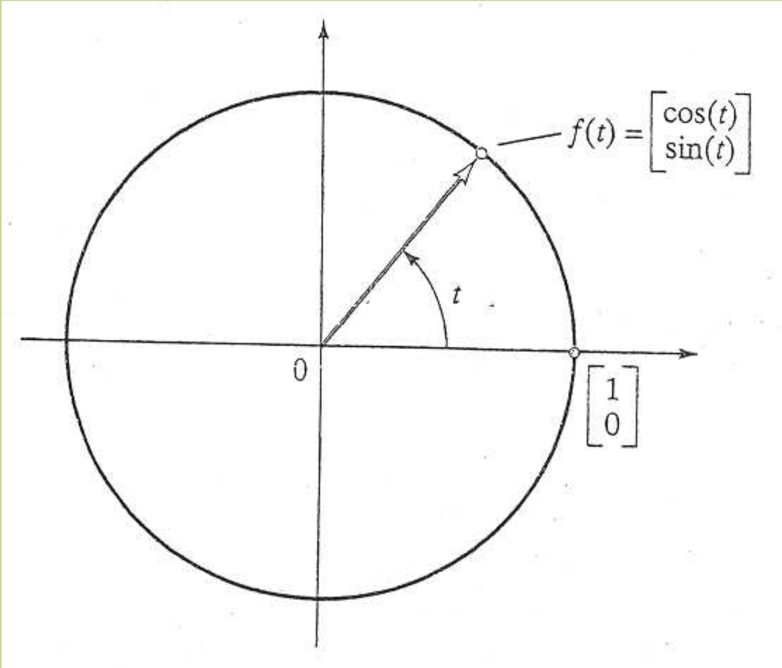


Figure 3.3: Cercle unité

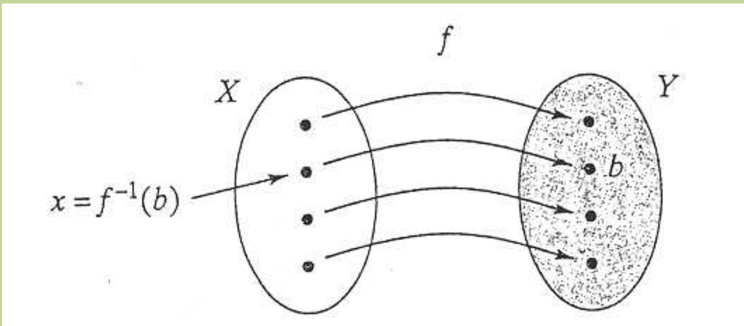


Figure 3.4: Exemple avec des ensembles de cardinal fini

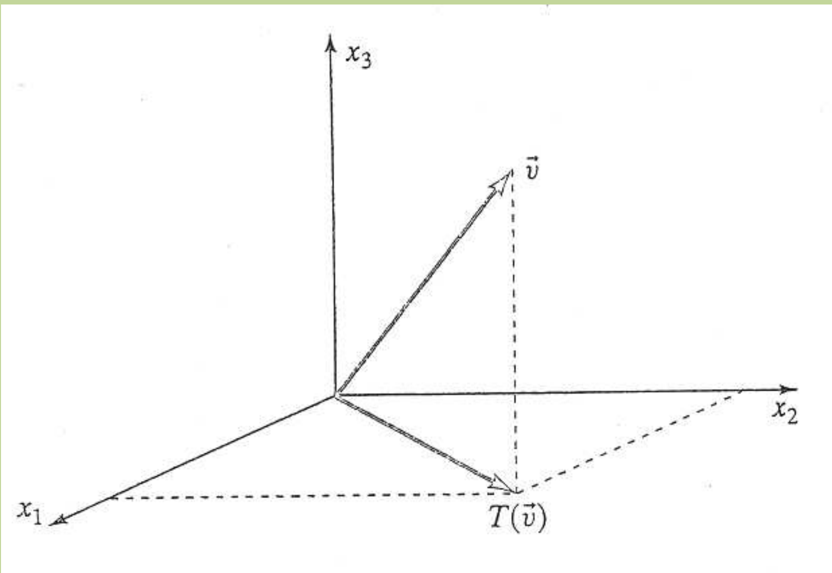


Figure 3.5: Projection orthogonale sur le plan x_1, x_2

Problème 3.1.1. *Déterminer* $\mathrm{Im}(T)$, *où* $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(\overrightarrow{x}) = A\overrightarrow{x}$, *avec* $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

Solution. L'image de T est l'ensemble des vecteurs "atteints" par T, c'est-à-dire tous les vecteurs de la forme

$$\begin{aligned} T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = &\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ &= x_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3x_2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = (x_1 + 3x_2)\begin{bmatrix} 1 \\ 2 \end{bmatrix}. \end{aligned}$$

L'image de T est la droite D ayant pour vecteur directeur le vecteur $\overrightarrow{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. En effet, ce qui précède assure que $\mathrm{Im}(T) \subset D$. Soit alors $\lambda \in \mathbb{R}$. Il existe une infinité de x_1 et x_2 dans $\mathbb{R}$, tels que $x_1 + 3x_2 = \lambda$.

Alors on a pour chaque x_1 et x_2 ainsi choisis, $T\begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = \lambda\overrightarrow{u}$. Donc $D \subset \mathrm{Im}(T)$ et finalement $D = \mathrm{Im}(T)$.

On notera que les deux vecteurs colonnes de la matrice A qui représente T, $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ et $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$ sont parallèles, ou bien encore "*colinéaires*"

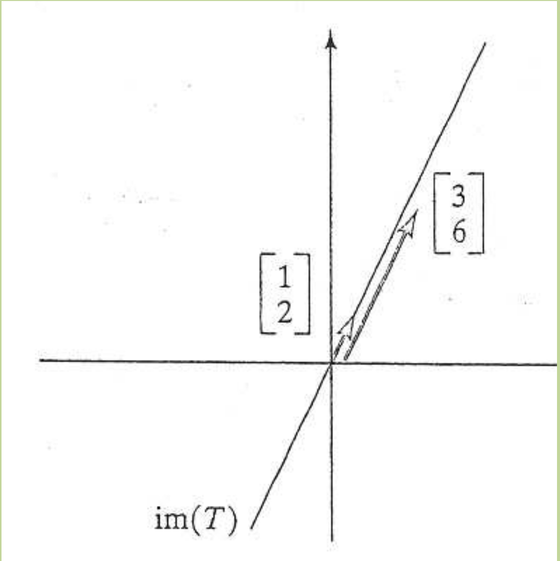


Figure 3.6: Image de T

□

Exercice 3.1.1. *Déterminer* Im(T), *où* $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(\vec{x}) = A\vec{x}$, *avec* $A = \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\end{bmatrix}$.

Solution. L'image de T est l'ensemble des vecteurs "atteints" par T, c'est-à-dire tous les vecteurs de la forme

$$T\begin{bmatrix}x_1\\x_2\end{bmatrix} = A\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = x_1\begin{bmatrix}1\\1\\1\end{bmatrix} + x_2\begin{bmatrix}1\\2\\3\end{bmatrix},$$

c'est-à-dire l'ensemble des combinaisons linéaires des vecteurs

$$\vec{v}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix}1\\2\\3\end{bmatrix}.$$

L'image de T est le plan **engendré** par les vecteurs $\vec{v}_1$ et $\vec{v}_2$, aussi décrit comme le plan qui passe par l'origine et les deux extrémités des vecteurs v_1 et v_2.

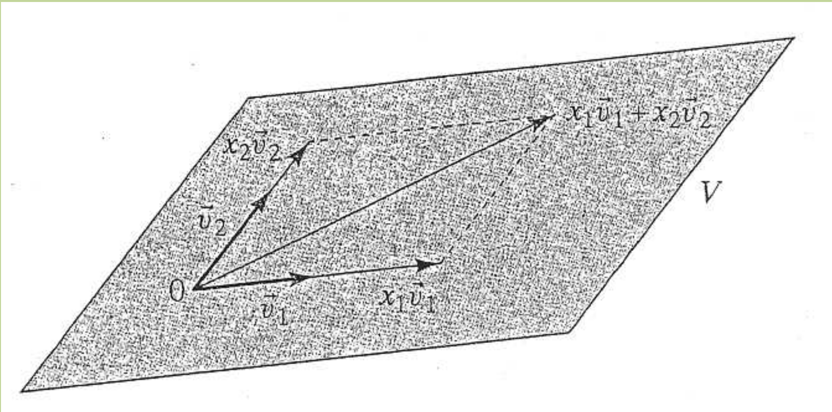


Figure 3.7: Image de T

Il faut noter que dans le cas de cet exemple, les deux vecteurs colonne de la matrice A ne sont pas **colinéaires** mais sont **linéairement indépendants**. $\square$

Définition 3.1.2 (Espace vectoriel engendré par une famille de vecteurs)**.** *Soit $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ une famille de vecteurs dans $\mathbb{R}^n$. L'ensemble des combinaisons linéaires de ces vecteurs est l'espace vectoriel engendré par cette famille de vecteurs. On le note*

$$\begin{aligned}&\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m) = \\ &\{\lambda_1\overrightarrow{v}_1 + \lambda_2\overrightarrow{v}_2 + \cdots + \lambda_m\overrightarrow{v}_m, \quad \lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{R}\}.\end{aligned}$$

Si $V = \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m)$, on dit que la famille $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ **engendre** *V, ou encore est une* **famille génératrice** *de V.*

La notion d' **espace vectoriel** sera détaillée plus en détail dans la section 3.2.

Proposition 3.1.1 (Image d'une application linéaire)**.** *L'image d'une application linéaire $T(\overrightarrow{x}) = A\overrightarrow{x}$ est l'espace vectoriel engendré par les vecteurs colonnes de A. On note cet espace* $\mathrm{Im}(T)$*.*

Démonstration. Pour démontrer ce résultat, on écrit l'application sous forme vectorielle comme dans les exemples précédents,

$$\begin{aligned} T(\vec{x}) = A\vec{x} &= \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \\ &= x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_m\vec{v}_m. \end{aligned}$$

Ce calcul montre que

$$\mathrm{Im}(T) = \mathrm{Vect}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m).$$

Soit

$$T : \begin{cases} \mathbb{R}^m \longrightarrow \mathbb{R}^n \\ \vec{x} \longrightarrow T(\vec{x}) = A\vec{x}, \end{cases}$$

une application linéaire. Son image a les propriétés suivantes.
(a) Le vecteur nul $\vec{0} \in \mathbb{R}^n$ est toujours dans l'image de T, car on a toujours $\vec{0} = A\vec{0} = T(\vec{0})$.
(b) Soient $\vec{v}_1 \in \mathrm{Im}(T)$ et $\vec{v}_2 \in \mathrm{Im}(T)$. Alors on a encore

$$\vec{v}_1 + \vec{v}_2 \in \mathrm{Im}(T).$$

Pour vérifier cette affirmation, il suffit de dire que puisque $\vec{v}_1 \in \mathrm{Im}(T)$ alors il existe $\vec{w}_1 \in \mathbb{R}^m$ tel que $\vec{v}_1 = T(\vec{w}_1)$, et de même il existe $\vec{w}_2 \in \mathbb{R}^m$ tel que $\vec{v}_2 = T(\vec{w}_2)$. Alors $\vec{v}_1 + \vec{v}_2 = T(\vec{w}_1) + T(\vec{w}_2) = T(\vec{w}_1 + \vec{w}_2)$ ce qui montre que $\vec{v}_1$ et $\vec{v}_2$ est bien dans $\mathrm{Im}(T)$.
(c) Soient $\vec{v} \in \mathrm{Im}(T)$ et $\lambda \in \mathbb{R}$. Alors on a encore $\lambda\vec{v} \in \mathrm{Im}(T)$.

En effet, puisque $\vec{v} \in \mathrm{Im}(T)$, il existe $\vec{w} \in \mathbb{R}^m$ tel que $\vec{v} = T(\vec{w})$. Par conséquent, $\lambda\vec{v} = \lambda T(\vec{w}) = T(\lambda\vec{w})$, ce qui montre que $\lambda\vec{v} \in \mathrm{Im}(T)$. □

Résumé 3.1.1 (Propriétés de l'image d'une application linéaire)**.** *L'image d'une application linéaire $T : \mathbb{R}^m \to \mathbb{R}^n$ a les propriétés suivantes :*

(a) $\vec{0} \in \mathrm{Im}(T)$,

(b) $\forall\, \vec{v}_1 \in \mathrm{Im}(T)$, $\forall\, \vec{v}_2 \in \mathrm{Im}(T)$, *alors* $\vec{v}_1 + \vec{v}_2 \in \mathrm{Im}(T)$, *(stabilité de l'addition)*

(c) $\forall\, \vec{v} \in \mathrm{Im}(T)$, $\forall\, \lambda \in \mathbb{R}$, *alors* $\lambda\vec{v} \in \mathrm{Im}(T)$ *(stabilité de la multiplication par un scalaire).*

Il résulte des propriétés (b) et (c) que l'image d'une application linéaire T est stable par combinaisons linéaires : si des vecteurs $\vec{v}_1, \cdots, \vec{v}_p$ sont dans l'image de T, $\lambda_1, \cdots, \lambda_p$ sont des scalaires, alors $\lambda\vec{v}_1 + \cdots + \lambda_p\vec{v}_p$ est encore dans l'image de T.

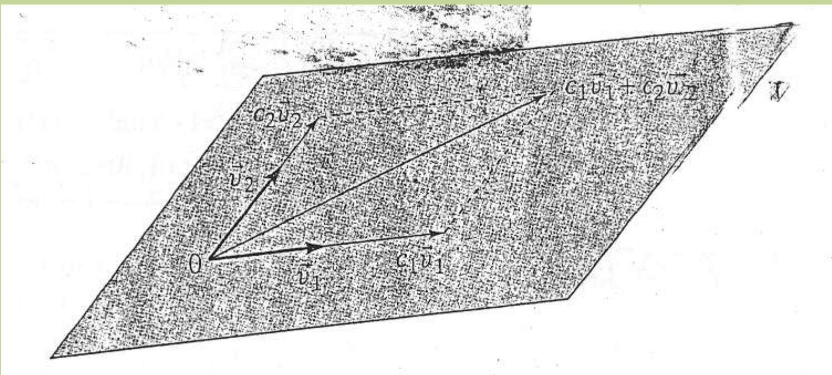


Figure 3.8: -

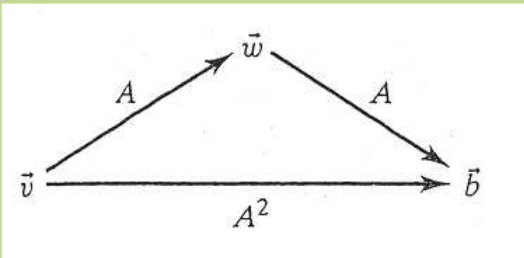


Figure 3.9: Illustration du "lemme des images"

Problème 3.1.2. *Soit $A \in M_n(\mathbb{R})$, une matrice carrée de taille $n \times n$. On note A^2 le produit de A par elle-même, i.e. $A^2 = AA$. Montrer que*

$$\mathrm{Im}(A^2) \subset \mathrm{Im}(A),$$

c'est-à-dire que chaque vecteur $\overrightarrow{b} \in \mathrm{Im}(A^2)$ est aussi dans $\mathrm{Im}(A)$.

Solution. Soit $\overrightarrow{b} \in \mathrm{Im}(A^2)$. Alors il existe $\overrightarrow{v} \in \mathbb{R}^n$ tel que $\overrightarrow{b} = A^2\overrightarrow{v} = A(A\overrightarrow{v})$. Donc en posant $\overrightarrow{w} = A\overrightarrow{v}$, on a $\overrightarrow{b} = A\overrightarrow{w}$, ce qui assure que $\overrightarrow{b} \in \mathrm{Im}(A)$ (voir figure 3.9).

□

3.1.2 Le noyau d'une application linéaire

Lorsque l'on étudie les fonctions réelles $y = f(x)$, on est souvent conduit à chercher les zéro de f (quand par exemple f est la dérivée d'une fonction).

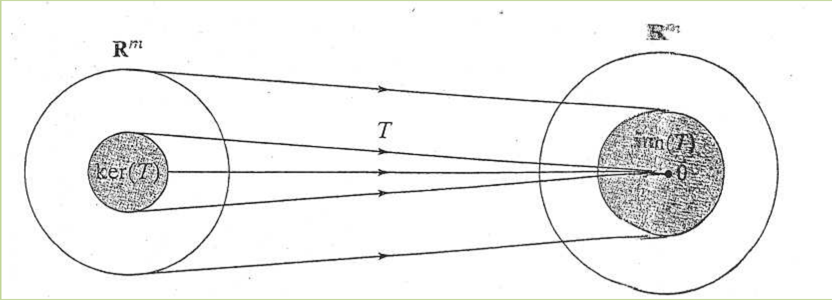


Figure 3.10: Noyau/Image

Par exemple la fonction $y = \sin(x)$ a une infinité de zéros qui sont les $k\pi$, $k \in \mathbb{Z}$.

Dans ce qui suit, on va étudier les zéros des applications linéaires.

Définition 3.1.3 (Noyau d'une application linéaire)**.** *Le noyau d'une application linéaire $T(\overrightarrow{x}) = A\overrightarrow{x}$ est l'ensemble des vecteurs $\overrightarrow{x}$ solutions de l'équation $T(\overrightarrow{x}) = A\overrightarrow{x} = \overrightarrow{0}$. Cet ensemble est le plus souvent noté* Ker(T) *ou encore* Ker(A) *(de l'anglais "Kernel").*

En d'autres termes le noyau de T est l'ensemble des solutions du système linéaire homogène

$$A\overrightarrow{x} = \overrightarrow{0}.$$

Résumé 3.1.2. *Étant donnée une application linéaire $T : \mathbb{R}^m \to \mathbb{R}^n$,*

- *L'image* Im(T) $\subset \mathbb{R}^n$ *est un sous-ensemble de l'ensemble d'arrivée $\mathbb{R}^n$.*
- *Le noyau* Ker(T) $\subset \mathbb{R}^m$ *est un sous-ensemble de l'ensemble de départ $\mathbb{R}^m$.*

Example 3.1.6. *Soit T la projection orthogonale sur le plan x_1, x_2 dans $\mathbb{R}^3$. Son noyau est l'ensemble des solutions de l'équation $T(\overrightarrow{x}) = \overrightarrow{0}$. Cet ensemble est constitué de l'axe x_3, autrement dit* Vect($\overrightarrow{e}_3$).

Exercice 3.1.2. *Trouver le noyau de l'application $T : \mathbb{R}^3 \to \mathbb{R}^2$*

$$T(\overrightarrow{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \overrightarrow{x}.$$

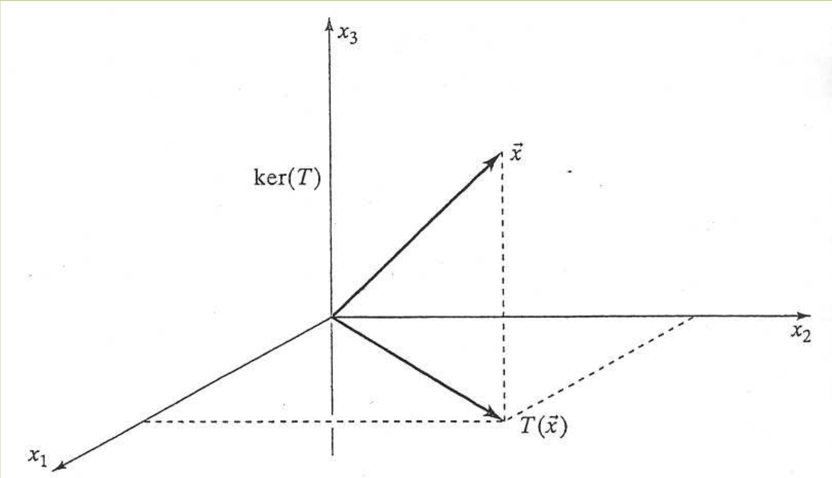


Figure 3.11: Projection orthogonale sur le plan x_1, x_2

Solution. On doit résoudre le système linéaire

$$T(\overrightarrow{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \overrightarrow{x} = \overrightarrow{0}.$$

En utilisant l'algorithme de Gauss-Jordan, on obtient

$$\text{Frel} \left[\begin{array}{ccc:c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array} \right] = \left[\begin{array}{ccc:c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right]$$

Par conséquent les solutions de ce système sont de la forme

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = t \overrightarrow{u}, \quad t \in \mathbb{R}.$$

Donc Ker(T) dans ce cas est la droite engendrée par le vecteur

$$\overrightarrow{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \in \mathbb{R}^3.$$

$\square$

Plus généralement, soit $T(\overrightarrow{x}) = A\overrightarrow{x}$ une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$ et où $m > n$. Dans ce cas, le système $A\overrightarrow{x} = 0$ a toujours des variables libres, et le système admet une infinité de solutions, d'après la proposition 1.3.2. Par conséquent, le noyau contient une infinité de vecteurs. Cela correspond à l'intuition qui nous dit qu'il y aura des "écrasements" si on cherche à plonger linéairement d'une manière ou d'une autre le "grand" espace $\mathbb{R}^m$ dans le "plus petit" espace $\mathbb{R}^n$.

Exercice 3.1.3. *Trouver le noyau de l'application linéaire $T = \mathbb{R}^5 \to \mathbb{R}^4$ définie par $T(\overrightarrow{x}) = A\overrightarrow{x}$, où A est la matrice*

$$\begin{bmatrix} 1 & 5 & 4 & 3 & 2 \\ 1 & 6 & 6 & 6 & 6 \\ 1 & 7 & 8 & 10 & 12 \\ 1 & 6 & 6 & 7 & 8 \end{bmatrix}.$$

Solution. Nous devons résoudre le système linéaire

$$T(\overrightarrow{x}) = A\overrightarrow{x} = \overrightarrow{0}.$$

On vérifie que

$$\mathrm{Frel}(A) = \begin{bmatrix} 1 & 0 & -6 & 0 & 6 \\ 0 & 1 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Il en résulte que le noyau de T est l'ensemble des solutions du système

$$\left| \begin{array}{ccccccccc} x_1 & & & - & 6x_3 & & + & 6x_5 & = & 0 \\ & & x_2 & + & 2x_3 & & - & 2x_5 & = & 0 \\ & & & & & x_4 & + & 2x_5 & = & 0 \end{array} \right|,$$

soit encore

$$\left| \begin{array}{ccccc} x_1 & = & 6x_3 & - & 6x_5 \\ x_2 & = & -2x_3 & + & 2x_5 \\ x_4 & = & & - & 2x_5 \end{array} \right|,$$

système qui admet comme solutions

$$\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 6s - 6t \\ -2s + 2t \\ s \\ -2t \\ t \end{bmatrix} = s \begin{bmatrix} 6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -6 \\ 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} = s\overrightarrow{u} + t\overrightarrow{v},$$

où $s \in \mathbb{R}$ et $t \in \mathbb{R}$ sont arbitraires.

En utilisant la notion "d'espace engendré par une famille de vecteurs" introduite plus haut, on peut écrire

$$\mathrm{Ker}(T) = \mathrm{Vect}(\overrightarrow{u}, \overrightarrow{v}).$$

□

Le noyau a les mêmes propriétés que l'image :

Proposition 3.1.2. *Le noyau d'une application linéaire $T : \mathbb{R}^m \to \mathbb{R}^n$ a les propriétés suivantes :*

(a) $\overrightarrow{0} \in$ *Ker* (T),

(b) $\forall \overrightarrow{v}_1 \in \mathrm{Ker}(T)$, $\forall \overrightarrow{v}_2 \in \mathrm{Ker}(T)$, *alors* $\overrightarrow{v}_1 + \overrightarrow{v}_2 \in \mathrm{Ker}(T)$, *(stabilité de l'addition)*

(c) $\forall \overrightarrow{v} \in \mathrm{Ker}(T)$, $\forall \lambda \in \mathbb{R}$, *alors* $\lambda \overrightarrow{v} \in \mathrm{Ker}(T)$ *(stabilité de la multiplication par un scalaire).*

N.B. Tout espace engendré par une famille de vecteurs vérifie les propriés (a), (b) et (c).

Problème 3.1.3. *Soit* $A \in M_n(\mathbb{R})$ *une matrice carrée de taille* $n \times n$ *inversible. Trouver* $\mathrm{Ker}(A)$.

Solution. Comme A est inversible, on sait que le système $A\overrightarrow{x} = \overrightarrow{0}$ admet une unique solution, $\overrightarrow{x} = \overrightarrow{0}$. Donc $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$.

Réciproquement, si A est non inversible, le système $A\overrightarrow{x} = \overrightarrow{0}$ admet une infinité de solutions, d'aprés la proposition 1.3.2 et donc $\mathrm{Ker}(A) \neq \{\overrightarrow{0}\}$. Comme on a toujours $\{\overrightarrow{0}\} \subset \mathrm{Ker}(A)$, on déduit par contraposée que si $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$, alors A est inversible. □

Résumé 3.1.3. *Soit* $A \in M_n(\mathbb{R})$ *une matrice carrée de taille* $n \times n$. *La matrice* A *est inversible si et seulement si* $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$.

Problème 3.1.4. *Quelles sont les matrices* $A \in M_{n,m}(\mathbb{R})$ *qui vérifient* $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$. *On donnera la réponse en terme du rang de* A.

Solution. On considère le système $A\overrightarrow{x} = \overrightarrow{0}$. Celui-ci admet $\overrightarrow{0}$ comme unique solution si et seulement si il n'y pas de variable libre, d'après la proposition 1.3.2, et donc si et seulement si $\mathrm{Rang}(A) = m$. □

En résumé :

Résumé 3.1.4. *Soit* $A \in M_{n,m}(\mathbb{R})$. *Alors* $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$ *si et seulement si* $\mathrm{Rang}(A) = m$ *(ce qui implique que* $m \leq n$ *puisque* $m = \mathrm{Rang}(A) \leq n$*).*

Soit $A \in M_n(\mathbb{R})$ *une matrice carrée de taille* $n \times n$. *La matrice* A *est inversible si et seulement si* $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$.

Résumé 3.1.5 (Différentes caractéristiques des matrices inversibles)**.** *Soit* $A \in M_n(\mathbb{R})$ *une matrice carrée de taille* $n \times n$. *Les affirmations suivantes sont équivalentes.*

(a) A *est inversible,*

(b) Le système linéaire $A\overrightarrow{x} = \overrightarrow{b}$ *admet une unique solution pour chaque* $\overrightarrow{b} \in \mathbb{R}^n$,

(c) Le système linéaire $A\overrightarrow{x} = \overrightarrow{b}_0$ *admet une unique solution pour un certain* $\overrightarrow{b}_0 \in \mathbb{R}^n$,

(d) $\mathrm{Frel}(A) = \mathrm{I}_n$,

(e) $\mathrm{Rang}(A) = n$,

(f) $\mathrm{Im}(A) = \mathbb{R}^n$,

(g) $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$.

Pour les matrices 2×2 on avait introduit le déterminant qui donnait une autre condition d'inversibilité pour une matrice (voir proposition 2.3.3).

Résumé 3.1.6 (Inversibilité et déterminant d'une matrice 2×2)**.** *La matrice* 2×2

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

est inversible si et seulement si le **déterminant** *de la matrice* A

$$\mathrm{Det}(A) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

est non nul.

Pour les matrices 3×3 on a un résultat similaire

Proposition 3.1.3 (Inversibilité et déterminant d'une matrice 3×3)**.** *La matrice* 3×3

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

est inversible si et seulement si le **déterminant** *de la matrice* A

$$\begin{aligned} \mathrm{Det}(A) &= \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} \\ &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ &\quad -a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} \end{aligned}$$

est non nul.

Démonstration. On note $\overrightarrow{v}_1$ $\overrightarrow{v}_2$ et $\overrightarrow{v}_3$ les trois vecteurs colonnes de A

Si $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ sont colinéaires, $\mathrm{Ker}(A) \neq \{\overrightarrow{0}\}$, A n'est pas inversible et $\mathrm{Det}(A) = 0$.

Si $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ ne sont pas colinéaires, on remarque que

$$\begin{aligned} &a_{1,1}a_{2,2}x_3 + a_{1,2}a_{2,3}x_1 + a_{1,3}a_{2,1}x_2 \\ &-a_{1,2}a_{2,1}x_3 - a_{1,3}a_{2,2}x_1 - a_{1,1}a_{2,3}x_2 \end{aligned}$$

est une équation linéaire dont tous les coefficients ne sont pas nuls. C'est l'équation du plan engendré par $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$. Donc $\overrightarrow{v}_3$ est dans le plan engendré par $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ si et seulement si $\mathrm{Det}(A) = 0$. Finalement, $\mathrm{Im}(A) \neq \mathbb{R}^3$ si et seulement si $\mathrm{Det}(A) = 0$. □

3.2 Sous-espaces vectoriel de $\mathbb{R}^n$, bases et indépendance linéaire

3.2.1 Sous-espaces vectoriels

On a vu dans la section 3.1 que l'image et le noyau d'une application linéaire avaient en commun :

(a) ils contiennent le vecteur nul (du domaine pour le noyau, du co-domaine pour l'image),

(b) ils sont stables pour l'addition

(c) ils sont stables pour la multiplication "externe" par un scalaire.

Définition 3.2.1 (Sous-espace vectoriel de l'espace vectoriel $\mathbb{R}^n$)**.** *Un* **sous-espace vectoriel de** $\mathbb{R}^n$ *est un sous-ensemble de* $\mathbb{R}^n$ *qui vérifie les trois propriétés suivantes :*

(a) il contient le vecteur nul ,

(b) il est stable pour l'addition

(c) il est stable pour la multiplication "externe" par un scalaire.

Les deux dernières propriétés montrent qu'une sous-espace vectoriel W est stable par combinaison linéaire. C'est-à-dire que pour chaques $\overrightarrow{w}_1 \in W$, $\overrightarrow{w}_2 \in W$, $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in \mathbb{R}$, alors $\lambda_1 \overrightarrow{w}_1 + \lambda_2 \overrightarrow{w}_2 \in W$.

Les résultats de la section 3.1 montrent les résultats suivants.

Résumé 3.2.1 (Image et noyau sont des sous-espaces vectoriels)**.** *Soit* $T : \mathbb{R}^m \to \mathbb{R}^n$ *une application linéaire. Alors*

- $\mathrm{Im}(T)$ *est un sous-espace vectoriel de* $\mathbb{R}^n$*,*
- $\mathrm{Ker}(T)$ *est un sous-espace vectoriel de* $\mathbb{R}^m$*.*

Le terme **espace vectoriel** peut être surprenant dans ce contexte. Nous avons vu dans les exemples que noyaux et images pouvaient être des droites ou des plans, que l'on n'appelle pas d'habitude des "espaces". Dans les mathématiques actuelles, le terme "espace vectoriel" est utilisé dans un contexte général et abstrait. Ce terme est souvent utilisé pour décrire des structures qui ne sont pas toujours en rapport avec notre "espace vectoriel physique" qui est $\mathbb{R}^3$. Par exemple, le terme "espace vectoriel" sera utilisé ici pour décrire $\mathbb{R}^n$ pour toute valeur de n, et pas seulement $n = 3$.

Exercice 3.2.1. *Soit* $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2,\ x \geq 0,\ y \geq 0 \right\}$. *Est-ce que* W *est un sous-espace vectoriel de* $\mathbb{R}^2$ *?*

Solution. L'ensemble W est constitué de tous les vecteurs dans le premier quadrant du plan x, y. Ce n'est pas un sous-espace vectoriel de $\mathbb{R}^2$ car non stable par la multiplication par un scalaire négatif.

$\square$

Problème 3.2.1. *Montrer que les seuls sous-espaces vectoriels de* $\mathbb{R}^2$ *sont*

- $\mathbb{R}^2$ *lui-même,*

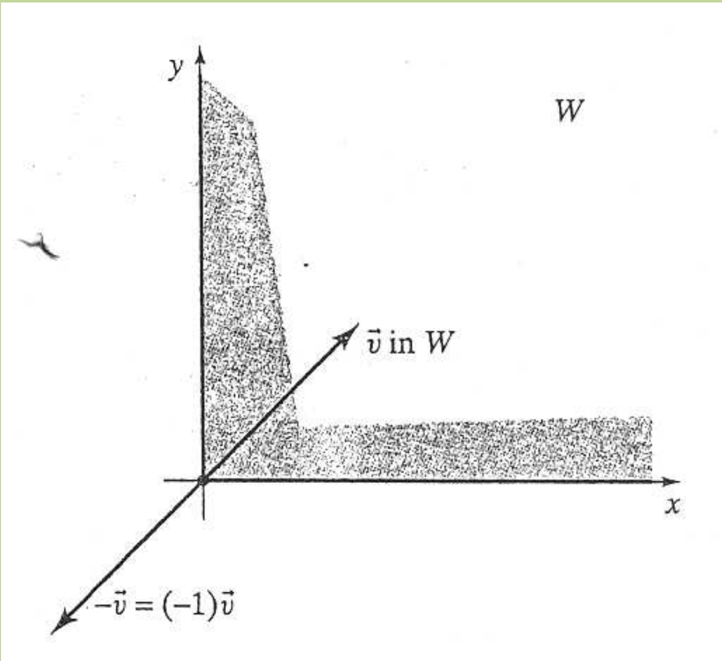


Figure 3.12: L'ensemble n'est pas stable par la multiplication par (-1).

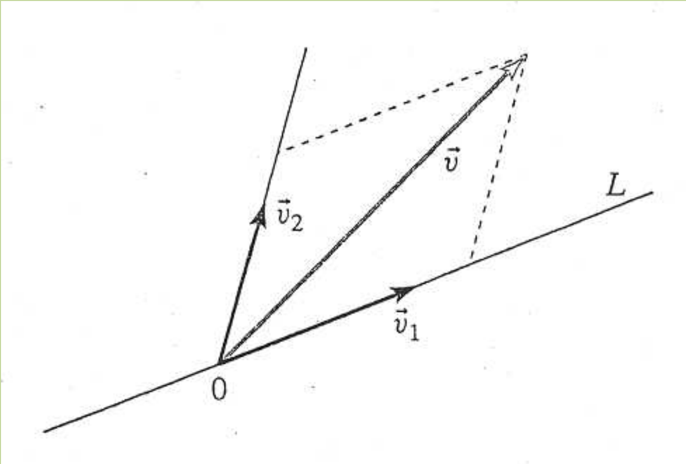


Figure 3.13: -

- $\{\overrightarrow{0}\}$,
- *toutes les droites passant par l'origine.*

Solution. On commence par remarquer que $\{\overrightarrow{0}\}$ est un sous-espace vectoriel de $\mathbb{R}^2$ (ce qui est facile à vérifier) et que $\mathbb{R}^2$ est bien un sous-espace vectoriel de lui-même.

Soit D une droite dans $\mathbb{R}^2$ passant par l'origine. Par définition, il existe un vecteur $\overrightarrow{u} \neq 0$ tel que $D = \text{Vect}(\overrightarrow{u})$. On vérifie facilement qu'il sagit bien d'un sous-espace vectoriel de $\mathbb{R}^2$.

Soit à présent W un sous-espace vectoriel de $\mathbb{R}^2$ qui ne soit ni une droite passant par l'origine, et qui ne soit pas réduit au vecteur nul, $W \neq \{0\}$. Donc il existe $\overrightarrow{v}_1 \neq 0$ tel que $\overrightarrow{v}_1 \in W$. Puisque W est stable par la multiplication externe par un scalaire, $D = \text{Vect}(\overrightarrow{v}_1) \subset W$. Comme W n'est pas une droite, W n'est pas réduit à D. Donc il existe $\overrightarrow{v}_2 \neq 0$, tel que $\overrightarrow{v}_2 \in W$, $\overrightarrow{v}_2 \notin D$. Soit $\overrightarrow{v}$ un vecteur non nul de $\mathbb{R}^2$. En utilisant un parallèlogramme, on peut écrire $\overrightarrow{v}$ comme combinaison linéaire de $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$. Comme W est stable par combinaisons linéaires, $\overrightarrow{v} \in W$, donc $\mathbb{R}^2 \subset W$ et finalement $W = \mathbb{R}^2$. $\square$

De même, les sous-espaces vectoriels de $\mathbb{R}^3$ sont

- $\{\overrightarrow{0}\}$ et $\mathbb{R}^3$ lui même,

- Les droites passant par l'origine (sous-espace vectoriel de dimension 1),
- Les plans passant par l'origine (sous-espace vectoriel de dimension 2),

On notera que la hiérarchie des sous-espaces vectoriels est liée à la **dimension** (concept qui sera explicite dans la section suivante).

	Sous-Espaces de $\mathbb{R}^2$	Sous-Espaces de $\mathbb{R}^3$
dimension 0	$\{\overrightarrow{0}\}$	$\{\overrightarrow{0}\}$
dimension 1	droites passant par $\overrightarrow{0}$	droites passant par $\overrightarrow{0}$
dimension 2	$\mathbb{R}^2$	plans passant par $\overrightarrow{0}$
dimension 3	aucun	$\mathbb{R}^3$

Exercice 3.2.2. *On considère le plan $V \subset \mathbb{R}^3$ donné par l'équation $x_1 + 2x_2 + 3x_3 = 0$.*

(a) Trouver une matrice A telle que $V = \mathrm{Ker}(A)$.

(b) Trouver une matrice B telle que $V = \mathrm{Im}(A)$.

Solution. (a) L'équation du plan peut sécrire $\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = 0$. Par conséquent $V = \mathrm{Ker}\begin{bmatrix}1 & 2 & 3\end{bmatrix}$.

(b) L'image d'une matrice est engendrée par ses vecteurs colonnes. Il suffit donc de décrire (si possible) V comme sous-espace vectoriel engendré par une famille de vecteurs.

Ici, comme la matrice est déjà réduite échelonnée par ligne, et qu'il y a deux variables libres x_2, x_3 on peut d'écrire toutes les vecteurs solutions de l'équation $x_1 + 2x_2 + 3x_3 = 0$ sous la forme forme

$$x_2\begin{bmatrix}-2\\1\\0\end{bmatrix} + x_3\begin{bmatrix}-3\\0\\1\end{bmatrix}.$$

Par conséquent,

$$V = \mathrm{Im}\begin{bmatrix}-2 & -3\\1 & 0\\0 & 1\end{bmatrix}.$$

Ceci est conforme à l'intuition géométrique: pour un plan, deux vecteurs non parallèles conviennent

□

Les sous-espaces vectoriels de $\mathbb{R}^n$ sont souvent définis par leurs équations, i.e. comme les solutions de systèmes linéaires homogènes (i.e. comme des noyaux d'applications linéaires), soit encore comme engendré par une famille de vecteurs (i.e. comme des images d'applications linéaires).

Parfois un sous-espace vectoriel défini comme un noyau doit être défini comme une image et vice-versa.

Problème 3.2.2. *Décrire le noyau d'une matrice comme l'image d'une autre matrice*

Solution. En utilisant la réduction de Gauss, on peut représenter les solutions du système linéaire homogène comme un sous-espace vectoriel engendré par une famille de vecteurs, i.e. une image, comme suit.

- Si un espace vectoriel V contenu dans $\mathbb{R}^m$ est noyau d'une matrice A à n lignes et m colonnes, on peut le décrire comme image d'une matrice B en calculant la forme réduite échelonnée par ligne de A, ceci donne une liste de variables libres (correspondant aux colonnes sans pivot) et en choisissant une variable libre égale à 1 et toutes les autres variables égales à 0 on obtient des vecteurs qui engendrent le noayau de A, ces vecteurs définissent une matrice B avec n lignes et $m - \text{rang}(A)$ colonnes.

$\square$

Problème 3.2.3. *Décrire l'image d'une matrice comme le noyaux d'une autre matrice*

Solution. L'idée est la suivante: on essaye de calculer l'inverse de la matrice et si on n'y arrive pas, c'est qu'on a trouvé les équations de l'image, donc qu'on a décrit l'image comme un noyau.

- Si un espace vectoriel V contenu dans $\mathbb{R}^n$ est image d'une matrice B à n lignes et m colonnes, on peut le décrire comme noyau d'une matrice A en calculant la forme réduite échelonnée par ligne de $C = [B \vdots \text{Id}_n]$. Les équations de V sont données par les lignes de C correspondant aux lignes nulles dans $\text{Frel}(A)$, donc comme le noyau d'une matrice avec $n - \text{rang}(A)$ lignes et n colonnes.

Cette méthode pour décrire l'image d'une matrice comme le noyau d'une autre est détaillée dans les exercices 3.1.42 et 3.1.43. $\square$

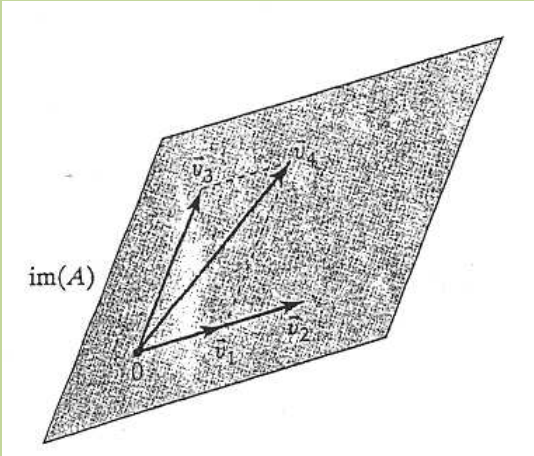


Figure 3.14: Image de A.

3.2.2 Bases et indépendance linéaire

Exercice 3.2.3. *Soit la matrice*

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

Trouver des vecteurs dans $\mathbb{R}^3$ *qui engendrent l'image de* A*. Quel est le plus petit nombre de vecteurs nécessaires pour décrire* $\mathrm{Im}(A)$ *?*

Solution. On sait que $\mathrm{Im}(A)$ est engendré par les vecteurs colonnes de A,

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

La figure 3.14 illustre que l'image de A (dans ce cas particulier) est un plan :

En effet, on note que

$$\vec{v}_2 = 2\vec{v}_1 \in \mathrm{Vect}(\vec{v}_1, \vec{v}_3), \quad \vec{v}_4 = \vec{v}_1 + \vec{v}_3 \in \mathrm{Vect}(\vec{v}_1, \vec{v}_3).$$

Par conséquent les vecteurs $\vec{v}_2$ et $\vec{v}_4$ sont redondants et on a

$$\mathrm{Vect}(\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4) = \mathrm{Vect}(\vec{v}_1, \vec{v}_3).$$

On observe enfin que les vecteurs $\overrightarrow{v}_1$ et $\overrightarrow{v}_3$ ne sont pas parallèles, de sorte que l'image de A est engendrée par deux vecteurs et pas par un unique vecteur. Vérifions par des calculs algébriques que $\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_4) = \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_3)$. On a naturellement $\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_3) \subset \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_4)$. On va vérifier l'inclusion inverse. Soit $\overrightarrow{v} \in \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_4)$ quelconque. Alors $\overrightarrow{v}$ est combinaison linéaire des vecteurs $\overrightarrow{v}_i$, et on a

$$\begin{aligned}
\overrightarrow{v} &= \lambda_1 \overrightarrow{v}_1 + \lambda_2 \overrightarrow{v}_2 + \lambda_3 \overrightarrow{v}_3 + \lambda_4 \overrightarrow{v}_4 \\
&= \lambda_1 \overrightarrow{v}_1 + \lambda_2 (2\overrightarrow{v}_1) + \lambda_3 \overrightarrow{v}_3 + \lambda_4 (\overrightarrow{v}_1 + \overrightarrow{v}_3) \\
&= (\lambda_1 + 2\lambda_2 + \lambda_4)\overrightarrow{v}_1 + (\lambda_3 + \lambda_4)\overrightarrow{v}_3 \in \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_3),
\end{aligned}$$

ce qui achève la preuve. □

L'étude précédente motive les définitions essentielles qui suivent.

Définition 3.2.2 (Vecteurs redondants, indépendance linéaires, bases)**.** *Soit une famille de vecteur $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ dans $\mathbb{R}^n$.*

- *Le vecteur $\overrightarrow{v}_1$ est redondant s'il est nul. Un vecteur $\overrightarrow{v}_j, j > 1$ est redondant si il est combinaison linéaire des vecteurs qui le précèdent dans la liste, $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_{j-1}$.*

- *On dit que les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont* **linéairement indépendants** *si aucun des vecteurs qui consitue la liste est redondant. Sinon, on dit que les vecteurs sont linéairement dépendants.*

- *On dit que la famille de vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ est une* **base** *d'un sous-espace vectoriel V de $\mathbb{R}^n$, si chaque vecteur $\overrightarrow{v}_j$ est dans V, si $V = \mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m)$ et si les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont linéairement indépendants.*

Lorsque que les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont **linéairement indépendants***, on dit aussi que la famille de vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ est* **une famille libre***.*

Lorsque que les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ ne sont pas linéairement indépendants, on dira aussi qu'ils forment une famille **liée***.*

On revient sur l'exercice 3.2.3. Dans la liste

$$\overrightarrow{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \overrightarrow{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad \overrightarrow{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \overrightarrow{v}_4 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

qui constitue la liste des vecteurs colonnes de A, les vecteurs $\overrightarrow{v}_2$ et $\overrightarrow{v}_4$ sont redondants, puisque $\overrightarrow{v}_2 = 2\overrightarrow{v}_1$ et $\overrightarrow{v}_4 = \overrightarrow{v}_1 + \overrightarrow{v}_3$. En retirant les vecteurs redondants, on observe que les vecteurs

$$\overrightarrow{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \overrightarrow{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

sont linéairement indépendants (on vérifie aussi facilement que le système $x\overrightarrow{v}_1 + y\overrightarrow{v}_3 = \overrightarrow{0}$ n'admet que $x = 0$ et $y = 0$ comme unique solution). Donc la famille $(\overrightarrow{v}_1, \overrightarrow{v}_3)$ constitue une base de $V = \mathrm{Im}(A)$.

Résumé 3.2.2 (Base de l'image d'une application linéaire)**.** *Soit A une matrice (ou de manière équivalente on peut raisonner sur l'application linéaire qu'elle représente). On obtient une base de l'image de A,* $\mathrm{Im}(A)$*, en retirant de la famillle des vecteurs colonnes de A tous les vecteurs redondants.*

Il se pose alors la question de savoir comment faire d'un point de vue pratique pour trouver les vecteurs redondants. Dans les cas faciles, on peut le faire à l'aide d'observations élémentaires. On développe dans la suite un algorithme systématique pour répondre à cette question. cet algorithme est basé sur la réduction de Gauss-Jordan.

Exercice 3.2.4. *Est-ce que la famille de vecteurs ci-dessous dans $\mathbb{R}^7$ est une famille libre, autrement dit est-ce que les vecteurs ci-dessous sont linéairement indépendants ?*

$$\overrightarrow{v}_1 = \begin{bmatrix} 7 \\ 0 \\ 4 \\ 0 \\ 1 \\ 9 \\ 0 \end{bmatrix}, \quad \overrightarrow{v}_2 = \begin{bmatrix} 6 \\ 0 \\ 7 \\ 1 \\ 4 \\ 8 \\ 0 \end{bmatrix}, \quad \overrightarrow{v}_3 = \begin{bmatrix} 5 \\ 0 \\ 6 \\ 2 \\ 3 \\ 1 \\ 7 \end{bmatrix}, \quad \overrightarrow{v}_4 = \begin{bmatrix} 4 \\ 5 \\ 3 \\ 3 \\ 2 \\ 2 \\ 4 \end{bmatrix}.$$

Solution. On cherche si il y a des vecteurs redondants dans la liste de vecteurs. Le vecteur $\overrightarrow{v}_1$ n'est pas redondant car il est différent de $\overrightarrow{0}$. Le vecteur $\overrightarrow{v}_2$ n'est pas un multiple (scalaire) de $\overrightarrow{v}_1$, ce que l'on voit en notant que la $4^{\text{ème}}$ composante de $\overrightarrow{v}_1$ est nulle, ce qui n'est pas le cas de la $4^{\text{ème}}$ composante de $\overrightarrow{v}_2$. Donc $\overrightarrow{v}_2$ n'est pas redondant. Le vecteur $\overrightarrow{v}_3$ ne l'est pas, car les deux dernières composantes de $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$ sont nulles, ce qui n'est pas le cas de celle de $\overrightarrow{v}_3$. Enfin, on voit que la deuxième composante des vecteurs $\overrightarrow{v}_1$, $\overrightarrow{v}_2$ et $\overrightarrow{v}_3$ est nulle, ce qui n'est pas le cas de le deuxième composante de $\overrightarrow{v}_4$, lequel ne peut pas être combinaison linéaire de $\overrightarrow{v}_1$, $\overrightarrow{v}_2$ et $\overrightarrow{v}_3$. Donc $\overrightarrow{v}_4$ n'est pas redondant. Ces vecteurs sont linéairement indépendants. □

Résumé 3.2.3 (Indépendance linéaire et composantes nulles)**.** *Soit une famille de vecteurs*

$$(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m).$$

On suppose que $\overrightarrow{v}_1$ n'est pas le vecteur nul, et que que chaque vecteur $\overrightarrow{v}_j$ admet une composante non nulle à un rang où les composantes correspondantes des vecteurs qui le précèdent, $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_{j-1}$, sont toutes nulles. Alors les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont linéairement indépendants.

On explicite autrement l'hypothèse faite dans l'énoncé qui précède. On note $v_{i,j}$ la $i^{\text{ième}}$ composante du vecteur $\overrightarrow{v}_j$. On suppose que pour chaque j entre 1 et m, il existe un indice i_0 (qui peut dépendre de j) tel que pour chaque $k < j$ on ait $v_{i_0,k} = 0$. Donc dans ce cas la famille est libre.

A titre d'illustration, on reprend l'exemple 3.2.4. Les vecteurs sont les suivants

$$\overrightarrow{v}_1 = \begin{bmatrix} 7 \\ \boxed{0} \\ 4 \\ \boxed{0} \\ 1 \\ 9 \\ \boxed{0} \end{bmatrix}, \quad \overrightarrow{v}_2 = \begin{bmatrix} 6 \\ \boxed{0} \\ 7 \\ \boxed{1} \\ 4 \\ 8 \\ \boxed{0} \end{bmatrix}, \quad \overrightarrow{v}_3 = \begin{bmatrix} 5 \\ \boxed{0} \\ 6 \\ 2 \\ 3 \\ 1 \\ \boxed{7} \end{bmatrix}, \quad \overrightarrow{v}_4 = \begin{bmatrix} 4 \\ \boxed{5} \\ 3 \\ 3 \\ 2 \\ 2 \\ 4 \end{bmatrix},$$

et on a encadré les coefficients correspondants.

Exercice 3.2.5. *Est ce que les vecteurs suivants sont linéairement indépendants ?*

$$\overrightarrow{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \overrightarrow{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \quad \overrightarrow{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

Solution. Dans ce cas, on ne peut pas utiliser le critère avec des 0. On note que $\overrightarrow{v}_1 \neq \overrightarrow{0}$, et on vérifie facilement que $\overrightarrow{v}_2$ n'est pas un multiple (scalaire) de $\overrightarrow{v}_1$. Donc ces vecteurs ne sont pas redondants. Pour savoir si $\overrightarrow{v}_3$ est redondant, nous devons étudier le système linéaire

$$\overrightarrow{v}_3 = x\overrightarrow{v}_1 + y\overrightarrow{v}_2,$$

lequel sécrit aussi

$$M = \left[\begin{array}{cc:c} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{array}\right], \quad \mathrm{Frel}(M) = \left[\begin{array}{cc:r} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right],$$

qui montre que ce système admet une et une seule solution $x = -1$ et $y = 2$, de sorte que

$$\overrightarrow{v}_3 = -\overrightarrow{v}_1 + 2\overrightarrow{v}_2.$$

Il en résulte que $\overrightarrow{v}_3$ est redondant et que les vecteurs $\overrightarrow{v}_1$, $\overrightarrow{v}_2$ et $\overrightarrow{v}_3$ sont linéairement dépendants. On dit aussi que la famille $(\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3)$ est une famille liée.

Pour de nombreuses raisons, les mathématiciens aiment écrirent leurs équations sous la forme

$$(\text{quelque chose}) = 0.$$

Dans le cas précédent, l'équation s'écrit

$$\overrightarrow{v}_1 - 2\overrightarrow{v}_2 + \overrightarrow{v}_3 = 0.$$

On appelle cette équation une **relation linéaire** portant sur les vecteurs $\overrightarrow{v}_1$, $\overrightarrow{v}_2$ et $\overrightarrow{v}_3$. □

Définition 3.2.3 (Relation linéaire)**.** *Soit $(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m)$ une famille de vecteurs. Une équation de la forme*

$$\lambda_1\overrightarrow{v}_1 + \lambda_2\overrightarrow{v}_2 + \cdots + \lambda_m\overrightarrow{v}_m = \overrightarrow{0}$$

est appelée **une relation linéaire** *entre les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$. Si les seuls coefficients λ_i satisfaisant cette relation sont tous égaux à zéro, on dit que cette relation est* **triviale**. *On dit que la relation est* **non triviale** *si il y a au moins un coefficient λ_i non nul.*

Proposition 3.2.1 (Relations et dépendance linéaire)**.** *Les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont linéairement dépendants si et seulement si il existe entre eux une relation linéaire non triviale.*

Démonstration. On suppose que les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ sont linéairement dépendants. Donc il existe un indice j tel que le vecteur $\overrightarrow{v}_j$ soit redondant. Il existe donc des scalaires λ_i tels que $\overrightarrow{v}_j = \lambda_{j-1}\overrightarrow{v}_{j-1} + \cdots + \lambda_1\overrightarrow{v}_1$ que l'on peut écrire encore $(-1)\overrightarrow{v}_j + \lambda_{j-1}\overrightarrow{v}_{j-1} + \cdots + \lambda_1\overrightarrow{v}_1 = \overrightarrow{0}$, ce qui est une relation non triviale puisque $-1 \neq 0$.

Réciproquement, supposons qu'il y ait entre les vecteurs $\overrightarrow{v}_j$ une relation non triviale $\lambda_m\overrightarrow{v}_m + \lambda_{m-1}\overrightarrow{v}_{m-1} + \cdots + \lambda_1\overrightarrow{v}_1 = \overrightarrow{0}$. Soit j le plus **grand indice** k tel que $\lambda_k \neq 0$. On peut écrire

$$\overrightarrow{v}_j = -\frac{\lambda_1}{\lambda_j}\overrightarrow{v}_1 - \frac{\lambda_2}{\lambda_j}\overrightarrow{v}_2 - \cdots - \frac{\lambda_{j-1}}{\lambda_j}\overrightarrow{v}_{j-1}.$$

Donc le vecteur $\overrightarrow{v}_j$ est redondant, ce qui achève la preuve. □

Problème 3.2.4. *Soit une matrice de taille $n \times m$. On suppose que les vecteurs colonne de A sont linéairement indépendants. Déterminer le noyau de A.*

Solution. Il faut résoudre l'équation $A\overrightarrow{x} = \overrightarrow{0}$, c'est-à-dire

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \text{ ou encore } \quad x_1\overrightarrow{v}_1 + \cdots + x_m\overrightarrow{v}_m = \overrightarrow{0}.$$

On constate que trouver le noyau de A, revient à étudier une relation linéaire entre les vecteurs colonnes de la matrice. Les vecteurs colonnes de A étant supposés indépendants, le système plus haut admet une unique solution qui est la solution triviale, ce qui fait que $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$. □

Résumé 3.2.4 (Noyau et relations linéaires)**.** *Soit une matrice de taille $n \times m$. Les vecteurs contenus dans le noyau de A correspondent à une relation linéaire entre les vecteurs colonnes de A, $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$. L'équation*

$$A\overrightarrow{x} = \overrightarrow{0} \quad \textit{est équivalente à} \quad x_1\overrightarrow{v}_1 + \cdots + x_m\overrightarrow{v}_m = \overrightarrow{0}.$$

En particulier, les vecteurs colonnes de A sont linéairement indépendants si et seulement si $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$*, ou bien encore de manière équivalente* $\mathrm{Rang}(A) = m$*. Cette condition implique en particulier que $m \leq n$.*

Donc : on ne peut pas trouver plus de n vecteurs linéairement indépendants dans $\mathbb{R}^n$.

Example 3.2.1. *Soit A la matrice*

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}.$$

$$\textit{vecteur redondant} \qquad \begin{bmatrix}7\\8\\9\end{bmatrix} = -\begin{bmatrix}1\\2\\3\end{bmatrix} + 2\begin{bmatrix}4\\5\\6\end{bmatrix},$$

$$\textit{relation linéaire} \quad 1\begin{bmatrix}1\\2\\3\end{bmatrix} - 2\begin{bmatrix}4\\5\\6\end{bmatrix} + 1\begin{bmatrix}7\\8\\9\end{bmatrix} = \overrightarrow{0},$$

$$\begin{bmatrix}1\\-2\\1\end{bmatrix} \in \operatorname{Ker}(A) \quad \textit{car} \quad \begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Résumé 3.2.5 (Caractérisation de l'indépendance linéaire). *Soit*

$$\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$$

une famille de vecteurs dans $\mathbb{R}^n$. *Les énoncés suivants sont équivalents.*

(a) Les vecteurs $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ *sont linéairement indépendants, (on dit aussi que la famille* $\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$ *est une famille libre.)*

(b) Aucun des vecteurs de la famille est redondant, à savoir qu'aucun vecteur n'est combinaison linéaire des vecteurs qui le précèdent dans l'énumération de la liste.

(c) Aucun des vecteurs $\overrightarrow{v}_j$ *n'est une combinaison linéaire des autres vecteurs* $\overrightarrow{v}_1, \cdot, \overrightarrow{v}_{j-1}, \overrightarrow{v}_{j+1}, \cdots, \overrightarrow{v}_m$.

(d) $\operatorname{Ker}\begin{bmatrix}\overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m\end{bmatrix} = \{\overrightarrow{0}\}.$

e. Le système $x_1\overrightarrow{v}_1 + \cdots + x_m\overrightarrow{v}_m = \overrightarrow{0}$ *n'admet que le vecteur nul pour solution, à savoir* $x_1 = 0, \cdots, x_m = 0$.

(f) $\operatorname{Rang}\begin{bmatrix}\overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m\end{bmatrix} = m.$

Donc: une famille libre ne peut jamais contenir le vecteur nul !

On termine cette section par une caractérisation utile des bases.

Problème 3.2.5. *Soit* $\mathcal{B} = (\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m)$ *une base d'un sous-espace vectoriel* V *de* $\mathbb{R}^n$ *et soit* $\overrightarrow{v} \in V$*. Combien de solutions* $c_1, \cdots, c_m$ *l'équation*

$$\overrightarrow{v} = c_1\overrightarrow{v}_1 + \cdots + c_m\overrightarrow{v}_m \tag{3.1}$$

admet-elle ?

Solution. Puisque $V = \text{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m)$ et que $\overrightarrow{v} \in V$, par définition, l'équation (3.1) admet au moins une solution $c_1, \cdots, c_m$. Soit $d_1, \cdots, d_m$ une autre solution éventuelle de l'équation (3.1) permettant de décomposer $\overrightarrow{v}$ suivant la base $\mathcal{B}$, à savoir

$$\overrightarrow{v} = d_1\overrightarrow{v}_1 + \cdots + d_m\overrightarrow{v}_m \tag{3.2}$$

En soustrayant (3.2) à (3.1), on obtient

$$(c_1 - d_1)\overrightarrow{v}_1 + \cdots + (c_m - d_m)\overrightarrow{v}_m = \overrightarrow{0}. \tag{3.3}$$

Puisque la famille $\mathcal{B}$ est une famille libre, on déduit de l'équation (3.3) que $c_1 - d_1 = 0, \cdots, c_m - d_m = 0$. Par conséquent les représentations données par (3.1) et (3.2) sont identiques. □

On a le résultat suivant.

Proposition 3.2.2 (Bases et unicité des représentations)**.** *Soit*

$$\mathcal{B} = \overrightarrow{v}_1, \overrightarrow{v}_2, \cdots, \overrightarrow{v}_m$$

une famille de vecteurs contenus dans un sous-espace vectoriel V *de* $\mathbb{R}^n$*.*

La famille $\mathcal{B}$ *est une base de* V *si et seulement si chaque vecteur* $\overrightarrow{v}$ *peut être décomposé de manière unique comme combinaison linéaire des vecteurs de* $\mathcal{B}$*, à savoir il existe un unique "m-uplet"* $c_1, \cdots, c_m$ *de nombres réels tel que*

$$\overrightarrow{v} = c_1\overrightarrow{v}_1 + \cdots + c_m\overrightarrow{v}_m.$$

Dans la suite on dira que les c_i *sont les coordonnées du vecteur* $\overrightarrow{v}$ *dans la base* $\mathcal{B}$*.*

Démonstration. On a déjà prouvé un sens dans cette afirmation. Il reste à prouver la réciproque, à savoir que l'unicité de la décomposition de chaque vecteur de V comme combinaison linéaire des vecteurs de $\mathcal{B}$ entraine que $\mathcal{B}$ est bien une base.

L'hypothèse entraine d'ores et déjà que $V = \mathrm{Vect}(\mathcal{B})$. Il reste à montrer que $\mathcal{B}$ est une famille libre pour prouver que c'est bien une base de V. Pour cela, considérons l'équation

$$c_1\overrightarrow{v}_1 + \cdots + c_m\overrightarrow{v}_m = \overrightarrow{0}. \tag{3.4}$$

Comme $\overrightarrow{0} \in V$, il se décompose de manière unique suivant $\mathcal{B}$,

$$\overrightarrow{0} = 0\overrightarrow{v}_1 + \cdots + 0\overrightarrow{v}_m.$$

Or $c_1 = 0, \cdots, c_m = 0$ est solution de (3.4), et l'unicité de la décomposition de $\overrightarrow{0}$ suivant $\mathcal{B}$ assure que la solution nulle est la seule de (3.4) et donc que $\mathcal{B}$ est libre, et donc une base de V puisque c'est déjà une famille génératrice de $\mathcal{B}$. □

3.3 Dimension d'un sous-espace vectoriel de $\mathbb{R}^n$

Soit V un plan dans $\mathbb{R}^3$. Intuitivement, on observe que toute base de V est constituée de deux vecteurs. En fait chaque paire de vecteurs non parallèles dans V fera l'affaire (voir figure 3.15).

En effet, un seul vecteur est insuffisant pour engendrer V et trois ou plus vecteurs dans V sont linéairement dépendants et donc ne forment pas une base de V.

3.3.1 Théorème de la dimension

On conjecture donc d'une manière générale que, étant donné un sous-espace vectoriel V de $\mathbb{R}^n$, chaque base de V possède le même nombre d'éléments. Cette conjecture est vraie, et pour démontrer ce théorème fondamental, on a besoin du lemme intermédiaire suivant.

Lemme 3.3.1. *Soit V un sous-espace vectoriel V de $\mathbb{R}^n$. On considère deux familles de vecteurs dans V, $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ et $\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q$. On suppose que la famille $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ est une famille libre et que la famille $\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q$ et génératrice, i.e. que $V = \mathrm{Vect}(\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q)$. Alors on a $p \leq q$.*

Par exemple, soit V un plan dans $\mathbb{R}^3$. Intuitivement on constate que l'on peut trouver au plus deux vecteurs linéairement indépendants dans V, mais par ailleurs que l'on a besoin d'au moins deux vecteurs pour engendrer V. Donc $q \geq 2$ tandis-que $p \leq 2$. Donc l'inégalité $p \leq q$ semble raisonnable dans ce cas.

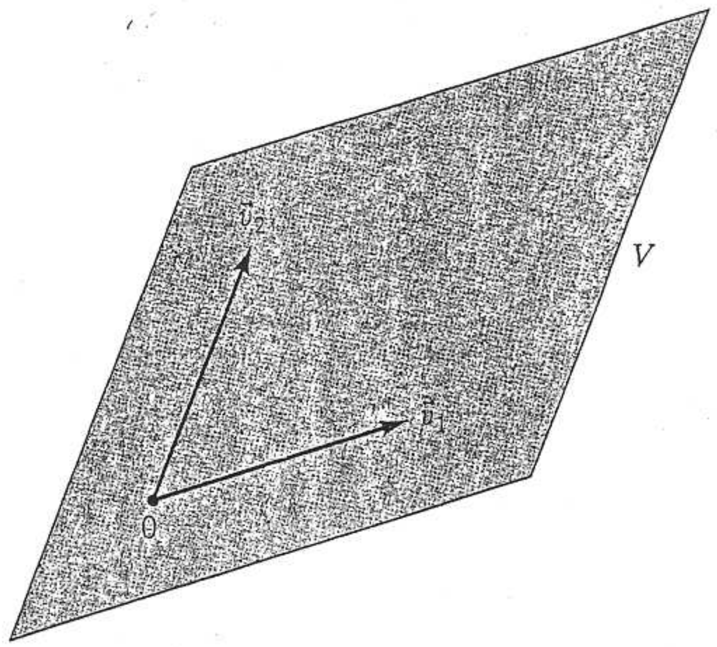


Figure 1 The vectors $\vec{v}_1$, $\vec{v}_2$ form a basis of V.

Figure 3.15: -

Démonstration. La preuve donnée ici est technique et même frustrante. Mais elle a le mérite de fonctionner avec les résultats développés dans ce cours.

Considérons les matrices

$$A = \begin{bmatrix} \overrightarrow{w}_1 & \overrightarrow{w}_2 & \cdots & \overrightarrow{w}_q \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_p \end{bmatrix}$$

Puisque $V = \text{Vect}(\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q)$, alors $\text{Im}(A) = V$. Par ailleurs, les vecteurs $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ sont dans $\text{Im}(A)$. Par conséquent il existe $\overrightarrow{u}_1, \cdots, \overrightarrow{u}_p$ des vecteurs dans $\mathbb{R}^q$ tels que

$$\overrightarrow{v}_1 = A\overrightarrow{u}_1, \cdots, \overrightarrow{v}_p = A\overrightarrow{u}_p.$$

On peut combiner ces équations, définir une matrice à q lignes et p colonnes

$$C = \begin{bmatrix} \overrightarrow{u}_1 & \overrightarrow{u}_2 & \cdots & \overrightarrow{u}_p \end{bmatrix}$$

et écrire

$$B = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_p \end{bmatrix} = A \begin{bmatrix} \overrightarrow{u}_1 & \overrightarrow{u}_2 & \cdots & \overrightarrow{u}_p \end{bmatrix} = A\,C$$

Le noyau de C est un sous-espace vectoriel du noyau de B : si $C\overrightarrow{x} = \overrightarrow{0}$, alors $B\overrightarrow{x} = A\,C\overrightarrow{x} = \overrightarrow{0}$. Mais comme les vecteurs $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ sont linéairement indépendants, $\text{Ker}(B) = \{\overrightarrow{0}\}$, de sorte que l'on a également $\text{Ker}(C) = \{\overrightarrow{0}\}$. Cela implique que la matrice C admet plus de lignes que de colonnes, c'est-à-dire $q \geq p$ comme annoncé. □

Théorème 6 (Théorème de la dimension.)**.** *Soit V un sous-espace vectoriel de $\mathbb{R}^n$. Chaque base de V admet le même nombre de vecteurs.*

Démonstration. On rappelle qu'une base est une famille libre et génératrice. Soit $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ et $\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q$ deux bases de V. Puisque $V = \text{Vect}(\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q)$ et que $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ est libre, alors

$$p \leq q.$$

Puisque $V = \text{Vect}(\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p)$ et que $\overrightarrow{w}_1, \cdots, \overrightarrow{w}_q$ est libre, alors

$$q \leq p.$$

Par conséquent,

$$p = q,$$

ce qui achève la démonstration. $\square$

Soient D et P respectivement une droite et un plan dans $\mathbb{R}^3$. Une base de D est constituée d'un vecteur non nul (n'importe quel vecteur non nul de D fait l'affaire), tandis-qu'une base de P est constituée de deux vecteurs. Une base de $\mathbb{R}^3$ est constituée de trois vecteurs (la base canonique déjà rencontrée, $\overrightarrow{e}_1$, $\overrightarrow{e}_2$, $\overrightarrow{e}_3$ est un choix possible). Dans chaque cas, le nombre de vecteurs de la base correspond à l'idée intuitive que l'on se fait de la notion de dimension.

Théorème 7. *Tout sous-espace vectorielde $\mathbb{R}^n$ admet une base.*

Démonstration. Soit V un sous-espace vectoriel de $\mathbb{R}^n$. Si $V \neq \{\overrightarrow{0}\}$, choisissons $\overrightarrow{v}_1$ non nul dans V. Si $\mathrm{Vect}(\overrightarrow{v}_1)$ n'est pas égal à V on peut chosir un vecteur $\overrightarrow{v}_2$ dans V qui ne soit pas dans la droite $\mathrm{Vect}(\overrightarrow{v}_1)$. Si $\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2)$ n'est pas égal à V on peut chosir un vecteur $\overrightarrow{v}_3$ dans V qui ne soit pas dans le plan $\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2)$. On peut continuer cette construction .... Elle se termine quand on a atteint une base de V, comprenant un nombre de vecteurs $\leq n$, puisqu'on a déjà vu qu'il ne peut pas y avoir dans $\mathbb{R}^n$ plus de n vecteurs linéairement indépendants. $\square$

Définition 3.3.1 (Dimension.). *Soit V un sous-espace vectoriel de $\mathbb{R}^n$. Le nombres de vecteurs constituant une base de V est appelé la dimension de V et notée* $\mathrm{Dim}(V)$.

La notion algébrique de dimension est une avancée majeure dans le dévelopement de l'algèbre linéaire, et dans les mathématiques en général. Cela permet de concevoir des espaces de dimension plus grande que 3. Cela correspond à une idée mystique populaire qui cherche par exemple "la $4^{\text{ième}}$ dimension".

Le mathématicien allemand Hermann Weyl (1855-1955) disait les choses dans ce sens :

"Rien ne nous oblige à chercher l'illumination par des doctrines mystiques ou spirituelles pour avoir une vision claire de la géométrie multidimensionnelle".

Le premier mathématicien qui a pensé à la notion de dimension était le mathématicien français Jean Le Rond d'Alembert (1717-1783). Dans un article sur la dimension dans *l'encyclopédie*, il écrit :

Le fait de considérer des structures ayant plus de 3 dimensions n'est jamais qu'un droit comme un autre, parce des lettres peuvent toujours représenter des nombres, qu'ils soient rationnels ou non. J'ai dit précédemment qu'il nétait pas possible de concevoir plus que 3 dimensions. Un homme d'esprit de ma connaissance croit qu'on pourrait cependant regarder la durée comme une quatrième dimension. Cette idée peut-être contestée, mais elle a, ce me semble, quelque mérite, quand ce serait que celui de la nouveauté.

Cet homme d'esprit était sans doutes d'Alembert, prudent puisqu'il savait qu'en ces temps une telle idée était risquée. Du reste, le non conformisme est toujours très risqué quand il remet des dogmes en question....

L'idée de dimension a été plus tard systématiquement étudiée par le mathématicien allemand Herman Günter Grassmann (1809-1877), qui a introduit la notion de sous-espace vectoriel de $\mathbb{R}^n$. En fait, la plupart des concepts étudiés dans la chapitre 3 suivent le travail de Grassman, présenté dans son livre en 1844 "*Die Lineare Ausdehnunglehre, ein neuer Zweig der Mathematik*" (La théorie de l'extension linéaire, une nouvelle branche des mathématiques). Le livre de Grassman étant mal rédigé, il a fallu du temps pour la diffusion de ses idées. Ce fait est essentiellement du à l'utilisation d'une terminologie allemande peu répandue. Par exemple, il écrivait "*Schatten*" (ombre) plutôt que "projection". Ses idées ont survécues, pas sa terminologie.

Des travaux analogues on été publiés par le mathématicien suisse Ludwig Schläfi (1814-1895), contemporain de Grasmann.

Aujourd'hui, le concept de dimension est central en mathématiques, aussi bien qu'en physique ou qu'en statistiques. Il peut aussi s'adapter à des strutures non linéaires comme les *variétés*, qui généralisent l'idée de courbes et de surfaces.

Après cette digression historique, revenons à des choses plus terre à terre: quelle est la dimension de $\mathbb{R}^n$? Bien entendu, n, puisque l'on sait que les vecteurs $\overrightarrow{e}_1, \cdots, \overrightarrow{e}_n$ forment une base de $\mathbb{R}^n$, la base dite "canonique".

Un plan dans $\mathbb{R}^3$ est de dimension 2. On a déjà dit qu'il fallait au moins 2 vecteurs pour engendrer un plan et que plus de deux vecteurs dans un plan formaient une famille liée. Si deux vecteurs dans un plan sont linéairement indépendants, alors ils forment une base du plan. De même, si deux vecteurs d'un plan engendrent ce plan, alors ils forment une base du plan.

On généralise ces idées comme suit.

Proposition 3.3.1 (Familles libres et familles génératrices d'un sous-espace vectoriel de $\mathbb{R}^n$.). *Soit V sous-espace vectoriel de $\mathbb{R}^n$ de dimension m (on a donc $m \leq n$).*

(a) On peut trouver au plus m vecteurs linéairement indépendants dans V,

(b) On a besoin d'au moins m vecteurs pour engendrer V,

(c) Si m vecteurs de V sont linéairement indépendants, alors ils forment une base de V,

(d) Si m vecteurs de V engendrent V, alors ils forment une base de V.

La partie (a) permet de définir la dimension comme le nombre maximal de vecteurs linéairement indépendants dans V. D'après (b), cela peut aussi être le nombre minimal de vecteurs nécessaires pour engendrer V.

Par définition une base doit être libre et génératrice. Cependant, si on travaille avec le bon nombre vecteurs, il suffit de vérifier l'une ou l'autre des deux propriétés, le reste suit automatiquement.

Démonstration. (a) Soit $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ une famille libre dans V et $\overrightarrow{w}_1, \cdots, \overrightarrow{w}_m$ une base de V. Puisque $V = \text{Vect}(\overrightarrow{w}_1, \cdots, \overrightarrow{w}_m)$, on a $p \leq m$ d'après la Propriété 3.3.1.
(b) Soit $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_p$ une famille génératrice de V, en lui supprimant ses vecteurs redondants on trouve une base de V qui a m éléments, donc $p \geq m$ d'après le Théorème de la Dimension page 3.
(c) Soit $m = \text{Dim}(V)$, $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$ une famille libre dans V. Nous devons prouver que $V = \text{Vect}(\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m)$. Soit $\overrightarrow{v}$ un vecteur de V. D'après (a), les $m+1$ vecteurs $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m, \overrightarrow{v}$ sont linéairement dépendants. Comme les vecteurs $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$ sont linéairement indépendants, il n'y a pas de vecteurs redondants dans cette liste, et donc seul $\overrightarrow{v}$ est redondant dans la famille $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m, \overrightarrow{v}$. Donc $\overrightarrow{v}$ est combinaison linéaire des vecteurs $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$. Ceci étant vrai pour chaque $\overrightarrow{v} \in V$, on en déduit que

$$V = \text{Vect}(\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m),$$

et donc $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$ est bien une base de V.
(d) Soient $(\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m)$ une famille génératrice de V. Si elle contenait un vecteur redondant, on trouverait une base de V qui avec au plus $m-1$ éléments, ce qui est impossible d'après le Théorème de la Dimension. Donc $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$ est une base. □

Dans ce qui suit on étudie comment trouver des bases des images et noyaux d'applications linéaires, et donc leur dimension.

Exercice 3.3.1. *Trouver une base respectivement du noyau et de l'image de l'application linéaire représentée par la matrice*

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 1 & 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 4 & 5 \end{bmatrix}$$

Solution. On sait que l'image de A est engendrée par ses vecteurs colonnes. Pour trouver une base de l'image, nous devons trouver les vecteurs redondants parmis les vecteurs colonnes de la matrice A, notés $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_5$. Dans cet exemple simple, on observe tout de suite que

$$\begin{aligned} \overrightarrow{v}_2 &= 2\overrightarrow{v}_1, \\ \overrightarrow{v}_3 &= \overrightarrow{0}, \\ \overrightarrow{v}_5 &= \overrightarrow{v}_1 + \overrightarrow{v}_4. \end{aligned}$$

On vérifie facilement que l'équation $x\overrightarrow{v}_1 + y\overrightarrow{v}_4 = \overrightarrow{0}$ n'admet que $x = y = 0$ comme solution, ce qui fait que la famille $(\overrightarrow{v}_1, \overrightarrow{v}_4)$ est une famille libre. Par conséquent la famille $(\overrightarrow{v}_1, \overrightarrow{v}_4)$ est une base de $\mathrm{Im}(A)$.

Pour trouver une base du noyau de A, on applique la méthode de l'exercice 3.2.4 de la section 3.2 aux vecteurs redondants $\overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_5$, ce qui va nous permettre de construire 3 vecteurs dans le noyau de A. On organise notre travail suivant une table.

Vecteurs redondants	**Relations**	**Vecteurs dans le noyau**
$\overrightarrow{v}_2 = 2\overrightarrow{v}_1$	$-2\overrightarrow{v}_1 + \overrightarrow{v}_2 = \overrightarrow{0}$	$\overrightarrow{w}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
$\overrightarrow{v}_3 = \overrightarrow{0}$	$\overrightarrow{v}_3 = \overrightarrow{0}$	$\overrightarrow{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$
$\overrightarrow{v}_5 = \overrightarrow{v}_1 + \overrightarrow{v}_4$	$-\overrightarrow{v}_1 - \overrightarrow{v}_4 + \overrightarrow{v}_5 = \overrightarrow{0}$	$\overrightarrow{w}_5 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$

On affirme que les trois vecteurs que nous avons construit forment une base du noyau de A. Le méthode des zéros développée dans la section précédente permet d'affirmer que ces vecteurs sont indépendants puisque la $i^{\text{ième}}$ composante $\overrightarrow{w}_i$ est égale à 1 tandis-que les autres correspondantes sont nulles.

Nous devons montrer que les trois vecteurs $\overrightarrow{w}_2$, $\overrightarrow{w}_3$ et $\overrightarrow{w}_5$ engendrent bien Ker(A). Soit $\overrightarrow{x}$ un vecteur quelconque du noyau. Il faut prouver qu'il est bien combinaison linéaire de $\overrightarrow{w}_2$, $\overrightarrow{w}_3$ et $\overrightarrow{w}_5$. Il est commode de considérer le vecteur auxiliaire

$$\begin{aligned}\overrightarrow{y} &= \overrightarrow{x} - x_2\overrightarrow{w}_2 - x_3\overrightarrow{w}_3 - x_5\overrightarrow{w}_5\\ &= \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{bmatrix} - \begin{bmatrix} -2x_2\\ x_2\\ 0\\ 0\\ 0\end{bmatrix} - \begin{bmatrix} 0\\ 0\\ x_3\\ 0\\ 0\end{bmatrix} - \begin{bmatrix} -x_5\\ 0\\ 0\\ -x_5\\ x_5\end{bmatrix} = \begin{bmatrix} y_1\\ 0\\ 0\\ y_4\\ 0\end{bmatrix},\end{aligned}$$

où

$$\begin{aligned} y_1 &= x_1 + 2x_2 + x_5,\\ y_4 &= x_4 + x_5.\end{aligned}$$

Puisque $\overrightarrow{y}$ est une combinaison linéaire de vecteurs dans Ker(A), il est lui-même dans Ker(A). Comme $\overrightarrow{x} \in$ Ker(A), on a

$$x_1\overrightarrow{v}_1 + x_2\overrightarrow{v}_2 + x_3\overrightarrow{v}_3 + x_4\overrightarrow{v}_4 + x_5\overrightarrow{v}_5 = \overrightarrow{0},$$

en rappellant que les $\overrightarrow{v}_i$ sont les vecteurs colonnes de la matrice A. En utilisant les relations entre les $\overrightarrow{v}_i$, on déduit de l'égalité plus haut

$$(x_1 + 2x_2 + x_5)\overrightarrow{v}_1 + (x_4 + x_5)\overrightarrow{v}_4 = \overrightarrow{0},$$

autrement dit

$$y_1\overrightarrow{v}_1 + y_4\overrightarrow{v}_4 = \overrightarrow{0}.$$

Comme la famille $(\overrightarrow{v}_1, \overrightarrow{v}_4)$ est une famille libre, on en déduit que $y_1 = y_4 = 0$. Par conséquent, $\overrightarrow{y} = \overrightarrow{0}$ et on a

$$\overrightarrow{x} = x_2\overrightarrow{w}_2 + x_3\overrightarrow{w}_3 + x_5\overrightarrow{w}_5,$$

ce qui implique bien que Ker$(A) =$ Vect$(\overrightarrow{w}_2, \overrightarrow{w}_3, \overrightarrow{w}_5)$ comme annoncé. $\square$

Si on applique la procédure développée dans l'exercice 3.2.4 à une matrice A quelconque de taille $n \times m$, on observe qu'une base du noyau de A contient autant de vecteur qu'il y a de vecteurs redondants que dans la famille des vecteurs colonnes, c'est-à-dire

$$\begin{aligned}&\mathrm{Dim}(\mathrm{Ker}(A)) = \\ &\qquad (\text{nombre des vecteurs redondants dans les vecteur colonnes}).\end{aligned}$$

Comme on a également vu que les vecteurs non redondants dans cette liste forme une base de l'image de A,

$$\begin{aligned}\mathrm{Dim}(\mathrm{Im}(A)) &= \text{nb de vecteurs colonnes non redondants} \\ &= \text{nb de colonnes} - (\text{nb des vecteurs colonnes redondants}) \\ &= m - \mathrm{Dim}(\mathrm{Ker}(A)).\end{aligned}$$

D'où la relation fondamentale satisfaite par toute matrice A de taille $n \times m$,

Théorème 8 (Formule fondamentale de l'algèbre linéaire)**.**

$$\boxed{m = \mathrm{Dim}(\mathrm{Ker}(A)) + \mathrm{Dim}(\mathrm{Im}(A))}$$

En principe on doit pouvoir déterminer des bases de l'image et du noyau en utilisant la méthode de l'exercice 3.3.1. Si jamais on ne voit pas rapidement les vecteurs redondants, on peut toujours utiliser l'algorithme de Gauss-Jordan comme l'illustre l'exercice suivant.

Exercice 3.3.2. *Déterminer les bases des noyaux et image de la matrice*

$$A = \begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ -1 & -2 & -1 & 1 & -1 \\ 4 & 8 & 5 & -8 & 9 \\ 3 & 6 & 1 & 5 & -7 \end{bmatrix},$$

en sachant que

$$\mathrm{Frel}(A) = B = \begin{bmatrix} 1 & 2 & 0 & 3 & -4 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Solution. Les vecteurs colonnes de A sont notés $\overrightarrow{a}_1, \cdots, \overrightarrow{a}_5$, tandis-que ceux de B sont notés $\overrightarrow{b}_1, \cdots, \overrightarrow{b}_5$.

On commence par remarquer que $\mathrm{Ker}(A) = \mathrm{Ker}(B)$. En effet, on sait que résoudre le système $A\overrightarrow{x} = \overrightarrow{0}$ est équivalent à résoudre le système $B\overrightarrow{x} = \overrightarrow{0}$. Il suffit alors de procéder comme dans le cas de l'exemple 3.3.1 pour trouver les vecteurs redondants parmis les vecteurs colonnes de B. Dans le cas d'une "Frel" c'est assez facile. Chaque colonne qui ne contient pas de pivot peut être exprimée comme combinaison linéaire de celles qui précèdent et qui contiennent des pivots.

Comme dans l'exemple 3.3.1, on organise le travail avec un tableau.

Vecteurs redondants	**Relations**	**Vecteurs dans le noyau**
$\overrightarrow{b}_2 = 2\overrightarrow{b}_1$	$-2\overrightarrow{b}_1 + \overrightarrow{b}_2 = \overrightarrow{0}$	$\overrightarrow{w}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
$\overrightarrow{b}_4 = 3\overrightarrow{b}_1 - 4\overrightarrow{b}_3$	$-3\overrightarrow{b}_1 + 4\overrightarrow{b}_3 + \overrightarrow{b}_4 = \overrightarrow{0}$	$\overrightarrow{w}_4 = \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix}$
$\overrightarrow{b}_5 = -4\overrightarrow{b}_1 + 5\overrightarrow{b}_3$	$4\overrightarrow{b}_1 - 5\overrightarrow{b}_3 + \overrightarrow{b}_5 = \overrightarrow{0}$	$\overrightarrow{w}_5 = \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix}$

Les vecteurs $\overrightarrow{w}_2$, $\overrightarrow{w}_4$ et $\overrightarrow{w}_5$ forment une base de $\mathrm{Ker}(B) = \mathrm{Ker}(A)$.

Pour construire une base de $\mathrm{Im}(A)$, il nous faut déterminer les vecteurs redondants parmis les vecteurs colonnes de A. Mais les vecteurs redondants de A correspondent à ceux de B dans le sens où si $\overrightarrow{a}_i$ est redondant alors $\overrightarrow{b}_i$ l'est aussi.

Vérifions ce point sur la dernière colonne, puisque nous savons que $\overrightarrow{b}_5$ est redondant, avec $\overrightarrow{b}_5 = -4\overrightarrow{b}_1 + 5\overrightarrow{b}_3$. Cela nous a fourni le vecteur

$$w_5 = \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix} \in \mathrm{Ker}(A) = \mathrm{Ker}(B),$$

d'où la relation $4\vec{a}_5 - 5\vec{a}_3 + \vec{a}_5 = \vec{0}$, *i.e.* $\vec{a}_5 = -4\vec{a}_1 + 5\vec{a}_3$, ce qui veut dire que $\vec{a}_5$ est redondant.

Comme on sait que $\vec{b}_2$, $\vec{b}_4$ et $\vec{b}_5$ sont redondants, il en résulte qu'une base de $\mathrm{Im}(A)$ est

$$\vec{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 4 \\ 3 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \\ 1 \end{bmatrix}.$$

□

En suivant l'exemple 3.3.1 on peut établir la règle générale suivante.

Algorithme 3 (Construction d'une base de l'image)**.** *Pour construire une base de l'image d'une matrice A, il suffit de prendre les vecteurs colonnes qui correspondent aux colonnes qui ont un pivot dans* $\mathrm{Frel}(A)$*.*

Il faut bien faire attention à prendre les vecteurs colonnes de A et pas ceux de $\mathrm{Frel}(A)$, car les matrices A et $\mathrm{Frel}(A)$ n'ont pas nécessairement les mêmes images. Par exemple si on prend

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathrm{Frel}(A) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Par contre A et $\mathrm{Frel}(A)$ ont toujours le même noyau. On note que pour trouver une base de l'image d'une matrice A, on prend en compte autant de colonnes de A qu'il y a de pivots dans la Frel correspondante, d'où un autre résultat essentiel :

Proposition 3.3.2 (Dimension de l'image)**.** *Soit A une matrice. Alors*

$$\mathrm{Dim}(\mathrm{Im}(A)) = \mathrm{Rang}(A).$$

On dit parfois que la dimension du noyau d'une matrice est la **nullité** de la matrice. Par conséquent la formule fondamentale de l'algèbre linéaire

$$\mathrm{Dim}(\mathrm{Im})(A) + \mathrm{Dim}(\mathrm{Ker})(A) = m$$

peut aussi sécrire

$$\text{Nullité de } A + \text{Rang de } A = m,$$

pour n'importe quelle matrice de taille $n \times m$.

Example 3.3.1. *Soit T la projection orthogonale de $\mathbb{R}^3$ sur un plan V.*

Ici $m = 3$, le noyau de T est la droite $V^\perp$ et l'image est le plan V. On a bien

$$1 + 2 = 3.$$

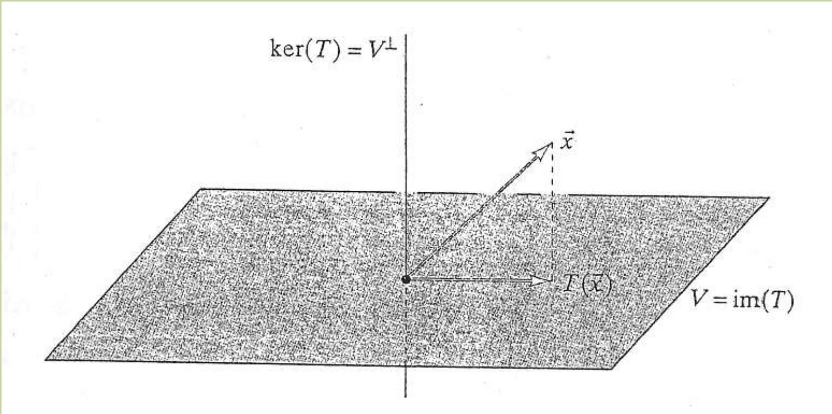


Figure 3.16: -

3.3.2 Bases de $\mathbb{R}^n$.

On sait qu'il faut n vecteurs indépendants pour constituer une base de $\mathbb{R}^n$, et on connait déjà la base canonique $\overrightarrow{e}_1, ..., \overrightarrow{e}_n$. La question est comment peut-on décider que n vecteurs $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ forment une base de $\mathbb{R}^n$?

On sait depuis la section 3.2 que la famille $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ est une base si et seulement si chaque vecteur $\overrightarrow{b} \in \mathbb{R}^n$ sécrit de manière unique sous forme d'une combinaison linéaire des vecteurs $\overrightarrow{v}_i$:

$$\overrightarrow{b} = c_1\overrightarrow{v}_1 + \cdots + c_n\overrightarrow{v}_n = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

On sait que le système

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \overrightarrow{b}$$

admet une unique solution si et seulement si la matrice

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix}$$

est inversible.

Résumé 3.3.1 (Bases de $\mathbb{R}^n$)**.** *La familllle* $(\overrightarrow{v}_1, ..., \overrightarrow{v}_n)$ *est une base de* $\mathbb{R}^n$ *si et seulement si la matrice*

$$\begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_n \end{bmatrix}$$

est inversible.

Exercice 3.3.3. *Pour quelle valeurs de* k *la famille suivante de vecteurs est-elle une base de* $\mathbb{R}^3$ *:*

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix}$$

Solution. On considère la matrice

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & k \\ 1 & 1 & k^2 \end{bmatrix},$$

qui se réduit à

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & (1-k)/2 \\ 0 & 0 & k^2-1 \end{bmatrix}.$$

On peut achever l'algorithme de Gauss pour réduire cette matrice en I_3 si et seulement si $k^2 - 1 \neq 0$, c'est-à-dire $k \neq 1$ et $k \neq -1$. Conclusion, ces vecteurs forment une base de $\mathbb{R}^3$ si et seulement si k est différent de 1 et de -1. □

On fait la synthèse de ces résultats dans le résumé suivant.

Résumé 3.3.2 (Différentes caractéristiques des matrices inversibles)**.** *Soit* $A \in M_n(\mathbb{R})$ *(i.e une matrice carrée de taille* $n \times n$*). Les assertions suivantes sont équivalentes :*

i) A *est inversible,*

ii) $\forall \overrightarrow{b} \in \mathbb{R}^n$, $A\overrightarrow{x} = \overrightarrow{b}$ *admet une et une seule solution,*

iii) $\exists \overrightarrow{b}_0 \in \mathbb{R}^n$, $A\overrightarrow{x} = \overrightarrow{b}_0$ *admet une et une seule solution,*

iv) $\mathrm{Frel}(A) = I_n$,

v) $\mathrm{Rang}(A) = n$,

vi) $\mathrm{Im}(A) = \mathbb{R}^n$,

vii) $\mathrm{Ker}(A) = \{\overrightarrow{0}\}$,

viii) *Les vecteurs colonnes de* A *forment une base de* $\mathbb{R}^n$,

ix) *Les vecteurs colonnes de* A *engendrent* $\mathbb{R}^n$,

x) *Les vecteurs colonnes de* A *sont linéairement indépendants.*

3.4 Coordonnées

Les coordonnées constituent "la grande idée" des mathématiques. On attribue à René Descartes (1596-1650) de les avoir introduites dans un appendice de son traité "*Discours de la méthode*"(Leyde, 1637). On dit que l'idée lui serait venue, alors qu'il était couché sur le dos dans son lit un dimanche matin à regarder une mouche voler au plafond. Il lui apparu qu'il pourrait décrire la position de la mouche en donnant sa distance par rapport à deux murs.

Pierre de Fermat (1601-1665) aurait développé les mêmes principes de géométrie analytique indépendamment au même moment, mais n'aurait pas publié ses travaux sur ce sujet.

On a utilisé des coordonnées cartésiennes dans le plan x, y et dans l'espace x, y, z dans les chapitres et sections précédents, sans trop insister sur la notion, pour représenter des vecteurs dans $\mathbb{R}^2$ et dans $\mathbb{R}^3$.

Dans cette section 3.4, on développe cette notion de façon systématique.

Exercice 3.4.1. *Soient les vecteurs de* $\mathbb{R}^3$

$$\overrightarrow{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad et \quad \overrightarrow{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

et considérons le plan $\mathrm{Vect}(\overrightarrow{v}_1, \overrightarrow{v}_2)$ *dans* $\mathbb{R}^3$. *Est-ce que le vecteur*

$$\overrightarrow{x} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$$

appartient à V *? Visualiser la réponse.*

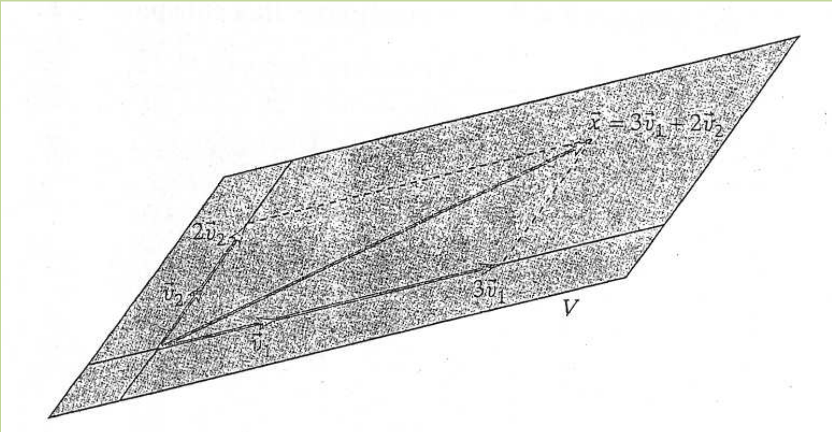


Figure 3.17: -

Solution. On se demande s'il existe deux scalaires c_1 et c_2 tels que $\overrightarrow{x} = c_1\overrightarrow{v}_1 + c_2\overrightarrow{v}_2$. Cela revient à considérer le système linéaire qui a pour matrice augmentée

$$M = \begin{bmatrix} 1 & 1 & \vdots & 5 \\ 1 & 2 & \vdots & 7 \\ 1 & 3 & \vdots & 9 \end{bmatrix} \quad \text{avec} \quad \mathrm{Frel}(M) = \begin{bmatrix} 1 & 0 & \vdots & 3 \\ 0 & 1 & \vdots & 2 \\ 0 & 0 & \vdots & 0 \end{bmatrix}.$$

Ce système est donc consistant et admet pour unique solution $c_1 = 3$ et $c_2 = 2$ de sorte que

$$\overrightarrow{x} = c_1\overrightarrow{v}_1 + c_2\overrightarrow{v}_2 = 3\overrightarrow{v}_1 + 2\overrightarrow{v}_2.$$

On a représenté géométriquement cette solution dans la figure 3.17, et il apparait clairement que $\overrightarrow{x}$ est bien dans le plan V.

Pour visualiser les coefficients 2 et 3 dans la combinaison linéaire $\overrightarrow{x} = 3\overrightarrow{v}_1 + 2\overrightarrow{v}_2$, il est commode d'introduire une **grille de coordonnées** dont les axes, notés c_1 et c_2, pointent dans les directions des vecteurs $\overrightarrow{v}_1$ et $\overrightarrow{v}_2$, comme sur la figure 2.18.

Le **vecteur coordonnée** de $\overrightarrow{v} = 3\overrightarrow{v}_1 + 2\overrightarrow{v}_2$ dans ce système de coordonnées est

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

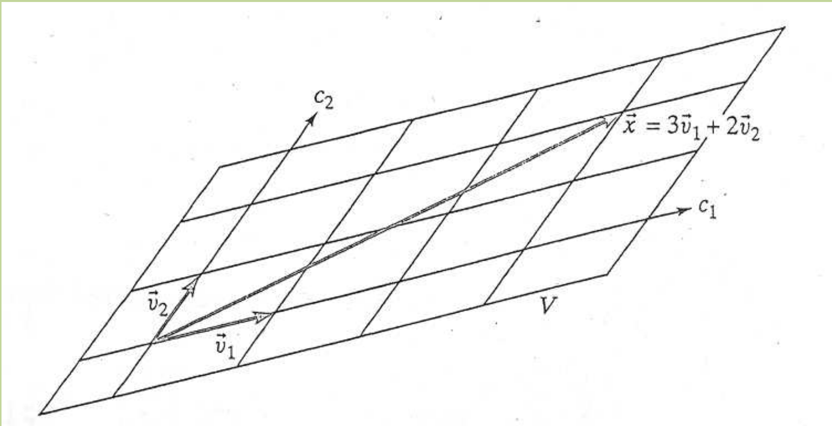


Figure 3.18: -

On peut imaginer que $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ est "l'adresse" de $\overrightarrow{x}$ dans le système de coordonnées c_1, c_2. En introduisant ce système de coordonnées, on a identifié V à $\mathbb{R}^2$. Il ne faut pas s'inquiéter du fait que les axes c_1 et c_2 ne soient pas perpendiculaires : les coordonnées cartésiennes ont un sens également dans le cas d'axes obliques.

On note $\mathcal{B}$ la base $\overrightarrow{v}_1, \overrightarrow{v}_2$ et $[\overrightarrow{x}]_{\mathcal{B}}$ le vecteur coordonnée de $\overrightarrow{x}$ dans la base $\mathcal{B}$.

Si

$$\overrightarrow{x} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} = c_1 \overrightarrow{v}_1 + c_2 \overrightarrow{v}_2 = 3\overrightarrow{v}_1 + 2\overrightarrow{v}_2,$$

alors

$$[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

□

On généralise l'idée introduite dans l'exercice 3.4.1 comme suit.

Définition 3.4.1 (Coordonnées dans un sous-espace de $\mathbb{R}^n$)**.** *Soit*

$$\mathcal{B} = \overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$$

une base d'un sous-espace vectoriel V de $\mathbb{R}^n$ et $\overrightarrow{x} \in V$. Le vecteur $\overrightarrow{x}$ peut sécrire de façon unique sous la forme

$$\overrightarrow{x} = c_1 \overrightarrow{v}_1 + \cdots + c_m \overrightarrow{v}_m.$$

les scalaires $c_1, \cdots, c_m$ sont appelées les $\mathcal{B}$-coordonnées de $\overrightarrow{x}$, et le vecteur

$$\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$$

est le $\mathcal{B}$-vecteur coordonnées de $\overrightarrow{x}$ noté $[\overrightarrow{x}]_{\mathcal{B}}$.

En particulier

$$[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$$

veut dire

$$\overrightarrow{x} = c_1 \overrightarrow{v}_1 + \cdots + c_m \overrightarrow{v}_m.$$

On notera que l'on a la relation matricielle

Proposition 3.4.1.

$$\overrightarrow{x} = P \cdot [\overrightarrow{x}]_{\mathcal{B}}, \quad \textit{avec} \quad P = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 & \cdots & \overrightarrow{v}_m \end{bmatrix},$$

P étant une matrice de taille $n \times m$.

L'équation $\overrightarrow{x} = P \cdot [\overrightarrow{x}]_{\mathcal{B}}$ résulte directement de la définition des coordonnées.

Dans l'exemple ??, nous avons considéré le cas

$$\overrightarrow{x} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}, \quad [\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix},$$

$$\overrightarrow{x} = P[\overrightarrow{x}]_{\mathcal{B}} \quad \text{ou encore} \quad \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Il en résulte une propriété importante de linéarité.

Proposition 3.4.2 (Linéarité des coordonnées)**.** *Soit $\mathcal{B}$ une base d'un sous-espace de $\mathbb{R}^n$. Alors on a :*

(a) $\forall \overrightarrow{x} \in V, \forall \overrightarrow{y} \in V, \qquad [\overrightarrow{x} + \overrightarrow{y}]_{\mathcal{B}} = [\overrightarrow{x}]_{\mathcal{B}} + [\overrightarrow{y}]_{\mathcal{B}},$

(b) $\forall \overrightarrow{x} \in V$, $\forall k \in \mathbb{R}$, $\quad [k\overrightarrow{x}]_{\mathcal{B}} = k[\overrightarrow{x}]_{\mathcal{B}}$.

Démonstration. On démontre le point (b) Le point (a) fait l'objet de l'exercice 51. Soit $\mathcal{B} = \overrightarrow{v}_1, \cdots, \overrightarrow{v}_m$ une base de V,

$$\overrightarrow{x} = c_1\overrightarrow{v}_1 + \cdots + c_m\overrightarrow{v}_m = \sum_{j=1}^{m} c_j\overrightarrow{v}_j \ \in V.$$

Alors

$$k\overrightarrow{x} = kc_1\overrightarrow{v}_1 + \cdots + kc_m\overrightarrow{v}_m = \sum_{j=1}^{m} kc_j\overrightarrow{v}_j,$$

de sorte que

$$[k\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} kc_1 \\ \vdots \\ kc_m \end{bmatrix} = k\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = k\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = k[\overrightarrow{x}]_{\mathcal{B}},$$

comme annoncé. □

Considérons la définition des coordonnées dans le cas particulier où $V = \mathbb{R}^n$. Il est souvent très utile de travailler avec d'autres bases que la base canonique $\overrightarrow{e}_1, \cdots, \overrightarrow{e}_n$. Par exemple, si on veut étudier l'ellipse de la figure 3.19, on voit que les axes $c_1 - c_2$ alignés avec les axes principaux de l'ellipse sont préférables aux axes standards x_1, x_2.

Exercice 3.4.2. *Considérons la base de* $\mathbb{R}^2$, $\mathcal{B} = \overrightarrow{v}_1, \overrightarrow{v}_2$, *où* $\overrightarrow{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ *et* $\overrightarrow{v}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

(a) Soit $\overrightarrow{x} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$. *Trouver* $[\overrightarrow{x}]_{\mathcal{B}}$.

(b) Soit $\overrightarrow{y}$ *tel que* $[\overrightarrow{y}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. *Déterminer* $\overrightarrow{y}$.

Proof. (a) Pour trouver les $\mathcal{B}$-coordonnées du vecteur $\overrightarrow{x}$, on écrit $\overrightarrow{x}$ comme combinaison linéaire des vecteurs de la base,

$$\overrightarrow{x} = c_1\overrightarrow{v}_1 + c_2\overrightarrow{v}_2 \quad \text{c'est-à-dire} \quad \begin{bmatrix} 10 \\ 10 \end{bmatrix} = c_1\begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

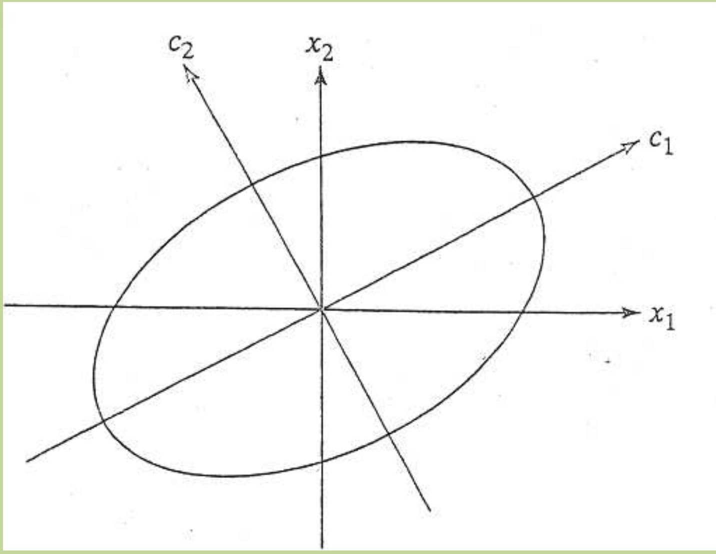


Figure 3.19: -

Ce système admet pour solution $c_1 = 4$, $c_2 = 2$ de sorte que $[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

Une autre méthode consiste à utiliser l'équation $\overrightarrow{x} = P[\overrightarrow{x}]_{\mathcal{B}}$ autrement dit $[\overrightarrow{x}]_{\mathcal{B}} = P^{-1}\overrightarrow{x}$, c'est-à-dire

$$[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}^{-1} \overrightarrow{x} = \frac{1}{10}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 10 \\ 10 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

(b) Par définition, $[\overrightarrow{y}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ se traduit par

$$\overrightarrow{y} = 2\overrightarrow{v}_1 + (-1)\overrightarrow{v}_2 = 2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}.$$

Autrement, on peut aussi directement utiliser la formule

$$\overrightarrow{y} = P[\overrightarrow{y}]_{\mathcal{B}} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}.$$

Ces résultats sont illustrés sur la figure 3.20 : □

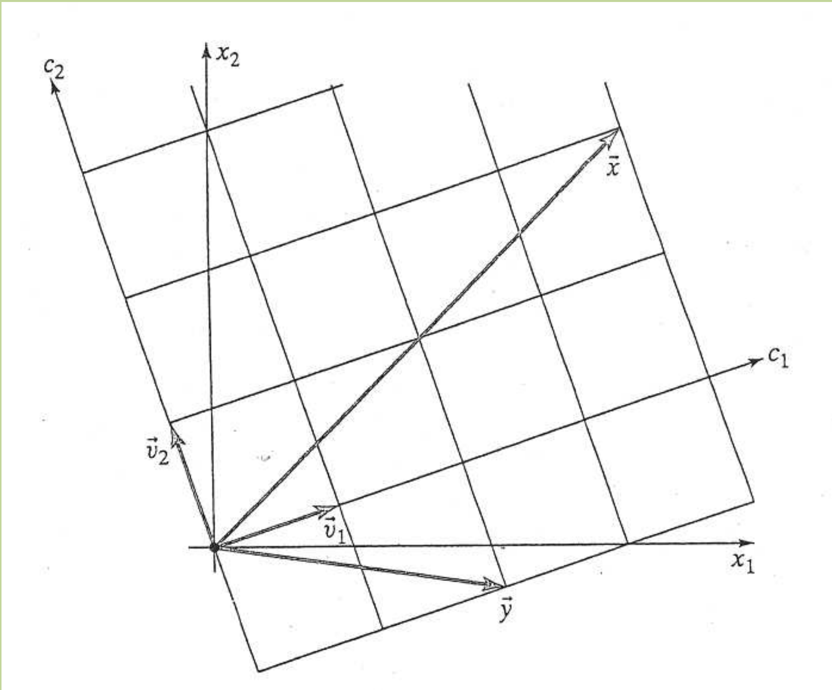


Figure 3.20: -

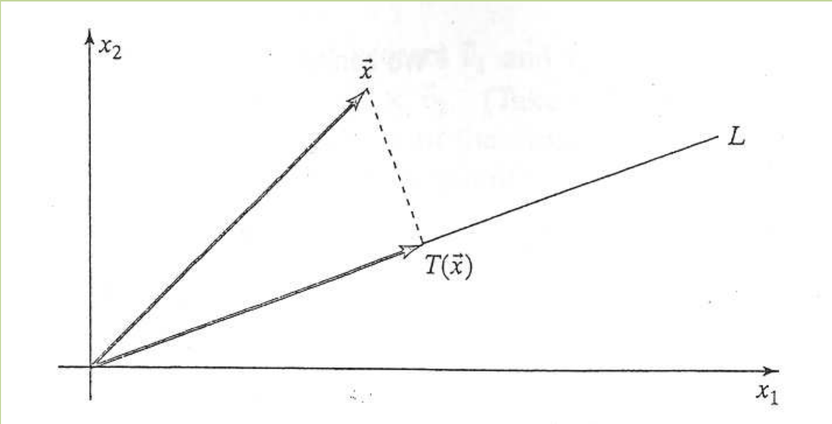


Figure 3.21: -

Example 3.4.1. *Soit $L \subset \mathbb{R}^2$ la droite engendrée par le vecteur $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Soit T l'application linéaire de $\mathbb{R}^2$ dans lui-même qui projette chaque vecteur $\overrightarrow{x}$ orthogonalement sur la droite L.*

On peut faciliter l'étude de T en introduisant un système de coordonnées dans lequel L serait un des axes (par exemple l'axe c_1) avec pour axe c_2 l'axe othogonal à L.

Suivant ce système de coordonnées, T transforme $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ en $\begin{bmatrix} c_1 \\ 0 \end{bmatrix}$.

Dans le système de coordonnées c_1, c_2, T est représenté par la matrice

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

car

$$\begin{bmatrix} c_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Précisions cela. Soit $\mathcal{B} = (\overrightarrow{v}_1, \overrightarrow{v}_2)$ une base de $\mathbb{R}^2$ telle que $\overrightarrow{v}_1$ est parallèle à la droite L et $\overrightarrow{v}_1$ est parallèle à la droite $L^\perp$. Par exemple $\overrightarrow{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ et $\overrightarrow{v}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

Si $\overrightarrow{x} = c_1 \overrightarrow{v}_1 + c_2 \overrightarrow{v}_2$, alors $T(\overrightarrow{x}) = Proj_L(\overrightarrow{x}) = c_1 \overrightarrow{v}_1$, ou encore

$$[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad \textit{alors} \quad [T(\overrightarrow{x})]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ 0 \end{bmatrix},$$

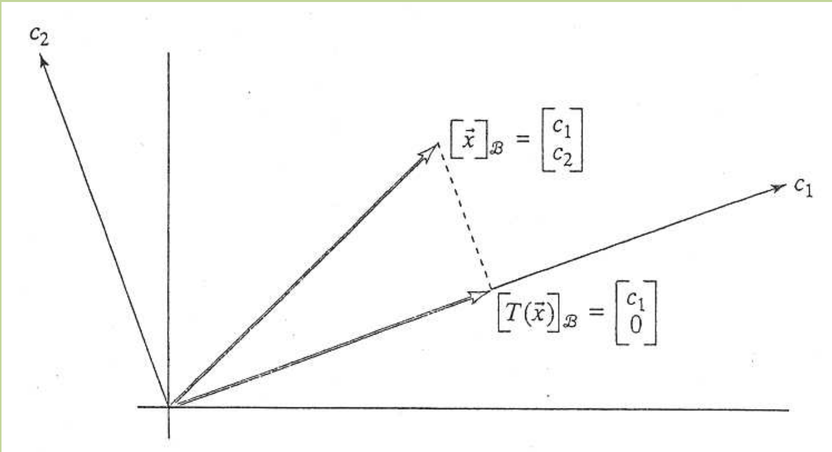


Figure 3.22: -

comme indiqué sur la figure 3.22.

La matrice $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *qui transforme* $[\overrightarrow{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ *en* $[T(\overrightarrow{x})]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ 0 \end{bmatrix}$ *est "***la matrice de** T **relativement à la base** $\mathcal{B}$*" (ou aussi* $\mathcal{B}$*-matrice de* T*) dans le sens où*

$$[T(\overrightarrow{x})]_{\mathcal{B}} = B[\overrightarrow{x}]_{\mathcal{B}}.$$

On peut représenter le travail sous forme d'un diagramme comme suit :

$$\begin{array}{ccc} \overrightarrow{x} & \xrightarrow{\quad T \quad} & T(\overrightarrow{x}) \\ \uparrow{\scriptstyle P} & & \uparrow{\scriptstyle P} \\ [\overrightarrow{x}]_{\mathcal{B}} & \xrightarrow{\quad B \quad} & [T(\overrightarrow{x})]_{\mathcal{B}} \end{array}$$

Définition 3.4.2 (Matrice d'une application linéaire)**.** *Soit* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *une application lin'eaire et* $\mathcal{B}$ *une base de* $\mathbb{R}^n$*. La matrice carrée* B *d'ordre* n *qui transforme* $[\overrightarrow{x}]_{\mathcal{B}}$ *en* $[T(\overrightarrow{x})]_{\mathcal{B}}$ *est la matrice de* T *relativement à la base* $\mathcal{B}$*,*

$$\forall \overrightarrow{x} \in \mathbb{R}^n, \quad [T(\overrightarrow{x})]_{\mathcal{B}} = B[\overrightarrow{x}]_{\mathcal{B}}.$$

La matrice B *est construite en colonnes de la manière suivante, en notant* $\mathcal{B} = (\overrightarrow{v}_1, \cdots, \overrightarrow{v}_n)$*,*

$$B = \begin{bmatrix} [T(\overrightarrow{v}_1)]_{\mathcal{B}} & \cdots & [T(\overrightarrow{v}_n)]_{\mathcal{B}} \end{bmatrix}.$$

Il faut vérifier que les colonnes de B sont bien les vecteurs $[T(\overrightarrow{v}_1)]_{\mathcal{B}} \cdots [T(\overrightarrow{v}_n)]_{\mathcal{B}}$. Soit $\overrightarrow{x} = c_1\overrightarrow{v}_1 + \cdots + c_n\overrightarrow{v}_n$. Comme T est linéaire, on a

$$T(\overrightarrow{x}) = c_1T(\overrightarrow{v}_1) + \cdots c_nT(\overrightarrow{v}_n)$$

et par conséquent,

$$\begin{aligned}[T(\overrightarrow{x})]_{\mathcal{B}} &= c_1[T(\overrightarrow{v}_1)]_{\mathcal{B}} + \cdots c_n[T(\overrightarrow{v}_n)]_{\mathcal{B}} \\ & \begin{bmatrix} \\ [T(\overrightarrow{v}_1)]_{\mathcal{B}} & \cdots & [T(\overrightarrow{v}_n)]_{\mathcal{B}} \\ \\ \end{bmatrix} [\overrightarrow{x}]_{\mathcal{B}} = B[\overrightarrow{x}]_{\mathcal{B}}.\end{aligned}$$

On peut utiliser cette méthode pour construire B, bien qu'il soit souvent plus simple d'utiliser un diagramme comme on l'a fait dans l'exemple 3.4.1.

On revient au cas de l'exemple 3.4.1. Soit $L \subset \mathbb{R}^2$ la droite engendrée par le vecteur $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Soit T l'application linéaire de $\mathbb{R}^2$ dans lui-même qui projète chaque vecteur $\overrightarrow{x}$ orthogonalement sur la droite L. On avait vu que la matrice de T relativement à la base

$$\mathcal{B} = \left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix} \right)$$

était la matrice

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Soit A la matrice "standard" de T telle qu'on l'a définie dans les sections précédentes, à savoir la matrice qui vérifie

$$T(\overrightarrow{x}) = A\overrightarrow{x}.$$

On remarque qu'avec avec cette nouvelle définition, la matrice A est la matrice de T relativement à la base canonique $\mathcal{U} = (\overrightarrow{e}_1, \overrightarrow{e}_2)$.

Exercice 3.4.3. *Quelle est la relation entre les matrices A et B ?*

Solution. On rappelle que l'on a par définition

$$\overrightarrow{x} = P[\overrightarrow{x}]_{\mathcal{B}}, \quad \text{où} \quad P = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}.$$

Considérons le diagramme suivant

$$\begin{array}{ccc} \overrightarrow{x} & \xrightarrow{\quad A \quad} & T(\overrightarrow{x}) \\ \uparrow \text{P} & & \uparrow \text{P} \\ [\overrightarrow{x}]_{\mathcal{B}} & \xrightarrow{\quad B \quad} & [T(\overrightarrow{x})]_{\mathcal{B}} \end{array}$$

On note que puisque

$$T(\vec{x}) = A\vec{x}, [T(\vec{x})]_{\mathcal{B}} = B[\vec{x}]_{\mathcal{B}}$$

$$T(\vec{x}) = A\vec{x} = AP[\vec{x}]_{\mathcal{B}},$$

$$T(\vec{x}) = P[T(\vec{x})]_{\mathcal{B}} = PB[\vec{x}]_{\mathcal{B}}$$

pour tout $\vec{x} \in \mathbb{R}^n$. Par conséquent

$$AP = PB, \quad B = P^{-1}AP \quad \text{et} \quad A = PBP^{-1}.$$

On dit que

P est la matrice de passage de la base $\mathcal{U}$ à la base $\mathcal{B}$,

(P comme "passage") et la formule est connue sous la forme

$$\boxed{PB = AP, \quad B = P^{-1}AP}$$

On peut utiliser cette formule pour trouver la matrice A,

$$A = PBP^{-1} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{1}{10} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix} \right) = \begin{bmatrix} 0,9 & 0,3 \\ 0,3 & 0,1 \end{bmatrix}.$$

□

Définition 3.4.3 (Matrice de passage)**.** *Soit $T : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire et $\mathcal{B} = (\vec{v}_1, \cdots, \vec{v}_n)$ une base de $\mathbb{R}^n$. Soit B la matrice de T relativement à la base $\mathcal{B}$, et A la matrice de T relativement à la base canonique $\mathcal{U} = (\vec{e}_1, \cdots, \vec{e}_n)$. Enfin soit*

$$P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$$

la **matrice de passage** *de la base $\mathcal{U}$ à la base $\mathcal{B}$. Alors*

$$AP = PB, \quad B = P^{-1}AP, \quad A = PBP^{-1}.$$

Ce qui précède motive la definition suivante :

Définition 3.4.4. *Soient A et B deux matrices carrées d'ordre n. On dit qu'elles sont* **semblables** *si il existe une matrice P* **inversible** *telle que*

$$AP = PB \quad \textit{ou bien de manière équivalente} \quad B = P^{-1}AP.$$

En clair, deux matrices sont semblables si elles représentent la même application linéaire mais dans des bases différentes.

Exercice 3.4.4. *Est-ce que les matrices* $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ *et* $B = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$ *sont semblables ?*

Solution. On cherche s'il existe une matrice inversible $P = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$ telle que l'on ait $AP = PB$. Cette dernière relation s'écrit composante par composante sous la forme

$$\begin{bmatrix} x + 2z & y + 2t \\ 4x + 3z & 4y + 3t \end{bmatrix} = \begin{bmatrix} 5x & -y \\ 5z & -t \end{bmatrix},$$

ce qui fournit un système qui se réduit à

$$z = 2x, \quad t = -y,$$

de sorte que n'importe quelle matrice de la forme

$$P = \begin{bmatrix} x & y \\ 2x & -y \end{bmatrix}$$

vérfie $AP = PB$. Cependant, pour répondre à la question posée, il faut vérifier que parmi ces matrices, certaines sont inversibles. Or on a $\det(P) = -3xy$. Donc P est inversible si et seulement si $xy \neq 0$, et par exemple en prenant $x = y = 1$, on voit que

$$P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

est inversible et vérifie $AP = PB$. Par conséquent, les matrices A et B sont semblables. □

Exercice 3.4.5. *Soient A et B deux matrices carrées d'ordre n semblables. Soit $k \in \mathbb{N}$. Alors les matrices A^k et B^k sont semblables.*

Solution. On rappelle que

$$A^k = \underbrace{A \times A \cdots \times A}_{k \text{ fois}}, \quad B^k = \underbrace{B \times B \cdots \times B}_{k \text{ fois}}.$$

Comme A et B sont semblables, il existe $P \in M_n(\mathbb{R})$ inversible telle que $B = P^{-1}AP$. Par conséquent

$$B^k = \underbrace{(P^{-1}AP) \times (P^{-1}AP) \cdots \times (P^{-1}AP)}_{k \text{ fois}}.$$

Or, en utilisant l'associativité de la multiplication des matrices, on a la relation

$$(P^{-1}AP) \times (P^{-1}AP) = (P^{-1}A) \times (PP^{-1}) \times (AP).$$

Puisque $PP^{-1} = I_n$, on en déduit que

$$(P^{-1}AP) \times (P^{-1}AP) = P^{-1}A^2P.$$

Donc

$$B^k = \underbrace{(P^{-1}AP) \times (P^{-1}AP) \cdots \times (P^{-1}AP)}_{k \text{ fois}}.$$

conduit de proche en proche à la relation

$$B^k = P^{-1}A^kP,$$

ce qui prouve bien que les matrices A^k et B^k sont semblables. □

Exercice 3.4.6. *Soient A une matrice de rotation et B une matrice de projection. Sont elles semblables ?*

Solution. Non, elles ne sont pas semblables. En effet la matrice A est inversible. Si A et B étaient semblables, il existerait P inversible telle que $B = P^{-1}AP$, et B serait également inversible avec $B^{-1} = P^{-1}A^{-1}P$. Or une matrice de projection n'est pas inversible. □

La notation $M_n(\mathbb{R})$ désigne l'ensembles des matrices carrées d'ordre n.

Proposition 3.4.3. *La relation "A et B sont semblables" est une relation d'équivalence sur $M_n(\mathbb{R})$.*

*(a) Soit $A \in M_n(\mathbb{R})$. Alors la matrice A est semblable à elle-même (**reflexivité**).*

*(b) Soient $A \in M_n(\mathbb{R})$ et $B \in M_n(\mathbb{R})$. Si B est semblable à A alors A est semblable à B (**symétrie**).*

*(c) Soient $A \in M_n(\mathbb{R})$, $B \in M_n(\mathbb{R})$ et $C \in M_n(\mathbb{R})$. Si A est semblable à B et B est semblable à C, alors A est semblable à C (**transitivité**).*

Démonstration. On démontre la transitivité, laissant la symétrie et la reflexivité en exercice. On sait qu'il existe $P \in M_n(\mathbb{R})$ et $Q \in M_n(\mathbb{R})$ inversibles telles que

$$AP = PB \quad \text{et} \quad BQ = QC.$$

En multipliant à droite par Q la première égalité, on obtient $APQ = PBQ$. Puis comme $BQ = QC$, on en déduit que $A(PQ) = (PQ)C$. Comme P et Q sont inversibles, on en déduit que le produit PQ est inversible, ce qui fait que les matrices A et C sont semblables. $\square$