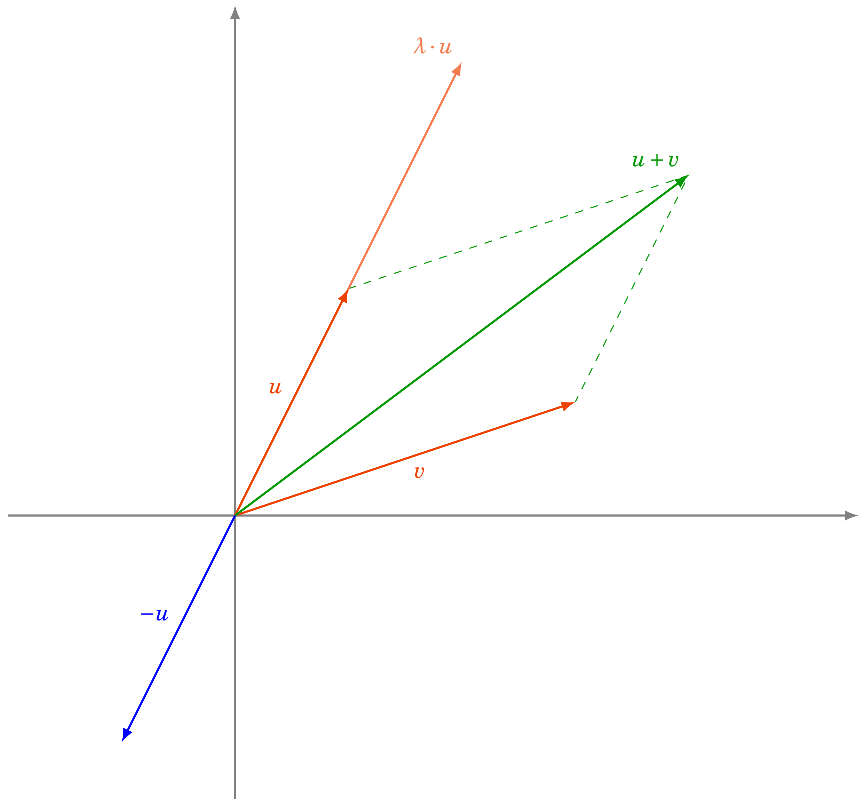
Cours de mathématiques
M22 Algèbre linéaire



Université Lille1
Sciences et Technologies

Exo7

Sommaire

Exo7

1 Systèmes linéaires

Vidéo ■ partie 1. Introduction aux systèmes d'équations linéaires
Vidéo ■ partie 2. Théorie des systèmes linéaires
Vidéo ■ partie 3. Résolution par la méthode du pivot de Gauss

1. Introduction aux systèmes d'équations linéaires

L'algèbre linéaire est un outil essentiel pour toutes les branches des mathématiques appliquées, en particulier lorsqu'il s'agit de modéliser puis résoudre numériquement des problèmes issus de divers domaines : des sciences physiques ou mécaniques, des sciences du vivant, de la chimie, de l'économie, des sciences de l'ingénieur,...
Les systèmes linéaires interviennent dans de nombreux contextes d'applications car ils forment la base calculatoire de l'algèbre linéaire. Ils permettent également de traiter une bonne partie de la théorie de l'algèbre linéaire en dimension finie. C'est pourquoi le présent cours commence avec une étude des équations linéaires et de leur résolution.
Ce chapitre a un but essentiellement pratique : résoudre des systèmes linéaires. La partie théorique sera revue et prouvée dans le chapitre « Matrices ».

1.1. Exemple : deux droites dans le plan

L'équation d'une droite dans le plan (Oxy) s'écrit

$$ax + by = e$$

où a, b et e sont des paramètres réels. Cette équation s'appelle ***équation linéaire*** linéaire (équation) dans les variables (ou inconnues) x et y.

Par exemple, $2x + 3y = 6$ est une équation linéaire, alors que les équations suivantes ne sont pas des équations linéaires :

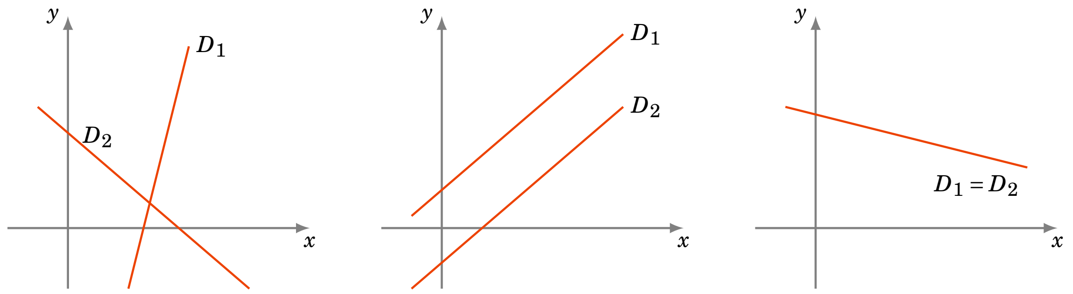
$$2x + y^2 = 1 \quad \text{ou} \quad y = \sin(x) \quad \text{ou} \quad x = \sqrt{y}.$$

Considérons maintenant deux droites D_1 et D_2 et cherchons les points qui sont simultanément sur ces deux droites. Un point (x, y) est dans l'intersection $D_1 \cap D_2$ s'il est solution du système :

$$\left\{\begin{array}{rcl} ax + by & = & e \\ cx + dy & = & f \end{array}\right. \qquad (S)$$

Trois cas se présentent alors :

1. Les droites D_1 et D_2 se coupent en un seul point. Dans ce cas, illustré par la figure de gauche, le système (S) a une seule solution.
2. Les droites D_1 et D_2 sont parallèles. Alors le système (S) n'a pas de solution. La figure du centre illustre cette situation.
3. Les droites D_1 et D_2 sont confondues et, dans ce cas, le système (S) a une infinité de solutions.



Nous verrons plus loin que ces trois cas de figure (une seule solution, aucune solution, une infinité de solutions) sont les seuls cas qui peuvent se présenter pour n'importe quel système d'équations linéaires.

1.2. Résolution par substitution

Pour savoir s'il existe une ou plusieurs solutions à un système linéaire, et les calculer, une première méthode est la ***substitution***. Par exemple pour le système :

$$\begin{cases} 3x+2y &= 1 \\ 2x-7y &= -2 \end{cases} \qquad (S)$$

Nous réécrivons la première ligne $3x+2y=1$ sous la forme $y=\frac{1}{2}-\frac{3}{2}x$. Et nous remplaçons (nous *substituons*) le y de la seconde équation, par l'expression $\frac{1}{2}-\frac{3}{2}x$. Nous obtenons un système équivalent :

$$\begin{cases} y &= \frac{1}{2}-\frac{3}{2}x \\ 2x-7(\frac{1}{2}-\frac{3}{2}x) &= -2 \end{cases}$$

La seconde équation est maintenant une expression qui ne contient que des x, et on peut la résoudre :

$$\begin{cases} y &= \frac{1}{2}-\frac{3}{2}x \\ (2+7\times\frac{3}{2})x &= -2+\frac{7}{2} \end{cases} \iff \begin{cases} y &= \frac{1}{2}-\frac{3}{2}x \\ x &= \frac{3}{25} \end{cases}$$

Il ne reste plus qu'à remplacer dans la première ligne la valeur de x obtenue :

$$\begin{cases} y &= \frac{8}{25} \\ x &= \frac{3}{25} \end{cases}$$

Le système (S) admet donc une solution unique $(\frac{3}{25},\frac{8}{25})$. L'ensemble des solutions est donc

$$\mathscr{S}=\left\{\left(\frac{3}{25},\frac{8}{25}\right)\right\}.$$

1.3. Exemple : deux plans dans l'espace

Dans l'espace $(0xyz)$, une équation linéaire est l'équation d'un plan :

$$ax+by+cz=d$$

L'intersection de deux plans dans l'espace correspond au système suivant à 2 équations et à 3 inconnues :

$$\begin{cases} ax+by+cz &= d \\ a'x+b'y+c'z &= d' \end{cases}$$

Trois cas se présentent alors :

- les plans sont parallèles (et distincts) et il n'y a alors aucune solution au système,
- les plans sont confondus et il y a une infinité de solutions au système,
- les plans se coupent en une droite et il y a une infinité de solutions.

Exemple 1

1. Le système $\begin{cases} 2x+3y-4z &= 7 \\ 4x+6y-8z &= -1 \end{cases}$ n'a pas de solution. En effet, en divisant par 2 la seconde équation, on obtient le système équivalent : $\begin{cases} 2x+3y-4z &= 7 \\ 2x+3y-4z &= -\frac{1}{2} \end{cases}$. Les deux lignes sont clairement incompatibles : aucun (x,y,z) ne peut vérifier à la fois $2x+3y-4z=7$ et $2x+3y-4z=-\frac{1}{2}$. L'ensemble des solutions est donc $\mathscr{S}=\varnothing$.

2. Pour le système $\begin{cases} 2x+3y-4z &= 7 \\ 4x+6y-8z &= 14 \end{cases}$, les deux équations définissent le même plan ! Le système est donc équivalent à une seule équation : $2x+3y-4z=7$. Si on récrit cette équation sous la forme $z=\frac{1}{2}x+\frac{3}{4}y-\frac{7}{4}$, alors on peut décrire l'ensemble des solutions sous la forme : $\mathscr{S}=\left\{(x,y,\frac{1}{2}x+\frac{3}{4}y-\frac{7}{4}) \mid x,y\in\mathbb{R}\right\}$.

3. Soit le système $\begin{cases} 7x+2y-2z &= 1 \\ 2x+3y+2z &= 1 \end{cases}$. Par substitution :

$$\begin{cases} 7x+2y-2z=1 \\ 2x+3y+2z=1 \end{cases} \iff \begin{cases} z=\frac{7}{2}x+y-\frac{1}{2} \\ 2x+3y+2\left(\frac{7}{2}x+y-\frac{1}{2}\right)=1 \end{cases} \iff \begin{cases} z=\frac{7}{2}x+y-\frac{1}{2} \\ 9x+5y=2 \end{cases}$$

$$\iff \begin{cases} z=\frac{7}{2}x+y-\frac{1}{2} \\ y=-\frac{9}{5}x+\frac{2}{5} \end{cases} \iff \begin{cases} z=\frac{17}{10}x-\frac{1}{10} \\ y=-\frac{9}{5}x+\frac{2}{5} \end{cases}$$

Pour décrire l'ensemble des solutions, on peut choisir x comme paramètre :

$$\mathscr{S}=\left\{\left(x,-\frac{9}{5}x+\frac{2}{5},\frac{17}{10}x-\frac{1}{10}\right) \mid x\in\mathbb{R}\right\}.$$

Géométriquement : nous avons trouvé une équation paramétrique de la droite définie par l'intersection de deux plans.

Du point de vue du nombre de solutions, nous constatons qu'il n'y a que deux possibilités, à savoir aucune solution ou une infinité de solutions. Mais les deux derniers cas ci-dessus sont néanmoins très différents géométriquement et il semblerait que dans le second cas (plans confondus), l'infinité de solutions soit plus grande que dans le troisième cas. Les chapitres suivants nous permettront de rendre rigoureuse cette impression.

Si on considère trois plans dans l'espace, une autre possibilité apparaît : il se peut que les trois plans s'intersectent en un seul point.

1.4. Résolution par la méthode de Cramer

On note $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ le ***déterminant***. On considère le cas d'un système de 2 équations à 2 inconnues :

$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases}$$

Si $ad - bc \neq 0$, on trouve une unique solution dont les coordonnées (x, y) sont :

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \qquad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

Notez que le dénominateur égale le déterminant pour les deux coordonnées et est donc non nul. Pour le numérateur de la première coordonnée x, on remplace la première colonne par le second membre ; pour la seconde coordonnée y, on remplace la seconde colonne par le second membre.

Exemple 2

Résolvons le système $\begin{cases} tx - 2y &= 1 \\ 3x + ty &= 1 \end{cases}$ suivant la valeur du paramètre $t \in \mathbb{R}$.

Le déterminant associé au système est $\begin{vmatrix} t & -2 \\ 3 & t \end{vmatrix} = t^2 + 6$ et ne s'annule jamais. Il existe donc une unique solution (x, y) et elle vérifie :

$$x = \frac{\begin{vmatrix} 1 & -2 \\ 1 & t \end{vmatrix}}{t^2 + 6} = \frac{t + 2}{t^2 + 6}, \qquad y = \frac{\begin{vmatrix} t & 1 \\ 3 & 1 \end{vmatrix}}{t^2 + 6} = \frac{t - 3}{t^2 + 6}.$$

Pour chaque t, l'ensemble des solutions est $\mathscr{S} = \left\{ \left(\frac{t+2}{t^2+6}, \frac{t-3}{t^2+6} \right) \right\}$.

1.5. Résolution par inversion de matrice

Pour ceux qui connaissent les matrices, le système linéaire

$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases}$$

est équivalent à

$$AX = Y \quad \text{où} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad Y = \begin{pmatrix} e \\ f \end{pmatrix}.$$

Si le déterminant de la matrice A est non nul, c'est-à-dire si $ad - bc \neq 0$, alors la matrice A est inversible et

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

et l'unique solution $X = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ du système est donnée par

$$X = A^{-1} Y.$$

Exemple 3

Résolvons le système $\begin{cases} x+y &= 1 \\ x+t^2y &= t \end{cases}$ suivant la valeur du paramètre $t \in \mathbb{R}$.

Le déterminant du système est $\begin{vmatrix} 1 & 1 \\ 1 & t^2 \end{vmatrix} = t^2-1$.

Premier cas. $t \neq +1$ **et** $t \neq -1$. Alors $t^2-1 \neq 0$. La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & t^2 \end{pmatrix}$ est inversible d'inverse $A^{-1} = \frac{1}{t^2-1}\begin{pmatrix} t^2 & -1 \\ -1 & 1 \end{pmatrix}$. Et la solution $X = \begin{pmatrix} x \\ y \end{pmatrix}$ est

$$X = A^{-1}Y = \frac{1}{t^2-1}\begin{pmatrix} t^2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ t \end{pmatrix} = \frac{1}{t^2-1}\begin{pmatrix} t^2-t \\ t-1 \end{pmatrix} = \begin{pmatrix} \frac{t}{t+1} \\ \frac{1}{t+1} \end{pmatrix}.$$

Pour chaque $t \neq \pm 1$, l'ensemble des solutions est $\mathscr{S} = \left\{\left(\frac{t}{t+1}, \frac{1}{t+1}\right)\right\}$.

Deuxième cas. $t = +1$. Le système s'écrit alors : $\begin{cases} x+y &= 1 \\ x+y &= 1 \end{cases}$ et les deux équations sont identiques. Il y a une infinité de solutions : $\mathscr{S} = \{(x, 1-x) \mid x \in \mathbb{R}\}$.

Troisième cas. $t = -1$. Le système s'écrit alors : $\begin{cases} x+y &= 1 \\ x+y &= -1 \end{cases}$, les deux équations sont clairement incompatibles et donc $\mathscr{S} = \varnothing$.

Mini-exercices

1. Tracer les droites et résoudre le système linéaire $\begin{cases} x-2y &= -1 \\ -x+3y &= 3 \end{cases}$ de trois façons différentes : substitution, méthode de Cramer, inverse d'une matrice. Idem avec $\begin{cases} 2x-y &= 4 \\ 3x+3y &= -5 \end{cases}$.
2. Résoudre suivant la valeur du paramètre $t \in \mathbb{R}$: $\begin{cases} 4x-3y &= t \\ 2x-y &= t^2 \end{cases}$.
3. Discuter et résoudre suivant la valeur du paramètre $t \in \mathbb{R}$: $\begin{cases} tx-y &= 1 \\ x+(t-2)y &= -1 \end{cases}$. Idem avec $\begin{cases} (t-1)x+y &= 1 \\ 2x+ty &= -1 \end{cases}$.

2. Théorie des systèmes linéaires

2.1. Définitions

Définition 1

On appelle ***équation linéaire*** dans les variables (ou ***inconnues***) $x_1, \ldots, x_p$ toute relation de la forme

$$a_1x_1 + \cdots + a_px_p = b, \tag{1.1}$$

où $a_1, \ldots, a_p$ et b sont des nombres réels donnés.

Remarque

- Il importe d'insister ici sur le fait que ces équations linéaires sont *implicites*, c'est-à-dire qu'elles décrivent des relations entre les variables, mais ne donnent pas directement les valeurs que peuvent prendre les variables.
- *Résoudre* une équation signifie donc la rendre *explicite*, c'est-à-dire rendre plus apparentes les valeurs que les variables peuvent prendre.
- On peut aussi considérer des équations linéaires de nombres rationnels ou de nombres complexes.

Soit $n \geqslant 1$ un entier.

Définition 2

Un ***système de n équations linéaires à p inconnues*** est une liste de n équations linéaires.

On écrit usuellement de tels systèmes en n lignes placées les unes sous les autres.

Exemple 4

Le système suivant a 2 équations et 3 inconnues :

$$\begin{cases} x_1 & - & 3x_2 & + & x_3 & = & 1 \\ -2x_1 & + & 4x_2 & - & 3x_3 & = & 9 \end{cases}$$

La forme générale d'un système linéaire de n équations à p inconnues est la suivante :

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +a_{13}x_3 & + & \cdots & +a_{1p}x_p & = & b_1 & (\leftarrow \text{équation } 1) \\ a_{21}x_1 & +a_{22}x_2 & +a_{23}x_3 & + & \cdots & +a_{2p}x_p & = & b_2 & (\leftarrow \text{équation } 2) \\ \vdots & \vdots & \vdots & & & \vdots & = & \vdots & \\ a_{i1}x_1 & +a_{i2}x_2 & +a_{i3}x_3 & + & \cdots & +a_{ip}x_p & = & b_i & (\leftarrow \text{équation } i) \\ \vdots & \vdots & \vdots & & & \vdots & = & \vdots & \\ a_{n1}x_1 & +a_{n2}x_2 & +a_{n3}x_3 & + & \cdots & +a_{np}x_p & = & b_n & (\leftarrow \text{équation } n) \end{cases}$$

Les nombres a_{ij}, $i = 1,\dots,n$, $j = 1,\dots,p$, sont les ***coefficients*** du système. Ce sont des données. Les nombres b_i, $i = 1,\dots,n$, constituent le ***second membre*** du système et sont également des données. Il convient de bien observer comment on a rangé le système en lignes (une ligne par équation) numérotées de 1 à n par l'indice i, et en colonnes : les termes correspondant à une même inconnue x_j sont alignés verticalement les uns sous les autres. L'indice j varie de 1 à p. Il y a donc p colonnes à gauche des signes d'égalité, plus une colonne supplémentaire à droite pour le second membre. La notation avec double indice a_{ij} correspond à ce rangement : le *premier* indice (ici i) est le numéro de *ligne* et le *second* indice (ici j) est le numéro de *colonne*. Il est extrêmement important de toujours respecter cette convention.

Dans l'exemple 4, on a $n = 2$ (nombre d'équations = nombre de lignes), $p = 3$ (nombre d'inconnues = nombre de colonnes à gauche du signe =) et $a_{11} = 1$, $a_{12} = -3$, $a_{13} = 1$, $a_{21} = -2$, $a_{22} = 4$, $a_{23} = -3$, $b_1 = 1$ et $b_2 = 9$.

Définition 3

Une ***solution*** du système linéaire est une liste de p nombres réels $(s_1, s_2, \dots, s_p)$ (un p-uplet) tels que si l'on substitue s_1 pour x_1, s_2 pour x_2, etc., dans le système linéaire, on obtient une égalité. L'***ensemble des solutions du système*** est l'ensemble de tous ces p-uplets.

Exemple 5

Le système

$$\begin{cases} x_1 & - & 3x_2 & + & x_3 & = & 1 \\ -2x_1 & + & 4x_2 & - & 3x_3 & = & 9 \end{cases}$$

admet comme solution $(-18, -6, 1)$, c'est-à-dire

$$x_1 = -18, \qquad x_2 = -6, \qquad x_3 = 1.$$

Par contre, $(7, 2, 0)$ ne satisfait que la première équation. Ce n'est donc pas une solution du système.

En règle générale, on s'attache à déterminer l'ensemble des solutions d'un système linéaire. C'est ce que l'on appelle ***résoudre*** le système linéaire. Ceci amène à poser la définition suivante.

Définition 4

On dit que deux systèmes linéaires sont ***équivalents*** s'ils ont le même ensemble de solutions.

À partir de là, le jeu pour résoudre un système linéaire donné consistera à le transformer en un système équivalent dont la résolution sera plus simple que celle du système de départ. Nous verrons plus loin comment procéder de façon systématique pour arriver à ce but.

2.2. Différents types de systèmes

Voici un résultat théorique important pour les systèmes linéaires.

Théorème 1

Un système d'équations linéaires n'a soit aucune solution, soit une seule solution, soit une infinité de solutions.

En particulier, si vous trouvez 2 solutions différentes à un système linéaire, alors c'est que vous pouvez en trouver une infinité ! Un système linéaire qui n'a aucune solution est dit ***incompatible***. La preuve de ce théorème sera vue dans un chapitre ultérieur (« Matrices »).

2.3. Systèmes homogènes

Un cas particulier important est celui des ***systèmes homogènes***, pour lesquels $b_1 = b_2 = \cdots = b_n = 0$, c'est-à-dire dont le second membre est nul. De tels systèmes sont toujours compatibles car ils admettent toujours la solution $s_1 = s_2 = \cdots = s_p = 0$. Cette solution est appelée *solution triviale*. Géométriquement, dans le cas 2×2, un système homogène correspond à deux droites qui passent par l'origine, $(0, 0)$ étant donc toujours solution.

Mini-exercices

1. Écrire un système linéaire de 4 équations et 3 inconnues qui n'a aucune solution. Idem avec une infinité de solution. Idem avec une solution unique.
2. Résoudre le système à n équations et n inconnues dont les équations sont $(L_i)\ : x_i - x_{i+1} = 1$ pour $i = 1,\dots,n-1$ et $(L_n)\ : x_n = 1$.
3. Résoudre les systèmes suivants :

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ \phantom{x_1 +{}} x_2 + 2x_3 + 3x_4 = 9 \\ \phantom{x_1 + x_2 +{}} x_3 + 2x_4 = 0 \end{cases} \quad \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + x_2 + x_3 = 2 \\ x_1 - x_2 + x_3 = 3 \end{cases} \quad \begin{cases} x_1 + x_2 = 1 \\ x_2 + x_3 = 2 \\ x_3 + x_4 = 3 \\ x_1 + 2x_2 + 2x_3 + x_4 = 0 \end{cases}$$

4. Montrer que si un système linéaire *homogène* a une solution $(x_1,\dots,x_p) \neq (0,\dots,0)$, alors il admet une infinité de solutions.

3. Résolution par la méthode du pivot de Gauss

3.1. Systèmes échelonnés

Définition 5

Un système est ***échelonné*** si :

- le nombre de coefficients nuls commençant une ligne croît strictement ligne après ligne.

Il est ***échelonné réduit*** si en plus :

- le premier coefficient non nul d'une ligne vaut 1 ;
- et c'est le seul élément non nul de sa colonne.

Exemple 6

- $\begin{cases} 2x_1 + 3x_2 + 2x_3 - x_4 = 5 \\ - x_2 - 2x_3 = 4 \\ 3x_4 = 1 \end{cases}$ est échelonné (mais pas réduit).
- $\begin{cases} 2x_1 + 3x_2 + 2x_3 - x_4 = 5 \\ - 2x_3 = 4 \\ \phantom{2x_1 + 3x_2 +{}} x_3 + x_4 = 1 \end{cases}$ n'est pas échelonné (la dernière ligne commence avec la même variable que la ligne au-dessus).

Il se trouve que les systèmes linéaires sous une forme échelonnée réduite sont particulièrement simples à résoudre.

Exemple 7

Le système linéaire suivant à 3 équations et 4 inconnues est échelonné et réduit.

$$\begin{cases} x_1 \phantom{{}+x_2} + 2x_3 = 25 \\ \phantom{x_1 +{}} x_2 - 2x_3 = 16 \\ \phantom{x_1 + x_2 + 2x_3 +{}} x_4 = 1 \end{cases}$$

Ce système se résout trivialement en

$$\begin{cases} x_1 = 25 - 2x_3 \\ x_2 = 16 + 2x_3 \\ x_4 = 1. \end{cases}$$

En d'autres termes, pour toute valeur de x_3 réelle, les valeurs de x_1, x_2 et x_4 calculées ci-dessus fournissent une solution du système, et on les a ainsi toutes obtenues. On peut donc décrire entièrement l'ensemble des solutions :

$$\mathscr{S} = \left\{ (25 - 2x_3, 16 + 2x_3, x_3, 1) \mid x_3 \in \mathbb{R} \right\}.$$

3.2. Opérations sur les équations d'un système

Nous allons utiliser trois opérations élémentaires sur les équations (c'est-à-dire sur les lignes) qui sont :

1. $L_i \leftarrow \lambda L_i$ avec $\lambda \neq 0$: on peut multiplier une équation par un réel non nul.
2. $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{R}$ (et $j \neq i$) : on peut ajouter à l'équation L_i un multiple d'une autre équation L_j.
3. $L_i \leftrightarrow L_j$: on peut échanger deux équations.

Ces trois opérations élémentaires ne changent pas les solutions d'un système linéaire ; autrement dit ces opérations transforment un système linéaire en un système linéaire équivalent.

Exemple 8

Utilisons ces opérations élémentaires pour résoudre le système suivant.

$$\begin{cases} x + y + 7z = -1 & (L_1) \\ 2x - y + 5z = -5 & (L_2) \\ -x - 3y - 9z = -5 & (L_3) \end{cases}$$

Commençons par l'opération $L_2 \leftarrow L_2 - 2L_1$: on soustrait à la deuxième équation deux fois la première équation. On obtient un système équivalent avec une nouvelle deuxième ligne (plus simple) :

$$\begin{cases} x + y + 7z = -1 \\ \quad -3y - 9z = -3 & L_2 \leftarrow L_2 - 2L_1 \\ -x - 3y - 9z = -5 \end{cases}$$

Puis $L_3 \leftarrow L_3 + L_1$:

$$\begin{cases} x + y + 7z = -1 \\ \quad -3y - 9z = -3 \\ \quad -2y - 2z = -6 & L_3 \leftarrow L_3 + L_1 \end{cases}$$

On continue pour faire apparaître un coefficient 1 en tête de la deuxième ligne ; pour cela on divise la ligne L_2 par -3 :

$$\begin{cases} x + y + 7z = -1 \\ \quad y + 3z = 1 & L_2 \leftarrow -\frac{1}{3} L_2 \\ \quad -2y - 2z = -6 \end{cases}$$

On continue ainsi

$$\begin{cases} x & +y & +7z & = & -1 \\ & y & +3z & = & 1 \\ & & 4z & = & -4 \end{cases} \quad {\scriptstyle L_3 \leftarrow L_3 + 2L_2} \qquad \begin{cases} x & +y & +7z & = & -1 \\ & y & +3z & = & 1 \\ & & z & = & -1 \end{cases} \quad {\scriptstyle L_3 \leftarrow \frac{1}{4} L_3}$$

$$\begin{cases} x & +y & +7z & = & -1 \\ & y & & = & 4 \\ & & z & = & -1 \end{cases} \quad {\scriptstyle L_2 \leftarrow L_2 - 3L_3} \qquad \begin{cases} x & +y & & = & 6 \\ & y & & = & 4 \\ & & z & = & -1 \end{cases} \quad {\scriptstyle L_1 \leftarrow L_1 - 7L_3}$$

On aboutit à un système réduit et échelonné :

$$\begin{cases} x & & & = & 2 \\ & y & & = & 4 \\ & & z & = & -1 \end{cases} \quad {\scriptstyle L_1 \leftarrow L_1 - L_2}$$

On obtient ainsi $x = 2$, $y = 4$ et $z = -1$ et l'unique solution du système est $(2, 4, -1)$.

La méthode utilisée pour cet exemple est reprise et généralisée dans le paragraphe suivant.

3.3. Méthode du pivot de Gauss

La méthode du pivot de Gauss permet de trouver les solutions de n'importe quel système linéaire. Nous allons décrire cet algorithme sur un exemple. Il s'agit d'une description précise d'une suite d'opérations à effectuer, qui dépendent de la situation et d'un ordre précis. Ce processus aboutit toujours (et en plus assez rapidement) à un système échelonné puis réduit, qui conduit immédiatement aux solutions du système.

Partie A. Passage à une forme échelonnée.

Soit le système suivant à résoudre :

$$\begin{cases} & -x_2 & +2x_3 & +13x_4 & = & 5 \\ x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 \\ -x_1 & +3x_2 & -3x_3 & -20x_4 & = & -1 \end{cases}$$

Pour appliquer la méthode du pivot de Gauss, il faut d'abord que le premier coefficient de la première ligne soit non nul. Comme ce n'est pas le cas ici, on échange les deux premières lignes par l'opération élémentaire $L_1 \leftrightarrow L_2$:

$$\begin{cases} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 \\ & -x_2 & +2x_3 & +13x_4 & = & 5 \\ -x_1 & +3x_2 & -3x_3 & -20x_4 & = & -1 \end{cases} \quad {\scriptstyle L_1 \leftrightarrow L_2}$$

Nous avons déjà un coefficient 1 devant le x_1 de la première ligne. On dit que nous avons un ***pivot*** en position $(1, 1)$ (première ligne, première colonne). Ce pivot sert de base pour éliminer tous les autres termes sur la même colonne.

Il n'y a pas de terme x_1 sur le deuxième ligne. Faisons disparaître le terme x_1 de la troisième ligne ; pour cela on fait l'opération élémentaire $L_3 \leftarrow L_3 + L_1$:

$$\begin{cases} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 \\ & -x_2 & +2x_3 & +13x_4 & = & 5 \\ & x_2 & & -3x_4 & = & 3 \end{cases} \quad {\scriptstyle L_3 \leftarrow L_3 + L_1}$$

On change le signe de la seconde ligne ($L_2 \leftarrow -L_2$) pour faire apparaître 1 au coefficient du pivot $(2,2)$ (deuxième ligne, deuxième colonne) :

$$\left\{\begin{array}{rrrrcrl} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 & \\ & x_2 & -2x_3 & -13x_4 & = & -5 & \scriptstyle L_2 \leftarrow -L_2 \\ & x_2 & & -3x_4 & = & 3 & \end{array}\right.$$

On fait disparaître le terme x_2 de la troisième ligne, puis on fait apparaître un coefficient 1 pour le pivot de la position $(3,3)$:

$$\left\{\begin{array}{rrrrcrl} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 & \\ & x_2 & -2x_3 & -13x_4 & = & -5 & \\ & & 2x_3 & +10x_4 & = & 8 & \scriptstyle L_3 \leftarrow L_3 - L_2 \end{array}\right. \qquad \left\{\begin{array}{rrrrcrl} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 & \\ & x_2 & -2x_3 & -13x_4 & = & -5 & \\ & & x_3 & +5x_4 & = & 4 & \scriptstyle L_3 \leftarrow \frac{1}{2} L_3 \end{array}\right.$$

Le système est maintenant sous forme échelonnée.

Partie B. Passage à une forme réduite.
Il reste à le mettre sous la forme échelonnée réduite. Pour cela, on ajoute à une ligne des multiples adéquats des lignes situées au-dessous d'elle, en allant du bas à droite vers le haut à gauche.
On fait apparaître des 0 sur la troisième colonne en utilisant le pivot de la troisième ligne :

$$\left\{\begin{array}{rrrrcrl} x_1 & -2x_2 & +3x_3 & +17x_4 & = & 4 & \\ & x_2 & & -3x_4 & = & 3 & \scriptstyle L_2 \leftarrow L_2 + 2L_3 \\ & & x_3 & +5x_4 & = & 4 & \end{array}\right. \qquad \left\{\begin{array}{rrrrcrl} x_1 & -2x_2 & & 2x_4 & = & -8 & \scriptstyle L_1 \leftarrow L_1 - 3L_3 \\ & x_2 & & -3x_4 & = & 3 & \\ & & x_3 & +5x_4 & = & 4 & \end{array}\right.$$

On fait apparaître des 0 sur la deuxième colonne (en utilisant le pivot de la deuxième ligne) :

$$\left\{\begin{array}{rrrrcrl} x_1 & & & -4x_4 & = & -2 & \scriptstyle L_1 \leftarrow L_1 + 2L_2 \\ & x_2 & & -3x_4 & = & 3 & \\ & & x_3 & +5x_4 & = & 4 & \end{array}\right.$$

Le système est sous forme échelonnée réduite.

Partie C. Solutions. Le système est maintenant très simple à résoudre. En choisissant x_4 comme variable libre, on peut exprimer x_1, x_2, x_3 en fonction de x_4 :

$$x_1 = 4x_4 - 2, \quad x_2 = 3x_4 + 3, \quad x_3 = -5x_4 + 4.$$

Ce qui permet d'obtenir toutes les solutions du système :

$$\mathscr{S} = \left\{(4x_4 - 2, 3x_4 + 3, -5x_4 + 4, x_4) \mid x_4 \in \mathbb{R}\right\}.$$

3.4. Systèmes homogènes

Le fait que l'on puisse toujours se ramener à un système échelonné réduit implique le résultat suivant :

Théorème 2

Tout système homogène d'équations linéaires dont le nombre d'inconnues est strictement plus grand que le nombre d'équations a une infinité de solutions.

Exemple 9

Considérons le système homogène

$$\left\{\begin{array}{rcrcrcrcrcl} 3x_1 & + & 3x_2 & - & 2x_3 & & & - & x_5 & = & 0\\ -x_1 & - & x_2 & + & x_3 & + & 3x_4 & + & x_5 & = & 0\\ 2x_1 & + & 2x_2 & - & x_3 & + & 2x_4 & + & 2x_5 & = & 0\\ & & & & x_3 & + & 8x_4 & + & 4x_5 & = & 0. \end{array}\right.$$

Sa forme échelonnée réduite est

$$\left\{\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & & & & & + & 13x_5 & = & 0\\ & & & & x_3 & & & + & 20x_5 & = & 0\\ & & & & & & x_4 & - & 2x_5 & = & 0. \end{array}\right.$$

On pose comme variables libres x_2 et x_5 pour avoir

$$x_1 = -x_2 - 13x_5, \qquad x_3 = -20x_5, \qquad x_4 = 2x_5,$$

et l'ensemble des solutions :

$$\mathscr{S} = \left\{(-x_2 - 13x_5, x_2, -20x_5, 2x_5, x_5) \mid x_2, x_5 \in \mathbb{R}\right\}$$

qui est bien infini.

Mini-exercices

1. Écrire un système linéaire à 4 équations et 5 inconnues qui soit échelonné mais pas réduit. Idem avec échelonné, non réduit, dont tous les coefficients sont 0 ou +1. Idem avec échelonné et réduit.
2. Résoudre les systèmes échelonnés suivants : $\left\{\begin{array}{rrrrcr} 2x_1 & -x_2 & & +x_4 & = & 1\\ & x_2 & +x_3 & -2x_4 & = & 3\\ & & 2x_3 & +x_4 & = & 4\\ & & & x_4 & = & -2 \end{array}\right.$

   $\left\{\begin{array}{rrrrcr} x_1 & +x_2 & & +x_4 & = & 0\\ & x_2 & +x_3 & & = & 0\\ & & 2x_3 & +x_4 & = & 0 \end{array}\right. \qquad \left\{\begin{array}{rrrrcr} x_1 & +2x_2 & & +x_4 & = & 0\\ & & 2x_3 & -3x_4 & = & 0 \end{array}\right.$
3. Si l'on passe d'un système (S) par une des trois opérations élémentaires à un système (S'), alors quelle opération permet de passer de (S') à (S) ?
4. Résoudre les systèmes linéaires suivants par la méthode du pivot de Gauss :

   $$\left\{\begin{array}{rcrcrcr} 2x & + & y & + & z & = & 3\\ x & - & y & + & 3z & = & 8\\ x & + & 2y & - & z & = & -3 \end{array}\right. \qquad \left\{\begin{array}{rcrcrcrcr} 2x_1 & + & 4x_2 & - & 6x_3 & - & 2x_4 & = & 2\\ 3x_1 & + & 6x_2 & - & 7x_3 & + & 4x_4 & = & 2\\ 5x_1 & + & 10x_2 & - & 11x_3 & + & 6x_4 & = & 3 \end{array}\right.$$
5. Résoudre le système suivant, selon les valeurs de $a, b \in \mathbb{R}$: $\left\{\begin{array}{rrrcr} x & +y & -z & = & a\\ -x & & +2z & = & b\\ & 2y & +2z & = & 4 \end{array}\right.$

Auteurs

- D'après un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- et un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties d'un cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- mixés et révisés par Arnaud Bodin, relu par Vianney Combet.

Exo7

2 Matrices

Vidéo ■ partie 1. Définition
Vidéo ■ partie 2. Multiplication de matrices
Vidéo ■ partie 3. Inverse d'une matrice : définition
Vidéo ■ partie 4. Inverse d'une matrice : calcul
Vidéo ■ partie 5. Inverse d'une matrice : systèmes linéaires et matrices élémentaires
Vidéo ■ partie 6. Matrices triangulaires, transposition, trace, matrices symétriques

Les matrices sont des tableaux de nombres. La résolution d'un certain nombre de problèmes d'algèbre linéaire se ramène à des manipulations sur les matrices. Ceci est vrai en particulier pour la résolution des systèmes linéaires.

Dans ce chapitre, $\mathbb{K}$ désigne un corps. On peut penser à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

1. Définition

1.1. Définition

Définition 6

- Une ***matrice***matrice A est un tableau rectangulaire d'éléments de $\mathbb{K}$.
- Elle est dite de ***taille*** $n \times p$ si le tableau possède n lignes et p colonnes.
- Les nombres du tableau sont appelés les ***coefficients*** de A.
- Le coefficient situé à la i-ème ligne et à la j-ème colonne est noté $a_{i,j}$.

Un tel tableau est représenté de la manière suivante :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i,1} & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,j} & \dots & a_{n,p} \end{pmatrix} \qquad \text{ou} \qquad A = \left(a_{i,j}\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} \qquad \text{ou} \qquad \left(a_{i,j}\right).$$

Exemple 10

$$A = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 3 & 7 \end{pmatrix}$$

est une matrice 2×3 avec, par exemple, $a_{1,1} = 1$ et $a_{2,3} = 7$.

Encore quelques définitions :

Définition 7

- Deux matrices sont ***égales*** lorsqu'elles ont la même taille et que les coefficients correspondants sont égaux.
- L'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est noté $M_{n,p}(\mathbb{K})$. Les éléments de $M_{n,p}(\mathbb{R})$ sont appelés ***matrices réelles***.

1.2. Matrices particulières

Voici quelques types de matrices intéressantes :

- Si $n = p$ (même nombre de lignes que de colonnes), la matrice est dite ***matrice carrée***. On note $M_n(\mathbb{K})$ au lieu de $M_{n,n}(\mathbb{K})$.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}$$

  Les éléments $a_{1,1}, a_{2,2}, \dots, a_{n,n}$ forment la ***diagonale principale*** de la matrice.
- Une matrice qui n'a qu'une seule ligne ($n = 1$) est appelée ***matrice ligne*** ou ***vecteur ligne***. On la note

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \end{pmatrix}.$$

- De même, une matrice qui n'a qu'une seule colonne ($p = 1$) est appelée ***matrice colonne*** ou ***vecteur colonne***. On la note

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

- La matrice (de taille $n \times p$) dont tous les coefficients sont des zéros est appelée la ***matrice nulle*** et est notée $0_{n,p}$ ou plus simplement 0. Dans le calcul matriciel, la matrice nulle joue le rôle du nombre 0 pour les réels.

1.3. Addition de matrices

Définition 8. Somme de deux matrices

somme ! de matrices Soient A et B deux matrices ayant la même taille $n \times p$. Leur ***somme*** $C = A + B$ est la matrice de taille $n \times p$ définie par

$$c_{ij} = a_{ij} + b_{ij}.$$

En d'autres termes, on somme coefficients par coefficients. Remarque : on note indifféremment a_{ij} où $a_{i,j}$ pour les coefficients de la matrice A.

Exemple 11

$$\text{Si} \quad A = \begin{pmatrix} 3 & -2 \\ 1 & 7 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 5 \\ 2 & -1 \end{pmatrix} \quad \text{alors} \quad A + B = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}.$$

$$\text{Par contre si} \quad B' = \begin{pmatrix} -2 \\ 8 \end{pmatrix} \quad \text{alors} \quad A + B' \quad \text{n'est pas définie.}$$

Définition 9. Produit d'une matrice par un scalaire

Le produit d'une matrice $A = \left(a_{ij}\right)$ de $M_{n,p}(\mathbb{K})$ par un scalaire $\alpha \in \mathbb{K}$ est la matrice $\left(\alpha a_{ij}\right)$ formée en multipliant chaque coefficient de A par α. Elle est notée $\alpha \cdot A$ (ou simplement αA).

Exemple 12

$$\text{Si} \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad \alpha = 2 \quad \text{alors} \quad \alpha A = \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 0 \end{pmatrix}.$$

La matrice $(-1)A$ est l'***opposée*** de A et est notée $-A$. La ***différence*** $A - B$ est définie par $A + (-B)$.

Exemple 13

$$\text{Si} \quad A = \begin{pmatrix} 2 & -1 & 0 \\ 4 & -5 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} -1 & 4 & 2 \\ 7 & -5 & 3 \end{pmatrix} \quad \text{alors} \quad A - B = \begin{pmatrix} 3 & -5 & -2 \\ -3 & 0 & -1 \end{pmatrix}.$$

L'addition et la multiplication par un scalaire se comportent sans surprises :

Proposition 1

Soient A, B et C trois matrices appartenant à $M_{n,p}(\mathbb{K})$. Soient $\alpha \in \mathbb{K}$ et $\beta \in \mathbb{K}$ deux scalaires.

1. $A + B = B + A$: la somme est commutative,
2. $A + (B + C) = (A + B) + C$: la somme est associative,
3. $A + 0 = A$: la matrice nulle est l'élément neutre de l'addition,
4. $(\alpha + \beta)A = \alpha A + \beta A$,
5. $\alpha(A + B) = \alpha A + \alpha B$.

Démonstration

Prouvons par exemple le quatrième point. Le terme général de $(\alpha+\beta)A$ est égal à $(\alpha+\beta)a_{ij}$. D'après les règles de calcul dans $\mathbb{K}$, $(\alpha+\beta)a_{ij}$ est égal à $\alpha a_{ij}+\beta a_{ij}$ qui est le terme général de la matrice $\alpha A+\beta A$.

Mini-exercices

1. Soient $A=\begin{pmatrix} -7 & 2 \\ 0 & -1 \\ 1 & -4 \end{pmatrix}$, $B=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{pmatrix}$, $C=\begin{pmatrix} 21 & -6 \\ 0 & 3 \\ -3 & 12 \end{pmatrix}$, $D=\frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $E=\begin{pmatrix} 1 & 2 \\ -3 & 0 \\ -8 & 6 \end{pmatrix}$. Calculer toutes les sommes possibles de deux de ces matrices. Calculer $3A+2C$ et $5B-4D$. Trouver α tel que $A-\alpha C$ soit la matrice nulle.
2. Montrer que si $A+B=A$, alors B est la matrice nulle.
3. Que vaut $0\cdot A$? et $1\cdot A$? Justifier l'affirmation : $\alpha(\beta A)=(\alpha\beta)A$. Idem avec $nA = A+A+\cdots+A$ (n occurrences de A).

2. Multiplication de matrices

2.1. Définition du produit

Le produit AB de deux matrices A et B est défini si et seulement si le nombre de colonnes de A est égal au nombre de lignes de B.

Définition 10. Produit de deux matrices

produit !matriciel Soient $A=(a_{ij})$ une matrice $n\times p$ et $B=(b_{ij})$ une matrice $p\times q$. Alors le produit $C=AB$ est une matrice $n\times q$ dont les coefficients c_{ij} sont définis par :

$$\boxed{c_{ij}=\sum_{k=1}^{p} a_{ik}b_{kj}}$$

On peut écrire le coefficient de façon plus développée, à savoir :

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ik}b_{kj}+\cdots+a_{ip}b_{pj}.$$

Il est commode de disposer les calculs de la façon suivante.

$$\begin{array}{rcl} & \begin{pmatrix} & & & \times & \\ & & & \times & \\ & & & \times & \\ & & & \times & \end{pmatrix} & \leftarrow B \\ A \rightarrow \begin{pmatrix} & & & \\ \times & \times & \times & \times \end{pmatrix} & \begin{pmatrix} & & & | & \\ & & & | & \\ - & - & - & c_{ij} & \end{pmatrix} & \leftarrow AB \end{array}$$

Avec cette disposition, on considère d'abord la ligne de la matrice A située à gauche du coefficient que l'on veut calculer (ligne représentée par des $\times$ dans A) et aussi la colonne de la matrice B située au-dessus du coefficient que l'on veut calculer (colonne représentée par des $\times$ dans B). On calcule le produit du premier coefficient de la ligne par le premier coefficient de la colonne ($a_{i1} \times b_{1j}$), que l'on ajoute au produit du deuxième coefficient de la ligne par le deuxième coefficient de la colonne ($a_{i2} \times b_{2j}$), que l'on ajoute au produit du troisième...

2.2. Exemples

Exemple 14

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

On dispose d'abord le produit correctement (à gauche) : la matrice obtenue est de taille 2×2. Puis on calcule chacun des coefficients, en commençant par le premier coefficient $c_{11} = 1 \times 1 + 2 \times (-1) + 3 \times 1 = 2$ (au milieu), puis les autres (à droite).

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array} \qquad \begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} 2 & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array} \qquad \begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} 2 & 7 \\ 3 & 11 \end{pmatrix} \end{array}$$

Un exemple intéressant est le produit d'un vecteur ligne par un vecteur colonne :

$$u = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \qquad v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Alors $u \times v$ est une matrice de taille 1×1 dont l'unique coefficient est $a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$. Ce nombre s'appelle le ***produit scalaire*** des vecteurs u et v.

Calculer le coefficient c_{ij} dans le produit $A \times B$ revient donc à calculer le produit scalaire des vecteurs formés par la i-ème ligne de A et la j-ème colonne de B.

2.3. Pièges à éviter

Premier piège. Le produit de matrices n'est pas commutatif en général.
En effet, il se peut que AB soit défini mais pas BA, ou que AB et BA soient tous deux définis mais pas de la même taille. Mais même dans le cas où AB et BA sont définis et de la même taille, on a en général $AB \neq BA$.

Exemple 15

$$\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 3 \\ -2 & -6 \end{pmatrix} \qquad \text{mais} \qquad \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 29 & -2 \end{pmatrix}.$$

Deuxième piège. $AB = 0$ **n'implique pas** $A = 0$ **ou** $B = 0$**.**
Il peut arriver que le produit de deux matrices non nulles soit nul. En d'autres termes, on peut avoir $A \neq 0$ et $B \neq 0$ mais $AB = 0$.

Exemple 16

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Troisième piège. $AB = AC$ **n'implique pas** $B = C$. On peut avoir $AB = AC$ et $B \neq C$.

Exemple 17

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & -1 \\ 5 & 4 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 5 \\ 5 & 4 \end{pmatrix} \qquad \text{et} \qquad AB = AC = \begin{pmatrix} -5 & -4 \\ 15 & 12 \end{pmatrix}.$$

2.4. Propriétés du produit de matrices

Malgré les difficultés soulevées au-dessus, le produit vérifie les propriétés suivantes :

Proposition 2

1. $A(BC) = (AB)C$: associativité du produit,
2. $A(B+C) = AB + AC$ et $(B+C)A = BA + CA$: distributivité du produit par rapport à la somme,
3. $A \cdot 0 = 0$ et $0 \cdot A = 0$.

Démonstration

Posons $A = (a_{ij}) \in M_{n,p}(\mathbb{K})$, $B = (b_{ij}) \in M_{p,q}(\mathbb{K})$ et $C = (c_{ij}) \in M_{q,r}(\mathbb{K})$. Prouvons que $A(BC) = (AB)C$ en montrant que les matrices $A(BC)$ et $(AB)C$ ont les mêmes coefficients.

Le terme d'indice (i,k) de la matrice AB est $x_{ik} = \sum_{\ell=1}^{p} a_{i\ell} b_{\ell k}$. Le terme d'indice (i,j) de la matrice $(AB)C$ est donc

$$\sum_{k=1}^{q} x_{ik} c_{kj} = \sum_{k=1}^{q} \left(\sum_{\ell=1}^{p} a_{i\ell} b_{\ell k} \right) c_{kj}.$$

Le terme d'indice (ℓ,j) de la matrice BC est $y_{\ell j} = \sum_{k=1}^{q} b_{\ell k} c_{kj}$. Le terme d'indice (i,j) de la matrice $A(BC)$ est donc

$$\sum_{\ell=1}^{p} a_{i\ell} \left(\sum_{k=1}^{q} b_{\ell k} c_{kj} \right).$$

Comme dans $\mathbb{K}$ la multiplication est distributive et associative, les coefficients de $(AB)C$ et $A(BC)$ coïncident. Les autres démonstrations se font comme celle de l'associativité.

2.5. La matrice identité

Ide@I_n, I - matrice identité matrice!identité La matrice carrée suivante s'appelle ***la matrice identité*** :

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

Ses éléments diagonaux sont égaux à 1 et tous ses autres éléments sont égaux à 0. Elle se note I_n ou simplement I. Dans le calcul matriciel, la matrice identité joue un rôle analogue à celui du nombre 1 pour les réels. C'est l'élément neutre pour la multiplication. En d'autres termes :

Proposition 3

Si A est une matrice $n \times p$, alors

$$I_n \cdot A = A \qquad \text{et} \qquad A \cdot I_p = A.$$

Démonstration

Nous allons détailler la preuve. Soit $A \in M_{n,p}(\mathbb{K})$ de terme général a_{ij}. La matrice unité d'ordre p est telle que tous les éléments de la diagonale principale sont égaux à 1, les autres étant tous nuls. On peut formaliser cela en introduisant le symbole de Kronecker. Si i et j sont deux entiers, on appelle ***symbole de Kronecker***, et on note $\delta_{i,j}$, le réel qui vaut 0 si i est différent de j, et 1 si i est égal à j. Donc

$$\delta_{i,j} = \begin{cases} 0 & \text{si } \; i \neq j \\ 1 & \text{si } \; i = j. \end{cases}$$

Alors le terme général de la matrice identité I_p est $\delta_{i,j}$ avec i et j entiers, compris entre 1 et p.
La matrice produit AI_p est une matrice appartenant à $M_{n,p}(\mathbb{K})$ dont le terme général c_{ij} est donné par la formule $c_{ij} = \sum_{k=1}^{p} a_{ik}\delta_{kj}$. Dans cette somme, i et j sont fixés et k prend toutes les valeurs comprises entre 1 et p. Si $k \neq j$ alors $\delta_{kj} = 0$, et si $k = j$ alors $\delta_{kj} = 1$. Donc dans la somme qui définit c_{ij}, tous les termes correspondant à des valeurs de k différentes de j sont nuls et il reste donc $c_{ij} = a_{ij}\delta_{jj} = a_{ij}1 = a_{ij}$. Donc les matrices AI_p et A ont le même terme général et sont donc égales. L'égalité $I_nA = A$ se démontre de la même façon.

2.6. Puissance d'une matrice

Dans l'ensemble $M_n(\mathbb{K})$ des matrices carrées de taille $n \times n$ à coefficients dans $\mathbb{K}$, la multiplication des matrices est une opération interne : si $A, B \in M_n(\mathbb{K})$ alors $AB \in M_n(\mathbb{K})$.
En particulier, on peut multiplier une matrice carrée par elle-même : on note $A^2 = A \times A$, $A^3 = A \times A \times A$.
On peut ainsi définir les puissances successives d'une matrice :

Définition 11

Pour tout $A \in M_n(\mathbb{K})$, on définit les puissances successives de A par $A^0 = I_n$ et $A^{p+1} = A^p \times A$ pour tout $p \in \mathbb{N}$. Autrement dit, $A^p = \underbrace{A \times A \times \cdots \times A}_{p \text{ facteurs}}$.

Exemple 18

On cherche à calculer A^p avec $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. On calcule A^2, A^3 et A^4 et on obtient :

$$A^2 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad A^3 = A^2 \times A = \begin{pmatrix} 1 & 0 & 7 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix} \qquad A^4 = A^3 \times A = \begin{pmatrix} 1 & 0 & 15 \\ 0 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix}.$$

L'observation de ces premières puissances permet de penser que la formule est : $A^p = \begin{pmatrix} 1 & 0 & 2^p - 1 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 2^p \end{pmatrix}$. Démontrons ce résultat par récurrence.
Il est vrai pour $p = 0$ (on trouve l'identité). On le suppose vrai pour un entier p et on va le démontrer pour $p+1$. On a, d'après la définition,

$$A^{p+1} = A^p \times A = \begin{pmatrix} 1 & 0 & 2^p - 1 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 2^p \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2^{p+1} - 1 \\ 0 & (-1)^{p+1} & 0 \\ 0 & 0 & 2^{p+1} \end{pmatrix}.$$

Donc la propriété est démontrée.

2.7. Formule du binôme

Comme la multiplication n'est pas commutative, les identités binomiales usuelles sont fausses. En particulier, $(A+B)^2$ ne vaut en général pas $A^2 + 2AB + B^2$, mais on sait seulement que

$$(A+B)^2 = A^2 + AB + BA + B^2.$$

Proposition 4. Calcul de $(A+B)^p$ lorsque $AB = BA$

Soient A et B deux éléments de $M_n(\mathbb{K})$ qui ***commutent***, c'est-à-dire tels que $AB = BA$. Alors, pour tout entier $p \geqslant 0$, on a la formule

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^{p-k} B^k$$

où $\binom{p}{k}$ désigne le coefficient du binôme.

La démonstration est similaire à celle de la formule du binôme pour $(a+b)^p$, avec $a, b \in \mathbb{R}$.

Exemple 19

Soit $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. On pose $N = A - I = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. La matrice N est nilpotente (c'est-à-dire il existe $k \in \mathbb{N}$ tel que $N^k = 0$) comme le montrent les calculs suivants :

$$N^2 = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad N^3 = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad N^4 = 0.$$

Comme on a $A = I + N$ et les matrices N et I commutent (la matrice identité commute avec toutes les matrices), on peut appliquer la formule du binôme de Newton. On utilise que $I^k = I$ pour tout k et surtout que $N^k = 0$ si $k \geqslant 4$. On obtient

$$A^p = \sum_{k=0}^{p} \binom{p}{k} N^k I^{p-k} = \sum_{k=0}^{3} \binom{p}{k} N^k = I + pN + \tfrac{p(p-1)}{2!} N^2 + \tfrac{p(p-1)(p-2)}{3!} N^3.$$

D'où

$$A^p = \begin{pmatrix} 1 & p & p^2 & p(p^2-p+1) \\ 0 & 1 & 2p & p(3p-2) \\ 0 & 0 & 1 & 3p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Mini-exercices

1. Soient $A = \left(\begin{smallmatrix} 0 & 2 & -2 \\ 6 & -4 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & -2 & -3 \end{smallmatrix}\right)$, $C = \left(\begin{smallmatrix} 8 & 2 \\ -3 & 2 \\ -5 & 5 \end{smallmatrix}\right)$, $D = \left(\begin{smallmatrix} 5 \\ 2 \\ -1 \end{smallmatrix}\right)$, $E = \begin{pmatrix} x & y & z \end{pmatrix}$. Quels produits sont possibles ? Les calculer !
2. Soient $A = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \end{smallmatrix}\right)$. Calculer A^2, B^2, AB et BA.
3. Soient $A = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \end{smallmatrix}\right)$. Calculer A^p et B^p pour tout $p \geqslant 0$. Montrer que $AB = BA$. Calculer $(A+B)^p$.

3. Inverse d'une matrice : définition

3.1. Définition

Définition 12. Matrice inverse

matrice !inverse matrice !inversible Soit A une matrice carrée de taille $n \times n$. S'il existe une matrice carrée B de taille $n \times n$ telle que

$$AB = I \qquad \text{et} \qquad BA = I,$$

on dit que A est ***inversible***. On appelle B l'***inverse de A*** et on la note A^{-1}.

On verra plus tard qu'il suffit en fait de vérifier une seule des conditions $AB = I$ ou bien $BA = I$.

- Plus généralement, quand A est inversible, pour tout $p \in \mathbb{N}$, on note :

$$A^{-p} = (A^{-1})^p = \underbrace{A^{-1}A^{-1}\cdots A^{-1}}_{p \text{ facteurs}}.$$

- L'ensemble des matrices inversibles de $M_n(\mathbb{K})$ est noté $GL_n(\mathbb{K})$.

3.2. Exemples

Exemple 20

Soit $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right)$. Étudier si A est inversible, c'est étudier l'existence d'une matrice $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ à coefficients dans $\mathbb{K}$, telle que $AB = I$ et $BA = I$. Or $AB = I$ équivaut à :

$$AB = I \iff \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} a+2c & b+2d \\ 3c & 3d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Cette égalité équivaut au système :

$$\begin{cases} a+2c=1 \\ b+2d=0 \\ 3c=0 \\ 3d=1 \end{cases}$$

Sa résolution est immédiate : $a=1$, $b=-\frac{2}{3}$, $c=0$, $d=\frac{1}{3}$. Il n'y a donc qu'une seule matrice possible, à savoir $B=\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$. Pour prouver qu'elle convient, il faut aussi montrer l'égalité $BA=I$, dont la vérification est laissée au lecteur. La matrice A est donc inversible et $A^{-1}=\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$.

Exemple 21

La matrice $A=\left(\begin{smallmatrix} 3 & 0 \\ 5 & 0 \end{smallmatrix}\right)$ n'est pas inversible. En effet, soit $B=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice quelconque. Alors le produit

$$BA=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 5 & 0 \end{pmatrix}=\begin{pmatrix} 3a+5b & 0 \\ 3c+5d & 0 \end{pmatrix}$$

ne peut jamais être égal à la matrice identité.

Exemple 22

- Soit I_n la matrice carrée identité de taille $n\times n$. C'est une matrice inversible, et son inverse est elle-même par l'égalité $I_nI_n=I_n$.
- La matrice nulle 0_n de taille $n\times n$ n'est pas inversible. En effet on sait que, pour toute matrice B de $M_n(\mathbb{K})$, on a $B0_n=0_n$, qui ne peut jamais être la matrice identité.

3.3. Propriétés

Unicité

Proposition 5

Si A est inversible, alors son inverse est unique.

Démonstration

La méthode classique pour mener à bien une telle démonstration est de supposer l'existence de deux matrices B_1 et B_2 satisfaisant aux conditions imposées et de démontrer que $B_1=B_2$.
Soient donc B_1 telle que $AB_1=B_1A=I_n$ et B_2 telle que $AB_2=B_2A=I_n$. Calculons $B_2(AB_1)$. D'une part, comme $AB_1=I_n$, on a $B_2(AB_1)=B_2$. D'autre part, comme le produit des matrices est associatif, on a $B_2(AB_1)=(B_2A)B_1=I_nB_1=B_1$. Donc $B_1=B_2$.

Inverse de l'inverse

Proposition 6

Soit A une matrice inversible. Alors A^{-1} est aussi inversible et on a :

$$\boxed{(A^{-1})^{-1} = A}$$

Inverse d'un produit

Proposition 7

Soient A et B deux matrices inversibles de même taille. Alors AB est inversible et

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

Il faut bien faire attention à l'inversion de l'ordre !

Démonstration

Il suffit de montrer $(B^{-1}A^{-1})(AB) = I$ et $(AB)(B^{-1}A^{-1}) = I$. Cela suit de

$$\begin{aligned} & (B^{-1}A^{-1})(AB) = B^{-1}(AA^{-1})B = B^{-1}IB = B^{-1}B = I, \\ \text{et} \quad & (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I. \end{aligned}$$

De façon analogue, on montre que si $A_1,\ldots,A_m$ sont inversibles, alors

$$(A_1A_2\cdots A_m)^{-1} = A_m^{-1}A_{m-1}^{-1}\cdots A_1^{-1}.$$

Simplification par une matrice inversible

Si C est une matrice quelconque de $M_n(\mathbb{K})$, nous avons vu que la relation $AC = BC$ où A et B sont des éléments de $M_n(\mathbb{K})$ n'entraîne pas forcément l'égalité $A = B$. En revanche, si C est une matrice inversible, on a la proposition suivante :

Proposition 8

Soient A et B deux matrices de $M_n(\mathbb{K})$ et C une matrice inversible de $M_n(\mathbb{K})$. Alors l'égalité $AC = BC$ implique l'égalité $A = B$.

Démonstration

Ce résultat est immédiat : si on multiplie à droite l'égalité $AC = BC$ par C^{-1}, on obtient l'égalité : $(AC)C^{-1} = (BC)C^{-1}$. En utilisant l'associativité du produit des matrices on a $A(CC^{-1}) = B(CC^{-1})$, ce qui donne d'après la définition de l'inverse $AI = BI$, d'où $A = B$.

Mini-exercices

1. Soient $A = \left(\begin{smallmatrix} -1 & -2 \\ 3 & 4 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 2 & 1 \\ 5 & 3 \end{smallmatrix}\right)$. Calculer A^{-1}, B^{-1} , $(AB)^{-1}$, $(BA)^{-1}$, A^{-2}.
2. Calculer l'inverse de $\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{smallmatrix}\right)$.

3. Soit $A = \begin{pmatrix} -1 & -2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer $2A - A^2$. Sans calculs, en déduire A^{-1}.

4. Inverse d'une matrice : calcul

Nous allons voir une méthode pour calculer l'inverse d'une matrice quelconque de manière efficace. Cette méthode est une reformulation de la méthode du pivot de Gauss pour les systèmes linéaires. Auparavant, nous commençons par une formule directe dans le cas simple des matrices 2×2.

4.1. Matrices 2×2

Considérons la matrice 2×2 : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Proposition 9

Si $ad - bc \neq 0$, alors A est inversible et

$$\boxed{A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

Démonstration

On vérifie que si $B = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ alors $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Idem pour BA.

4.2. Méthode de Gauss pour inverser les matrices

La méthode pour inverser une matrice A consiste à faire des opérations élémentaires sur les lignes de la matrice A jusqu'à la transformer en la matrice identité I. On fait simultanément les mêmes opérations élémentaires en partant de la matrice I. On aboutit alors à une matrice qui est A^{-1}. La preuve sera vue dans la section suivante.

En pratique, on fait les deux opérations en même temps en adoptant la disposition suivante : à côté de la matrice A que l'on veut inverser, on rajoute la matrice identité pour former un tableau $(A \mid I)$. Sur les lignes de cette matrice augmentée, on effectue des opérations élémentaires jusqu'à obtenir le tableau $(I \mid B)$. Et alors $B = A^{-1}$.

Ces opérations élémentaires sur les lignes sont :

1. $L_i \leftarrow \lambda L_i$ avec $\lambda \neq 0$: on peut multiplier une ligne par un réel non nul (ou un élément de $\mathbb{K} \setminus \{0\}$).
2. $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{K}$ (et $j \neq i$) : on peut ajouter à la ligne L_i un multiple d'une autre ligne L_j.
3. $L_i \leftrightarrow L_j$: on peut échanger deux lignes.

N'oubliez pas : tout ce que vous faites sur la partie gauche de la matrice augmentée, vous devez aussi le faire sur la partie droite.

4.3. Un exemple

Calculons l'inverse de $A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}$.
Voici la matrice augmentée, avec les lignes numérotées :

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 4 & 0 & -1 & 0 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array}$$

On applique la méthode de Gauss pour faire apparaître des 0 sur la première colonne, d'abord sur la deuxième ligne par l'opération élémentaire $L_2 \leftarrow L_2 - 4L_1$ qui conduit à la matrice augmentée :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \; L_2 \leftarrow L_2 - 4L_1$$

Puis un 0 sur la première colonne, à la troisième ligne, avec $L_3 \leftarrow L_3 + L_1$:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) \; L_3 \leftarrow L_3 + L_1$$

On multiplie la ligne L_2 afin qu'elle commence par 1 :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) \; L_2 \leftarrow -\tfrac{1}{8}L_2$$

On continue afin de faire apparaître des 0 partout sous la diagonale, et on multiplie la ligne L_3. Ce qui termine la première partie de la méthode de Gauss :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{2} & -1 & \frac{1}{2} & 1 \end{array}\right) \; L_3 \leftarrow L_3 - 4L_2 \qquad \text{puis} \qquad \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \; L_3 \leftarrow 2L_3$$

Il ne reste plus qu'à « remonter » pour faire apparaître des zéros au-dessus de la diagonale :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \; L_2 \leftarrow L_2 - \tfrac{5}{8}L_3 \qquad \text{puis} \qquad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \; L_1 \leftarrow L_1 - 2L_2 - L_3$$

Ainsi l'inverse de A est la matrice obtenue à droite et après avoir factorisé tous les coefficients par $\frac{1}{4}$, on a obtenu :

$$A^{-1} = \frac{1}{4}\begin{pmatrix} -2 & 2 & 2 \\ 7 & -3 & -5 \\ -8 & 4 & 8 \end{pmatrix}$$

Pour se rassurer sur ses calculs, on n'oublie pas de vérifier rapidement que $A \times A^{-1} = I$.

Mini-exercices

1. Si possible calculer l'inverse des matrices : $\left(\begin{smallmatrix} 3 & 1 \\ 7 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & -3 \\ -5 & 4 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & 2 \\ 3 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} \alpha+1 & 1 \\ 2 & \alpha \end{smallmatrix}\right)$.
2. Soit $A(\theta) = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Calculer $A(\theta)^{-1}$.
3. Calculer l'inverse des matrices : $\left(\begin{smallmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \\ -2 & 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & -2 & 1 \\ 3 & 0 & 5 \\ 1 & 1 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 0 \\ -1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 3 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 5 & 3 \end{smallmatrix}\right)$.

5. Inverse d'une matrice : systèmes linéaires et matrices élémentaires

5.1. Matrices et systèmes linéaires

Le système linéaire

$$\begin{cases} a_{11}\,x_1 & + & a_{12}\,x_2 & + & \cdots & + & a_{1p}\,x_p & = & b_1 \\ a_{21}\,x_1 & + & a_{22}\,x_2 & + & \cdots & + & a_{2p}\,x_p & = & b_2 \\ & & \cdots & & & & & & \\ a_{n1}\,x_1 & + & a_{n2}\,x_2 & + & \cdots & + & a_{np}\,x_p & = & b_n \end{cases}$$

peut s'écrire sous forme matricielle :

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1p} \\ a_{21} & \dots & a_{2p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{B}.$$

On appelle $A \in M_{n,p}(\mathbb{K})$ la matrice des coefficients du système. $B \in M_{n,1}(\mathbb{K})$ est le vecteur du second membre. Le vecteur $X \in M_{p,1}(\mathbb{K})$ est une solution du système si et seulement si

$$AX = B.$$

Nous savons que :

Théorème 3

Un système d'équations linéaires n'a soit aucune solution, soit une seule solution, soit une infinité de solutions.

5.2. Matrices inversibles et systèmes linéaires

Considérons le cas où le nombre d'équations égale le nombre d'inconnues :

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{B}.$$

Alors $A \in M_n(\mathbb{K})$ est une matrice carrée et B un vecteur de $M_{n,1}(\mathbb{K})$. Pour tout second membre, nous pouvons utiliser les matrices pour trouver la solution du système linéaire.

Proposition 10

Si la matrice A est inversible, alors la solution du système $AX = B$ est unique et est :

$$\boxed{X = A^{-1}B.}$$

La preuve est juste de vérifier que si $X = A^{-1}B$, alors $AX = A\left(A^{-1}B\right) = \left(AA^{-1}\right)B = I \cdot B = B$. Réciproquement si $AX = B$, alors nécessairement $X = A^{-1}B$. Nous verrons bientôt que si la matrice n'est pas inversible, alors soit il n'y a pas de solution, soit une infinité.

5.3. Les matrices élémentaires

Pour calculer l'inverse d'une matrice A, et aussi pour résoudre des systèmes linéaires, nous avons utilisé trois opérations élémentaires sur les lignes qui sont :

1. $L_i \leftarrow \lambda L_i$ avec $\lambda \neq 0$: on peut multiplier une ligne par un réel non nul (ou un élément de $\mathbb{K} \setminus \{0\}$).
2. $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{K}$ (et $j \neq i$) : on peut ajouter à la ligne L_i un multiple d'une autre ligne L_j.
3. $L_i \leftrightarrow L_j$: on peut échanger deux lignes.

Nous allons définir trois matrices élémentaires $E_{L_i \leftarrow \lambda L_i}$, $E_{L_i \leftarrow L_i + \lambda L_j}$, $E_{L_i \leftrightarrow L_j}$ correspondant à ces opérations. Plus précisément, le produit $E \times A$ correspondra à l'opération élémentaire sur A. Voici les définitions accompagnées d'exemples.

1. La matrice $E_{L_i \leftarrow \lambda L_i}$ est la matrice obtenue en multipliant par λ la i-ème ligne de la matrice identité I_n, où λ est un nombre réel non nul.

$$E_{L_2 \leftarrow 5L_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2. La matrice $E_{L_i \leftarrow L_i + \lambda L_j}$ est la matrice obtenue en ajoutant λ fois la j-ème ligne de I_n à la i-ème ligne de I_n.

$$E_{L_2 \leftarrow L_2 - 3L_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3. La matrice $E_{L_i \leftrightarrow L_j}$ est la matrice obtenue en permutant les i-ème et j-ème lignes de I_n.

$$E_{L_2 \leftrightarrow L_4} = E_{L_4 \leftrightarrow L_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Les opérations élémentaires sur les lignes sont réversibles, ce qui entraîne l'inversibilité des matrices élémentaires.

Le résultat de la multiplication d'un matrice élémentaire E par A est la matrice obtenue en effectuant l'opération élémentaire correspondante sur A. Ainsi :

1. La matrice $E_{L_i \leftarrow \lambda L_i} \times A$ est la matrice obtenue en multipliant par λ la i-ème ligne de A.
2. La matrice $E_{L_i \leftarrow L_i + \lambda L_j} \times A$ est la matrice obtenue en ajoutant λ fois la j-ème ligne de A à la i-ème ligne de A.
3. La matrice $E_{L_i \leftrightarrow L_j} \times A$ est la matrice obtenue en permutant les i-ème et j-ème lignes de A.

Exemple 23

1.
$$E_{L_2 \leftarrow \frac{1}{3}L_2} \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{1}{3}y_1 & \frac{1}{3}y_2 & \frac{1}{3}y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

2.
$$E_{L_1 \leftarrow L_1 - 7L_3} \times A = \begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 - 7z_1 & x_2 - 7z_2 & x_3 - 7z_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

3.
$$E_{L_2 \leftrightarrow L_3} \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ z_1 & z_2 & z_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

5.4. Équivalence à une matrice échelonnée

Définition 13

matrices équivalentes par lignes Deux matrices A et B sont dites ***équivalentes par lignes*** si l'une peut être obtenue à partir de l'autre par une suite d'opérations élémentaires sur les lignes. On note $A \sim B$.

Définition 14

Une matrice est ***échelonnée*** si :
- le nombre de zéros commençant une ligne croît strictement ligne par ligne jusqu'à ce qu'il ne reste plus que des zéros.

Elle est ***échelonnée réduite*** si en plus :
- le premier coefficient non nul d'une ligne (non nulle) vaut 1 ;
- et c'est le seul élément non nul de sa colonne.

Exemple d'une matrice échelonnée (à gauche) et échelonnée réduite (à droite) ; les $*$ désignent des coefficients quelconques, les $+$ des coefficients non nuls :

$$\begin{pmatrix} + & * & * & * & * & * & * \\ 0 & 0 & + & * & * & * & * \\ 0 & 0 & 0 & + & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Théorème 4

Étant donnée une matrice $A \in M_{n,p}(\mathbb{K})$, il existe une unique matrice échelonnée réduite U obtenue à partir de A par des opérations élémentaires sur les lignes.

Ce théorème permet donc de se ramener par des opérations élémentaires à des matrices dont la structure est beaucoup plus simple : les matrices échelonnées réduites.

Démonstration

Nous admettons l'unicité.
L'existence se démontre grâce à l'algorithme de Gauss. L'idée générale consiste à utiliser des substitutions de lignes pour placer des zéros là où il faut de façon à créer d'abord une forme échelonnée, puis une forme échelonnée réduite.

Soit A une matrice $n \times p$ quelconque.

Partie A. Passage à une forme échelonnée.
Étape A.1. *Choix du pivot.*
On commence par inspecter la première colonne. Soit elle ne contient que des zéros, auquel cas on passe directement à l'étape A.3, soit elle contient au moins un terme non nul. On choisit alors un tel terme, que l'on appelle le ***pivot***. Si c'est le terme a_{11}, on passe directement à l'étape A.2 ; si c'est un terme a_{i1} avec $i \neq 1$, on échange les lignes 1 et i ($L_1 \leftrightarrow L_i$) et on passe à l'étape A.2.
Au terme de l'étape A.1, soit la matrice A a sa première colonne nulle (à gauche) ou bien on obtient une matrice équivalente dont le premier coefficient a'_{11} est non nul (à droite) :

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1j} & \cdots & a_{1p} \\ 0 & a_{22} & \cdots & a_{2j} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{n2} & \cdots & a_{nj} & \cdots & a_{np} \end{pmatrix} = A \qquad \text{ou bien} \qquad \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1p} \\ a'_{21} & a'_{22} & \cdots & a'_{2j} & \cdots & a'_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{i1} & a'_{i2} & \cdots & a'_{ij} & \cdots & a'_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nj} & \cdots & a'_{np} \end{pmatrix} \sim A.$$

Étape A.2. *Élimination.*
On ne touche plus à la ligne 1, et on se sert du pivot a'_{11} pour éliminer tous les termes a'_{i1} (avec $i \geqslant 2$) situés sous le pivot. Pour cela, il suffit de remplacer la ligne i par elle-même moins $\frac{a'_{i1}}{a'_{11}} \times$ la ligne 1, ceci pour $i = 2, \ldots, n$: $L_2 \leftarrow L_2 - \frac{a'_{21}}{a'_{11}} L_1$, $L_3 \leftarrow L_3 - \frac{a'_{31}}{a'_{11}} L_1$,...
Au terme de l'étape A.2, on a obtenu une matrice de la forme

$$\begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1p} \\ 0 & a''_{22} & \cdots & a''_{2j} & \cdots & a''_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{i2} & \cdots & a''_{ij} & \cdots & a''_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{n2} & \cdots & a''_{nj} & \cdots & a''_{np} \end{pmatrix} \sim A.$$

Étape A.3. *Boucle.*

Au début de l'étape A.3, on a obtenu dans tous les cas de figure une matrice de la forme

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ 0 & a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix} \sim A$$

dont la première colonne est bien celle d'une matrice échelonnée. On va donc conserver cette première colonne. Si $a_{11}^1 \neq 0$, on conserve aussi la première ligne, et l'on repart avec l'étape A.1 en l'appliquant cette fois à la sous-matrice $(n-1) \times (p-1)$ (ci-dessous à gauche : on « oublie » la première ligne et la première colonne de A) ; si $a_{11}^1 = 0$, on repart avec l'étape A.1 en l'appliquant à la sous-matrice $n \times (p-1)$ (à droite, on « oublie » la première colonne) :

$$\begin{pmatrix} a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & & \vdots & & \vdots \\ a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & & \vdots & & \vdots \\ a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix} \qquad \begin{pmatrix} a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & & \vdots & & \vdots \\ a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & & \vdots & & \vdots \\ a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix}$$

Au terme de cette deuxième itération de la boucle, on aura obtenu une matrice de la forme

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ 0 & a_{22}^2 & \cdots & a_{2j}^2 & \cdots & a_{2p}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{ij}^2 & \cdots & a_{ip}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nj}^2 & \cdots & a_{np}^2 \end{pmatrix} \sim A,$$

et ainsi de suite.
Comme chaque itération de la boucle travaille sur une matrice qui a une colonne de moins que la précédente, alors au bout d'au plus $p-1$ itérations de la boucle, on aura obtenu une matrice échelonnée.

Partie B. Passage à une forme échelonnée réduite.
Étape B.1. *Homothéties.*
On repère le premier élément non nul de chaque ligne non nulle, et on multiplie cette ligne par l'inverse de cet élément. Exemple : si le premier élément non nul de la ligne i est $\alpha \neq 0$, alors on effectue $L_i \leftarrow \frac{1}{\alpha} L_i$. Ceci crée une matrice échelonnée avec des 1 en position de pivots.

Étape B.2. *Élimination.*
On élimine les termes situés au-dessus des positions de pivot comme précédemment, en procédant à partir du bas à droite de la matrice. Ceci ne modifie pas la structure échelonnée de la matrice en raison de la disposition des zéros dont on part.

Exemple 24

Soit
$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

A. Passage à une forme échelonnée.
Première itération de la boucle, étape A.1. Le choix du pivot est tout fait, on garde $a^1_{11} = 1$.
Première itération de la boucle, étape A.2. On ne fait rien sur la ligne 2 qui contient déjà un zéro en bonne position et on remplace la ligne 3 par $L_3 \leftarrow L_3 + L_1$. On obtient
$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 2 & 4 & 4 \end{pmatrix}.$$

Deuxième itération de la boucle, étape A.1. Le choix du pivot est tout fait, on garde $a^2_{22} = 2$.
Deuxième itération de la boucle, étape A.2. On remplace la ligne 3 avec l'opération $L_3 \leftarrow L_3 - L_2$. On obtient
$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Cette matrice est échelonnée.

B. Passage à une forme échelonnée réduite.
Étape B.1, homothéties. On multiplie la ligne 2 par $\frac{1}{2}$ et la ligne 3 par $-\frac{1}{2}$ et l'on obtient
$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Étape B.2, première itération. On ne touche plus à la ligne 3 et on remplace la ligne 2 par $L_2 \leftarrow L_2 - 3L_3$ et $L_1 \leftarrow L_1 - 4L_3$. On obtient
$$A \sim \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Étape B.2, deuxième itération. On ne touche plus à la ligne 2 et on remplace la ligne 1 par $L_1 \leftarrow L_1 - 2L_2$. On obtient
$$A \sim \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
qui est bien échelonnée et réduite.

5.5. Matrices élémentaires et inverse d'une matrice

Théorème 5

Soit $A \in M_n(\mathbb{K})$. La matrice A est inversible si et seulement si sa forme échelonnée réduite est la matrice identité I_n.

Démonstration

Notons U la forme échelonnée réduite de A. Et notons E le produit de matrices élémentaires tel que $EA = U$.

$\Longleftarrow$ Si $U = I_n$ alors $EA = I_n$. Ainsi par définition, A est inversible et $A^{-1} = E$.

$\Longrightarrow$ Nous allons montrer que si $U \neq I_n$, alors A n'est pas inversible.

- Supposons $U \neq I_n$. Alors la dernière ligne de U est nulle (sinon il y aurait un pivot sur chaque ligne donc ce serait I_n).
- Cela entraîne que U n'est pas inversible : en effet, pour tout matrice carrée V, la dernière ligne de UV est nulle ; on n'aura donc jamais $UV = I_n$.
- Alors, A n'est pas inversible non plus : en effet, si A était inversible, on aurait $U = EA$ et U serait inversible comme produit de matrices inversibles (E est inversible car c'est un produit de matrices élémentaires qui sont inversibles).

Remarque

Justifions maintenant notre méthode pour calculer A^{-1}.
Nous partons de $(A|I)$ pour arriver par des opérations élémentaires sur les lignes à $(I|B)$. Montrons que $B = A^{-1}$. Faire une opération élémentaire signifie multiplier à gauche par une des matrices élémentaires. Notons E le produit de ces matrices élémentaires. Dire que l'on arrive à la fin du processus à I signifie $EA = I$. Donc $A^{-1} = E$. Comme on fait les mêmes opérations sur la partie droite du tableau, alors on obtient $EI = B$. Donc $B = E$. Conséquence : $B = A^{-1}$.

Corollaire 1

Les assertions suivantes sont équivalentes :

(i) La matrice A est inversible.

(ii) Le système linéaire $AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ a une unique solution $X = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

(iii) Pour tout second membre B, le système linéaire $AX = B$ a une unique solution X.

Démonstration

Nous avons déjà vu $(i) \Longrightarrow (ii)$ et $(i) \Longrightarrow (iii)$.
Nous allons seulement montrer $(ii) \Longrightarrow (i)$. Nous raisonnons par contraposée : nous allons montrer la proposition équivalente non$(i) \Longrightarrow$ non(ii). Si A n'est pas inversible, alors sa forme échelonnée réduite U contient un premier zéro sur sa diagonale, disons à la place ℓ. Alors U à la forme suivante

$$\begin{pmatrix} 1 & 0 & \cdots & c_1 & * & \cdots & * \\ 0 & \ddots & 0 & \vdots & & \cdots & * \\ 0 & 0 & 1 & c_{\ell-1} & & \cdots & * \\ 0 & \cdots & 0 & \color{red}{0} & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \cdots & 0 & \ddots & \vdots \\ 0 & \cdots & & & \cdots & 0 & * \end{pmatrix}. \qquad \text{On note} \qquad X = \begin{pmatrix} -c_1 \\ \vdots \\ -c_{\ell-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Alors X n'est pas le vecteur nul, mais UX est le vecteur nul. Comme $A = E^{-1}U$, alors AX est le vecteur nul. Nous avons donc trouvé un vecteur non nul X tel que $AX = 0$.

Mini-exercices

1. Exprimer les systèmes linéaires suivants sous forme matricielle et les résoudre en inversant la matrice : $\begin{cases} 2x+4y=7 \\ -2x+3y=-14 \end{cases}$, $\begin{cases} x+z=1 \\ -2y+3z=1 \\ x+z=1 \end{cases}$, $\begin{cases} x+t=\alpha \\ x-2y=\beta \\ x+y+t=2 \\ y+t=4 \end{cases}$.
2. Écrire les matrices 4×4 correspondant aux opérations élémentaires : $L_2 \leftarrow \frac{1}{3}L_2$, $L_3 \leftarrow L_3 - \frac{1}{4}L_2$, $L_1 \leftrightarrow L_4$. Sans calculs, écrire leurs inverses. Écrire la matrice 4×4 de l'opération $L_1 \leftarrow L_1 - 2L_3 + 3L_4$.
3. Écrire les matrices suivantes sous forme échelonnée, puis échelonnée réduite : $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 0 \\ -2 & -2 & -3 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & -2 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 & -2 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -2 & 1 & 4 \\ -1 & 2 & -1 & -2 \end{pmatrix}$.

6. Matrices triangulaires, transposition, trace, matrices symétriques

6.1. Matrices triangulaires, matrices diagonales

matrice !triangulaire Soit A une matrice de taille $n \times n$. On dit que A est ***triangulaire inférieure*** si ses éléments au-dessus de la diagonale sont nuls, autrement dit :

$$i < j \implies a_{ij} = 0.$$

Une matrice triangulaire inférieure a la forme suivante :

$$\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

On dit que A est ***triangulaire supérieure*** si ses éléments en-dessous de la diagonale sont nuls, autrement dit :

$$i > j \implies a_{ij} = 0.$$

Une matrice triangulaire supérieure a la forme suivante :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & \dots & a_{2n} \\ \vdots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & a_{nn} \end{pmatrix}$$

Exemple 25

Deux matrices triangulaires inférieures (à gauche), une matrice triangulaire supérieure (à droite) :

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & -2 & 3 \end{pmatrix} \qquad \begin{pmatrix} 5 & 0 \\ 1 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}$$

matrice !diagonalediagonal(e) !matrice Une matrice qui est triangulaire inférieure ***et*** triangulaire supérieure est dite ***diagonale***. Autrement dit : $i \neq j \implies a_{ij} = 0$.

Exemple 26

Exemples de matrices diagonales :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

Exemple 27. Puissances d'une matrice diagonale

Si D est une matrice diagonale, il est très facile de calculer ses puissances D^p (par récurrence sur p) :

$$D = \begin{pmatrix} \alpha_1 & 0 & \dots & \dots & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \alpha_n \end{pmatrix} \implies D^p = \begin{pmatrix} \alpha_1^p & 0 & \dots & \dots & 0 \\ 0 & \alpha_2^p & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1}^p & 0 \\ 0 & \dots & \dots & 0 & \alpha_n^p \end{pmatrix}$$

Théorème 6

Une matrice A de taille $n \times n$, triangulaire, est inversible si et seulement si ses éléments diagonaux sont tous non nuls.

Démonstration

Supposons que A soit triangulaire supérieure.

- Si les éléments de la diagonale sont tous non nuls, alors la matrice A est déjà sous la forme échelonnée. En multipliant chaque ligne i par l'inverse de l'élément diagonal a_{ii}, on obtient des 1 sur la diagonale. De ce fait, la forme échelonnée réduite de A sera la matrice identité. Le théorème 5 permet de conclure que A est inversible.
- Inversement, supposons qu'au moins l'un des éléments diagonaux soit nul et notons $a_{\ell\ell}$ le premier élément nul de la diagonale. En multipliant les lignes 1 à $\ell - 1$ par l'inverse de leur

élément diagonal, on obtient une matrice de la forme

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

Il est alors clair que la colonne numéro ℓ de la forme échelonnée réduite ne contiendra pas de 1 comme pivot. La forme échelonnée réduite de A ne peut donc pas être I_n et par le théorème 5, A n'est pas inversible.

Dans le cas d'une matrice triangulaire inférieure, on utilise la transposition (qui fait l'objet de la section suivante) et on obtient une matrice triangulaire supérieure. On applique alors la démonstration ci-dessus.

6.2. La transposition

Soit A la matrice de taille $n \times p$

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

Définition 15

matrice !transposée On appelle ***matrice transposée*** de A la matrice A^T de taille $p \times n$ définie par : AT@A^T - matrice transposée

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

Autrement dit : le coefficient à la place (i,j) de A^T est a_{ji}. Ou encore la i-ème ligne de A devient la i-ème colonne de A^T (et réciproquement la j-ème colonne de A^T est la j-ème ligne de A).

Notation : La transposée de la matrice A se note aussi souvent ${}^tA$.

Exemple 28

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & -6 \\ -7 & 8 & 9 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 & -7 \\ 2 & 5 & 8 \\ 3 & -6 & 9 \end{pmatrix} \qquad \begin{pmatrix} 0 & 3 \\ 1 & -5 \\ -1 & 2 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -5 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}$$

L'opération de transposition obéit aux règles suivantes :

Théorème 7

1. $(A+B)^T = A^T + B^T$
2. $(\alpha A)^T = \alpha A^T$
3. $(A^T)^T = A$
4. $\boxed{(AB)^T = B^T A^T}$
5. Si A est inversible, alors A^T l'est aussi et on a $(A^T)^{-1} = (A^{-1})^T$.

Notez bien l'inversion : $(AB)^T = B^T A^T$, comme pour $(AB)^{-1} = B^{-1}A^{-1}$.

6.3. La trace

Dans le cas d'une matrice carrée de taille $n \times n$, les éléments a_{11}, $a_{22}, \ldots, a_{nn}$ sont appelés les ***éléments diagonaux***. diagonal(e)!élément Sa ***diagonale principale*** est la diagonale $(a_{11}, a_{22}, \ldots, a_{nn})$.

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Définition 16

La ***trace*** de la matrice A est le nombre obtenu en additionnant les éléments diagonaux de A. Autrement dit,

$$\boxed{\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}.}$$

Exemple 29

- Si $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 5 \end{smallmatrix}\right)$, alors $\operatorname{tr} A = 2 + 5 = 7$.
- Pour $B = \left(\begin{smallmatrix} 1 & 1 & 2 \\ 5 & 2 & 8 \\ 11 & 0 & -10 \end{smallmatrix}\right)$, $\operatorname{tr} B = 1 + 2 - 10 = -7$.

Théorème 8

Soient A et B deux matrices $n \times n$. Alors :

1. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$,
2. $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr} A$ pour tout $\alpha \in \mathbb{K}$,
3. $\operatorname{tr}(A^T) = \operatorname{tr} A$,
4. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Démonstration

1. Pour tout $1 \leqslant i \leqslant n$, le coefficient (i,i) de $A+B$ est $a_{ii}+b_{ii}$. Ainsi, on a bien $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.
2. On a $\mathrm{tr}(\alpha A) = \alpha a_{11} + \cdots + \alpha a_{nn} = \alpha(a_{11} + \cdots + a_{nn}) = \alpha \,\mathrm{tr}\, A$.
3. Étant donné que la transposition ne change pas les éléments diagonaux, la trace de A est égale à la trace de A^T.
4. Notons c_{ij} les coefficients de AB. Alors par définition
$$c_{ii} = a_{i1}b_{1i} + a_{i2}b_{2i} + \cdots + a_{in}b_{ni}.$$
Ainsi,
$$\begin{array}{rcllll} \mathrm{tr}(AB) & = & a_{11}b_{11} & +a_{12}b_{21} & +\cdots & +a_{1n}b_{n1} \\ & & +a_{21}b_{12} & +a_{22}b_{22} & +\cdots & +a_{2n}b_{n2} \\ & & \vdots & & & \\ & & +a_{n1}b_{1n} & +a_{n2}b_{2n} & +\cdots & +a_{nn}b_{nn}. \end{array}$$
On peut réarranger les termes pour obtenir
$$\begin{array}{rcllll} \mathrm{tr}(AB) & = & a_{11}b_{11} & +a_{21}b_{12} & +\cdots & +a_{n1}b_{1n} \\ & & +a_{12}b_{21} & +a_{22}b_{22} & +\cdots & +a_{n2}b_{2n} \\ & & \vdots & & & \\ & & +a_{1n}b_{n1} & +a_{2n}b_{n2} & +\cdots & +a_{nn}b_{nn}. \end{array}$$
En utilisant la commutativité de la multiplication dans $\mathbb{K}$, la première ligne devient
$$b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1n}a_{n1}$$
qui vaut le coefficient $(1,1)$ de BA. On note d_{ij} les coefficients de BA. En faisant de même avec les autres lignes, on voit finalement que
$$\mathrm{tr}(AB) = d_{11} + \cdots + d_{nn} = \mathrm{tr}(BA).$$

6.4. Matrices symétriques

Définition 17

matrice !symétrique Une matrice A de taille $n \times n$ est ***symétrique*** si elle est égale à sa transposée, c'est-à-dire si
$$A = A^T,$$
ou encore si $a_{ij} = a_{ji}$ pour tout $i,j = 1,\ldots,n$. Les coefficients sont donc symétriques par rapport à la diagonale.

Exemple 30

Les matrices suivantes sont symétriques :
$$\begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 5 \\ 0 & 2 & -1 \\ 5 & -1 & 0 \end{pmatrix}$$

Exemple 31

Pour une matrice B quelconque, les matrices $B \cdot B^T$ et $B^T \cdot B$ sont symétriques.
Preuve : $(BB^T)^T = (B^T)^T B^T = BB^T$. Idem pour $B^T B$.

6.5. Matrices antisymétriques

Définition 18

matrice !antisymétrique Une matrice A de taille $n \times n$ est ***antisymétrique*** si

$$A^T = -A,$$

c'est-à-dire si $a_{ij} = -a_{ji}$ pour tout $i, j = 1, \ldots, n$.

Exemple 32

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 4 & 2 \\ -4 & 0 & -5 \\ -2 & 5 & 0 \end{pmatrix}$$

Remarquons que les éléments diagonaux d'une matrice antisymétrique sont toujours tous nuls.

Exemple 33

Toute matrice est la somme d'une matrice symétrique et d'une matrice antisymétrique.

Preuve : Soit A une matrice. Définissons $B = \frac{1}{2}(A + A^T)$ et $C = \frac{1}{2}(A - A^T)$. Alors d'une part $A = B + C$; d'autre part B est symétrique, car $B^T = \frac{1}{2}(A^T + (A^T)^T) = \frac{1}{2}(A^T + A) = B$; et enfin C est antisymétrique, car $C^T = \frac{1}{2}(A^T - (A^T)^T) = -C$.
Exemple :

$$\text{Pour} \quad A = \begin{pmatrix} 2 & 10 \\ 8 & -3 \end{pmatrix} \qquad \text{alors} \qquad A \;=\; \underbrace{\begin{pmatrix} 2 & 9 \\ 9 & -3 \end{pmatrix}}_{\text{symétrique}} \;+\; \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\text{antisymétrique}}.$$

Mini-exercices

1. Montrer que la somme de deux matrices triangulaires supérieures reste triangulaire supérieure. Montrer que c'est aussi valable pour le produit.
2. Montrer que si A est triangulaire supérieure, alors A^T est triangulaire inférieure. Et si A est diagonale ?
3. Soit $A = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Calculer $A^T \cdot A$, puis $A \cdot A^T$.
4. Soit $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Calculer $\mathrm{tr}(A \cdot A^T)$.
5. Soit A une matrice de taille 2×2 inversible. Montrer que si A est symétrique, alors A^{-1} aussi. Et si A est antisymétrique ?
6. Montrer que la décomposition d'une matrice sous la forme « symétrique + antisymé-

trique » est unique.

Auteurs

- D'après un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- et un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties de cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- mixés et révisés par Arnaud Bodin, relu par Vianney Combet.

Exo7

3 L'espace vectoriel $\mathbb{R}^n$

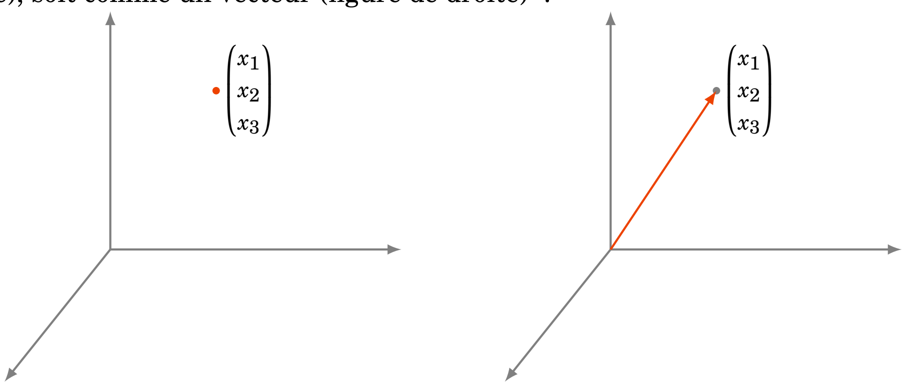
Ce chapitre est consacré à l'ensemble $\mathbb{R}^n$ vu comme espace vectoriel. Il peut être vu de plusieurs façons :

- un cours minimal sur les espaces vectoriels pour ceux qui n'auraient besoin que de $\mathbb{R}^n$,
- une introduction avant d'attaquer le cours détaillé sur les espaces vectoriels,
- une source d'exemples à lire en parallèle du cours sur les espaces vectoriels.

1. Vecteurs de $\mathbb{R}^n$

1.1. Opérations sur les vecteurs

- L'ensemble des nombres réels $\mathbb{R}$ est souvent représenté par une droite. C'est un espace de dimension 1.
- Le plan est formé des couples $\binom{x_1}{x_2}$ de nombres réels. Il est noté $\mathbb{R}^2$. C'est un espace à deux dimensions.
- L'espace de dimension 3 est constitué des triplets de nombres réels $\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$. Il est noté $\mathbb{R}^3$. Le symbole $\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$ a deux interprétations géométriques : soit comme un point de l'espace (figure de gauche), soit comme un vecteur (figure de droite) :



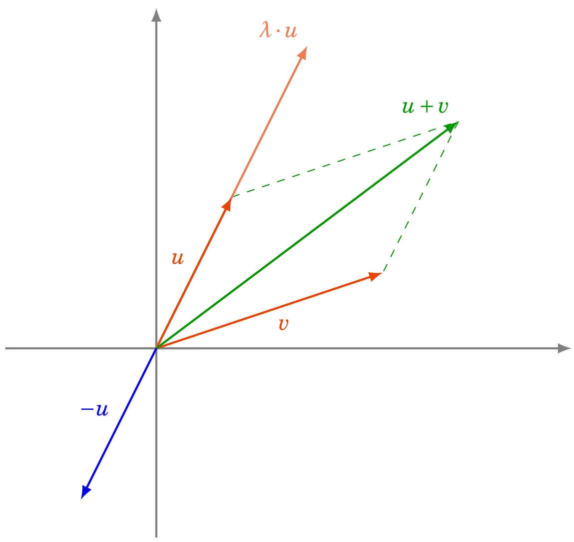
On généralise ces notions en considérant des espaces de dimension n pour tout entier positif $n = 1, 2, 3, 4, \dots$ Les éléments de l'espace de dimension n sont les n-uples $\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ de nombres réels. L'espace de dimension n est noté $\mathbb{R}^n$. Comme en dimensions 2 et 3, le n-uple $\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ dénote aussi bien un point qu'un vecteur de l'espace de dimension n.

Soient $u = \begin{pmatrix}u_1\\u_2\\\vdots\\u_n\end{pmatrix}$ et $v = \begin{pmatrix}v_1\\v_2\\\vdots\\v_n\end{pmatrix}$ deux vecteurs de $\mathbb{R}^n$.

Définition 19

- ***Somme de deux vecteurs.*** Leur somme est par définition le vecteur $u+v=\begin{pmatrix} u_1+v_1 \\ \vdots \\ u_n+v_n \end{pmatrix}$.
- ***Produit d'un vecteur par un scalaire.*** Soit $\lambda \in \mathbb{R}$ (appelé un ***scalaire***) : $\lambda \cdot u = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix}$.
- Le ***vecteur nul*** de $\mathbb{R}^n$ est le vecteur $0=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.
- L'***opposé*** du vecteur $u=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ est le vecteur $-u=\begin{pmatrix} -u_1 \\ \vdots \\ -u_n \end{pmatrix}$.

Voici des vecteurs dans $\mathbb{R}^2$ (ici $\lambda = 2$) :



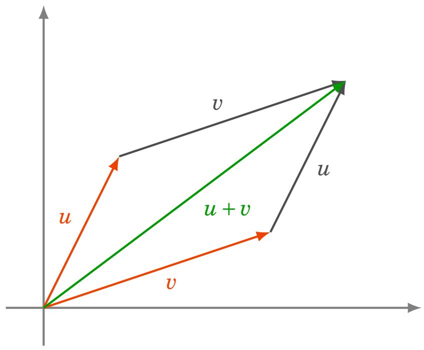
Dans un premier temps, vous pouvez noter $\vec{u}, \vec{v}, \vec{0}$ au lieu de u, v, 0. Mais il faudra s'habituer rapidement à la notation sans flèche. De même, si λ est un scalaire et u un vecteur, on notera souvent λu au lieu de $\lambda \cdot u$.

Théorème 9

Soient $u=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, $v=\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ et $w=\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ des vecteurs de $\mathbb{R}^n$ et $\lambda, \mu \in \mathbb{R}$. Alors :

1. $u+v=v+u$
2. $u+(v+w)=(u+v)+w$
3. $u+0=0+u=u$
4. $u+(-u)=0$
5. $1 \cdot u = u$
6. $\lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u$
7. $\lambda \cdot (u+v) = \lambda \cdot u + \lambda \cdot v$

8. $(\lambda+\mu)\cdot u = \lambda\cdot u+\mu\cdot u$



Chacune de ces propriétés découle directement de la définition de la somme et de la multiplication par un scalaire. Ces huit propriétés font de $\mathbb{R}^n$ un ***espace vectoriel***. Dans le cadre général, ce sont ces huit propriétés qui définissent ce qu'est un espace vectoriel.

1.2. Représentation des vecteurs de $\mathbb{R}^n$

Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ un vecteur de $\mathbb{R}^n$. On l'appelle ***vecteur colonne*** et on considère naturellement u comme une matrice de taille $n\times 1$. Parfois, on rencontre aussi des ***vecteurs lignes*** : on peut voir le vecteur u comme une matrice $1\times n$, de la forme $(u_1,\ldots,u_n)$. En fait, le vecteur ligne correspondant à u est le transposé u^T du vecteur colonne u.
Les opérations de somme et de produit par un scalaire définies ci-dessus pour les vecteurs coïncident parfaitement avec les opérations définies sur les matrices :

$$u+v = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1+v_1 \\ \vdots \\ u_n+v_n \end{pmatrix} \qquad \text{et} \qquad \lambda u = \lambda \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix}.$$

1.3. Produit scalaire

Soient $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ deux vecteurs de $\mathbb{R}^n$. On définit leur ***produit scalaire*** par

$$\langle u \mid v \rangle = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

C'est un scalaire (un nombre réel). Remarquons que cette définition généralise la notion de produit scalaire dans le plan $\mathbb{R}^2$ et dans l'espace $\mathbb{R}^3$.
Une autre écriture :

$$\langle u \mid v \rangle = u^T \times v = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Soient $A = (a_{ij})$ une matrice de taille $n\times p$, et $B = (b_{ij})$ une matrice de taille $p\times q$. Nous savons que l'on peut former le produit matriciel AB. On obtient une matrice de taille $n\times q$. L'élément d'indice ij de la matrice AB est

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}.$$

Remarquons que ceci est aussi le produit matriciel :

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{ip} \end{pmatrix} \times \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}.$$

Autrement dit, c'est le produit scalaire du i-ème vecteur ligne de A avec le j-ème vecteur colonne de B. Notons $\ell_1, \ldots, \ell_n$ les vecteurs lignes formant la matrice A, et $c_1, \ldots, c_q$ les vecteurs colonnes formant la matrice B. On a alors

$$AB = \begin{pmatrix} \langle \ell_1 \mid c_1 \rangle & \langle \ell_1 \mid c_2 \rangle & \cdots & \langle \ell_1 \mid c_q \rangle \\ \langle \ell_2 \mid c_1 \rangle & \langle \ell_2 \mid c_2 \rangle & \cdots & \langle \ell_2 \mid c_q \rangle \\ \vdots & \vdots & & \vdots \\ \langle \ell_n \mid c_1 \rangle & \langle \ell_n \mid c_2 \rangle & \cdots & \langle \ell_n \mid c_q \rangle \end{pmatrix}.$$

Mini-exercices

1. Faire un dessin pour chacune des 8 propriétés qui font de $\mathbb{R}^2$ un espace vectoriel.
2. Faire la même chose pour $\mathbb{R}^3$.
3. Montrer que le produit scalaire vérifie $\langle u \mid v \rangle = \langle v \mid u \rangle$, $\langle u + v \mid w \rangle = \langle u \mid w \rangle + \langle v \mid w \rangle$, $\langle \lambda u \mid v \rangle = \lambda \langle u \mid v \rangle$ pour tout $u, v, w \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$.
4. Soit $u \in \mathbb{R}^n$. Montrer que $\langle u \mid u \rangle \geqslant 0$. Montrer $\langle u \mid u \rangle = 0$ si et seulement si u est le vecteur nul.

2. Exemples d'applications linéaires

Soient

$$f_1 : \mathbb{R}^p \longrightarrow \mathbb{R} \qquad f_2 : \mathbb{R}^p \longrightarrow \mathbb{R} \qquad \cdots \qquad f_n : \mathbb{R}^p \longrightarrow \mathbb{R}$$

n fonctions de p variables réelles à valeurs réelles ; chaque f_i est une fonction :

$$f_i : \mathbb{R}^p \longrightarrow \mathbb{R}, \qquad (x_1, x_2, \ldots, x_p) \mapsto f_i(x_1, \ldots, x_p)$$

On construit une application

$$f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$$

définie par

$$f(x_1, \ldots, x_p) = \big(f_1(x_1, \ldots, x_p), \ldots, f_n(x_1, \ldots, x_p)\big).$$

2.1. Applications linéaires

Définition 20

Une application $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ définie par $f(x_1, \ldots, x_p) = (y_1, \ldots, y_n)$ est dite une ***application***

linéaire si

$$\left\{\begin{array}{ccccccccc} y_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1p}x_p \\ y_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2p}x_p \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ y_n & = & a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{np}x_p\,. \end{array}\right.$$

En notation matricielle, on a

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix},$$

ou encore, si on note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ et $A \in M_{n,p}(\mathbb{R})$ la matrice (a_{ij}),

$$\boxed{f(X) = AX.}$$

Autrement dit, une application linéaire $\mathbb{R}^p \to \mathbb{R}^n$ peut s'écrire $X \mapsto AX$. La matrice $A \in M_{n,p}(\mathbb{R})$ est appelée la ***matrice de l'application linéaire*** f.

Remarque

- On a toujours $f(0,\ldots,0) = (0,\ldots,0)$. Si on note 0 pour le vecteur nul dans $\mathbb{R}^p$ et aussi dans $\mathbb{R}^n$, alors une application linéaire vérifie toujours $f(0) = 0$.
- Le nom complet de la matrice A est : la matrice de l'application linéaire f de la base canonique de $\mathbb{R}^p$ vers la base canonique de $\mathbb{R}^n$!

Exemple 34

La fonction $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ définie par

$$\left\{\begin{array}{ccl} y_1 & = & -2x_1 + 5x_2 + 2x_3 - 7x_4 \\ y_2 & = & 4x_1 + 2x_2 - 3x_3 + 3x_4 \\ y_3 & = & 7x_1 - 3x_2 + 9x_3 \end{array}\right.$$

s'exprime sous forme matricielle comme suit :

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & 2 & -7 \\ 4 & 2 & -3 & 3 \\ 7 & -3 & 9 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Exemple 35

- Pour l'application linéaire identité $\mathbb{R}^n \to \mathbb{R}^n$, $(x_1,\ldots,x_n) \mapsto (x_1,\ldots,x_n)$, sa matrice associée est l'identité I_n (car $I_n X = X$).
- Pour l'application linéaire nulle $\mathbb{R}^p \to \mathbb{R}^n$, $(x_1,\ldots,x_p) \mapsto (0,\ldots,0)$, sa matrice associée est la matrice nulle $0_{n,p}$ (car $0_{n,p}X = 0$).

2.2. Exemples d'applications linéaires

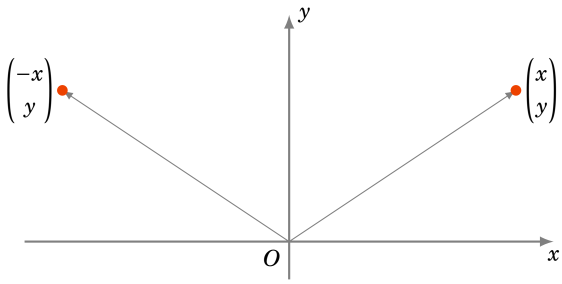
Réflexion par rapport à l'axe (Oy)

La fonction

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}$$

est la réflexion par rapport à l'axe des ordonnées (Oy), et sa matrice est

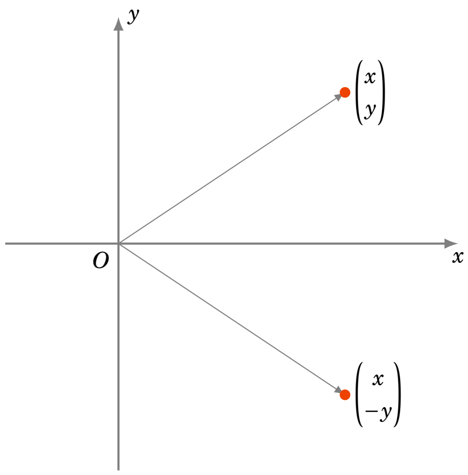
$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{car} \qquad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$



Réflexion par rapport à l'axe (Ox)

La réflexion par rapport à l'axe des abscisses (Ox) est donnée par la matrice

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



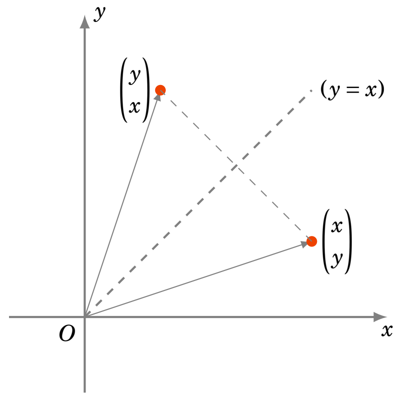
Réflexion par rapport à la droite $(y = x)$

La réflexion par rapport à la droite $(y = x)$ est donnée par

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$$

et sa matrice est

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



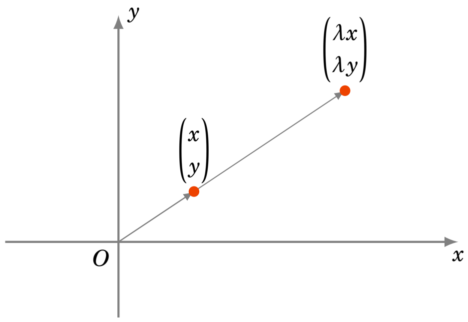
Homothéties

L'homothétie de rapport λ centrée à l'origine est :

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}.$$

On peut donc écrire $f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$. Alors la matrice de l'homothétie est :

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$



Remarque

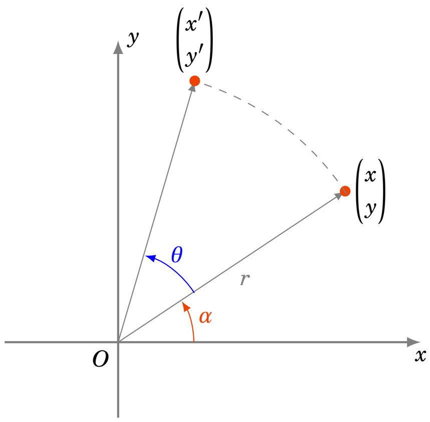
La translation de vecteur $\left(\begin{smallmatrix} u_0 \\ v_0 \end{smallmatrix}\right)$ est l'application

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} x + u_0 \\ y + v_0 \end{pmatrix}.$$

Si c'est une translation de vecteur non nul, c'est-à-dire $\left(\begin{smallmatrix} u_0 \\ v_0 \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, alors ***ce n'est pas*** une application linéaire, car $f\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$.

Rotations

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la rotation d'angle θ, centrée à l'origine.



Si le vecteur $\binom{x}{y}$ fait un angle α avec l'horizontale et que le point $\binom{x}{y}$ est à une distance r de l'origine, alors

$$\begin{cases} x = r\cos\alpha \\ y = r\sin\alpha \end{cases}.$$

Si $\binom{x'}{y'}$ dénote l'image de $\binom{x}{y}$ par la rotation d'angle θ, on obtient :

$$\begin{cases} x' = r\cos(\alpha+\theta) \\ y' = r\sin(\alpha+\theta) \end{cases} \quad \text{donc} \quad \begin{cases} x' = r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ y' = r\cos\alpha\sin\theta + r\sin\alpha\cos\theta \end{cases}$$

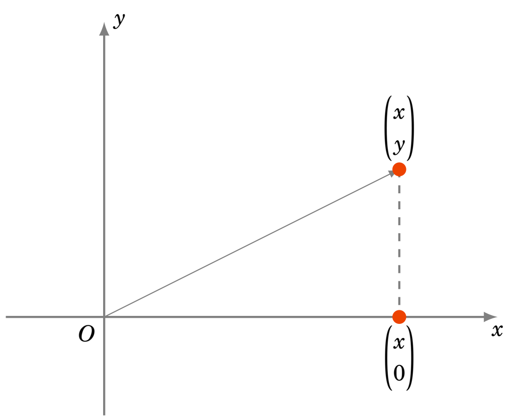
(où l'on a appliqué les formules de trigonométrie pour $\cos(\alpha+\theta)$ et $\sin(\alpha+\theta)$).
On aboutit à

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases} \quad \text{donc} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Autrement dit, la rotation d'angle θ est donnée par la matrice

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Projections orthogonales



L'application

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

est la projection orthogonale sur l'axe (Ox). C'est une application linéaire donnée par la matrice

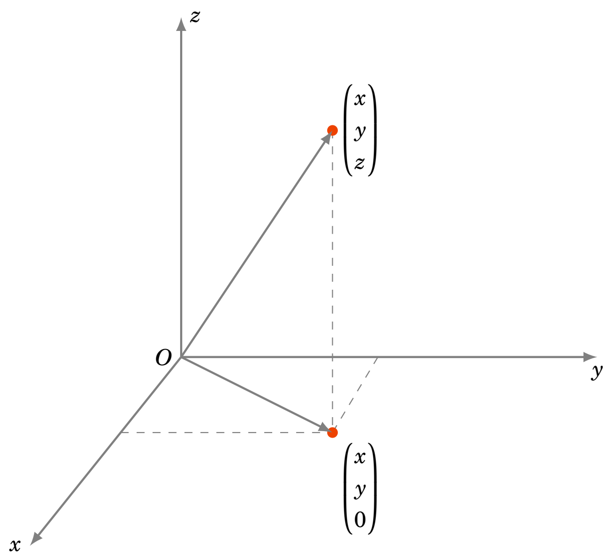
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

L'application linéaire

$$f:\mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \begin{pmatrix} x\\ y\\ z\end{pmatrix} \mapsto \begin{pmatrix} x\\ y\\ 0\end{pmatrix}$$

est la projection orthogonale sur le plan (Oxy) et sa matrice est

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



De même, la projection orthogonale sur le plan (Oxz) est donnée par la matrice de gauche ; la projection orthogonale sur le plan (Oyz) par la matrice de droite :

$$\begin{pmatrix} 1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Réflexions dans l'espace

L'application

$$f:\mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \begin{pmatrix} x\\ y\\ z\end{pmatrix} \mapsto \begin{pmatrix} x\\ y\\ -z\end{pmatrix}$$

est la réflexion par rapport au plan (Oxy). C'est une application linéaire et sa matrice est

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

De même, les réflexions par rapport aux plans (Oxz) (à gauche) et (Oyz) (à droite) sont données par les matrices :

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

Mini-exercices

1. Soit $A = \left(\begin{smallmatrix}1&2\\1&3\end{smallmatrix}\right)$ et soit f l'application linéaire associée. Calculer et dessiner l'image par f de $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, puis $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ et plus généralement de $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)$. Dessiner l'image par f du carré de sommets $\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. Dessiner l'image par f du cercle inscrit dans ce carré.
2. Soit $A = \left(\begin{smallmatrix}1&2&-1\\0&1&0\\2&1&1\end{smallmatrix}\right)$ et soit f l'application linéaire associée. Calculer l'image par f de $\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right)$ et plus généralement de $\left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right)$.
3. Écrire la matrice de la rotation du plan d'angle $\frac{\pi}{4}$ centrée à l'origine. Idem dans l'espace avec la rotation d'angle $\frac{\pi}{4}$ d'axe (Ox).
4. Écrire la matrice de la réflexion du plan par rapport à la droite $(y = -x)$. Idem dans l'espace avec la réflexion par rapport au plan d'équation $(y = -x)$.
5. Écrire la matrice de la projection orthogonale de l'espace sur l'axe (Oy).

3. Propriétés des applications linéaires

3.1. Composition d'applications linéaires et produit de matrices

Soient

$$f : \mathbb{R}^p \longrightarrow \mathbb{R}^n \qquad \text{et} \qquad g : \mathbb{R}^q \longrightarrow \mathbb{R}^p$$

deux applications linéaires. Considérons leur composition :

$$\mathbb{R}^q \xrightarrow{g} \mathbb{R}^p \xrightarrow{f} \mathbb{R}^n \qquad\qquad f \circ g : \mathbb{R}^q \longrightarrow \mathbb{R}^n.$$

L'application $f \circ g$ est une application linéaire. Notons :

- $A = \mathrm{Mat}(f) \in M_{n,p}(\mathbb{R})$ la matrice associée à f,
- $B = \mathrm{Mat}(g) \in M_{p,q}(\mathbb{R})$ la matrice associée à g,
- $C = \mathrm{Mat}(f \circ g) \in M_{n,q}(\mathbb{R})$ la matrice associée à $f \circ g$.

On a pour un vecteur $X \in \mathbb{R}^q$:

$$(f \circ g)(X) = f\big(g(X)\big) = f\big(BX\big) = A(BX) = (AB)X.$$

Donc la matrice associée à $f \circ g$ est $C = AB$.
Autrement dit, la matrice associée à la composition de deux applications linéaires est égale au produit de leurs matrices :

$$\boxed{\mathrm{Mat}(f \circ g) = \mathrm{Mat}(f) \times \mathrm{Mat}(g)}$$

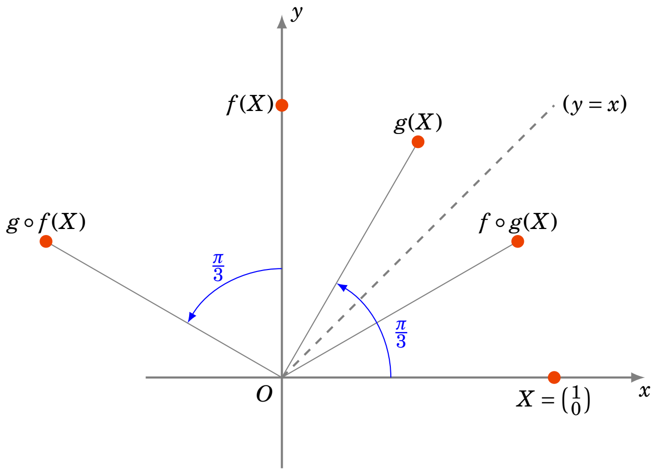
En fait le produit de matrices, qui au premier abord peut sembler bizarre et artificiel, est défini exactement pour vérifier cette relation.

Exemple 36

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la réflexion par rapport à la droite $(y = x)$ et soit $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la rotation d'angle $\theta = \frac{\pi}{3}$ (centrée à l'origine).
Les matrices sont

$$A = \mathrm{Mat}(f) = \begin{pmatrix}0&1\\1&0\end{pmatrix} \qquad \text{et} \qquad B = \mathrm{Mat}(g) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} = \begin{pmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}.$$

Voici pour $X = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ les images $f(X)$, $g(X)$, $f \circ g(X)$, $g \circ f(X)$:



Alors

$$C=\mathrm{Mat}(f\circ g)=\mathrm{Mat}(f)\times\mathrm{Mat}(g)=\begin{pmatrix}0&1\\1&0\end{pmatrix}\times\begin{pmatrix}\frac12&-\frac{\sqrt3}{2}\\\frac{\sqrt3}{2}&\frac12\end{pmatrix}=\begin{pmatrix}\frac{\sqrt3}{2}&\frac12\\\frac12&-\frac{\sqrt3}{2}\end{pmatrix}.$$

Notons que si l'on considère la composition $g\circ f$ alors

$$D=\mathrm{Mat}(g\circ f)=\mathrm{Mat}(g)\times\mathrm{Mat}(f)=\begin{pmatrix}\frac12&-\frac{\sqrt3}{2}\\\frac{\sqrt3}{2}&\frac12\end{pmatrix}\times\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}-\frac{\sqrt3}{2}&\frac12\\\frac12&\frac{\sqrt3}{2}\end{pmatrix}.$$

Les matrices $C=AB$ et $D=BA$ sont distinctes, ce qui montre que la composition d'applications linéaires, comme la multiplication des matrices, n'est pas commutative en général.

3.2. Application linéaire bijective et matrice inversible

Théorème 10

Une application linéaire $f:\mathbb{R}^n\to\mathbb{R}^n$ est bijective si et seulement si sa matrice associée $A=\mathrm{Mat}(f)\in M_n(\mathbb{R})$ est inversible.

L'application f est définie par $f(X)=AX$. Donc si f est bijective, alors d'une part $f(X)=Y \iff X=f^{-1}(Y)$, mais d'autre part $AX=Y \iff X=A^{-1}Y$. Conséquence : la matrice de f^{-1} est A^{-1}.

Corollaire 2

Si f est bijective, alors

$$\boxed{\mathrm{Mat}(f^{-1})=\big(\mathrm{Mat}(f)\big)^{-1}.}$$

Exemple 37

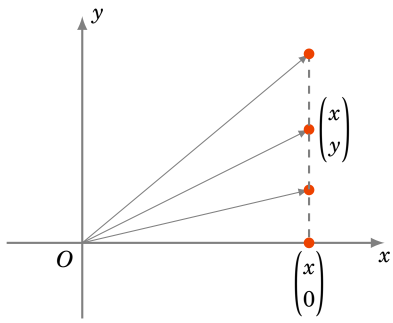
Soit $f:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ la rotation d'angle θ. Alors $f^{-1}:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ est la rotation d'angle $-\theta$.
On a

$$\mathrm{Mat}(f)=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix},$$

$$\mathrm{Mat}(f^{-1}) = \big(\mathrm{Mat}(f)\big)^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix}.$$

Exemple 38

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la projection sur l'axe (Ox). Alors f n'est pas injective. En effet, pour x fixé et tout $y \in \mathbb{R}$, $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$. L'application f n'est pas non plus surjective : ceci se vérifie aisément car aucun point en-dehors de l'axe (Ox) n'est dans l'image de f.



La matrice de f est $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$; elle n'est pas inversible.

La preuve du théorème 10 est une conséquence directe du théorème suivant, vu dans le chapitre sur les matrices :

Théorème 11

Les assertions suivantes sont équivalentes :

(i) La matrice A est inversible.

(ii) Le système linéaire $AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ a une unique solution $X = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

(iii) Pour tout second membre Y, le système linéaire $AX = Y$ a une unique solution X.

Voici donc la preuve du théorème 10.

Démonstration

- Si A est inversible, alors pour tout vecteur Y le système $AX = Y$ a une unique solution X, autrement dit pour tout Y, il existe un unique X tel que $f(X) = AX = Y$. f est donc bijective.
- Si A n'est pas inversible, alors il existe un vecteur X non nul tel que $AX = 0$. En conséquence on a $X \neq 0$ mais $f(X) = f(0) = 0$. f n'est pas injective donc pas bijective.

3.3. Caractérisation des applications linéaires

Théorème 12

Une application $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ est linéaire si et seulement si pour tous les vecteurs u, v de $\mathbb{R}^p$ et pour tout scalaire $\lambda \in \mathbb{R}$, on a

(i) $f(u+v) = f(u) + f(v)$,

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

Dans le cadre général des espaces vectoriels, ce sont ces deux propriétés (i) et (ii) qui définissent une application linéaire.

Définition 21

Les vecteurs

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \cdots \qquad e_p = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

sont appelés les ***vecteurs de la base canonique*** de $\mathbb{R}^p$.

La démonstration du théorème impliquera :

Corollaire 3

Soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ une application linéaire, et soient $e_1, \ldots, e_p$ les vecteurs de base canonique de $\mathbb{R}^p$. Alors la matrice de f (dans les bases canoniques de $\mathbb{R}^p$ vers $\mathbb{R}^n$) est donnée par

$$\mathrm{Mat}(f) = \begin{pmatrix} f(e_1) & f(e_2) & \cdots & f(e_p) \end{pmatrix} ;$$

autrement dit les vecteurs colonnes de $\mathrm{Mat}(f)$ sont les images par f des vecteurs de la base canonique $(e_1, \ldots, e_p)$.

Exemple 39

Considérons l'application linéaire $f : \mathbb{R}^3 \to \mathbb{R}^4$ définie par

$$\begin{cases} y_1 &= 2x_1 & +x_2 & -x_3 \\ y_2 &= -x_1 & -4x_2 & \\ y_3 &= 5x_1 & +x_2 & +x_3 \\ y_4 &= & 3x_2 & +2x_3 \,. \end{cases}$$

Calculons les images des vecteurs de la base canonique $\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right)$:

$$f\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 5 \\ 0 \end{pmatrix} \qquad f\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 1 \\ 3 \end{pmatrix} \qquad f\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

Donc la matrice de f est :

$$\operatorname{Mat}(f) = \begin{pmatrix} 2 & 1 & -1 \\ -1 & -4 & 0 \\ 5 & 1 & 1 \\ 0 & 3 & 2 \end{pmatrix}.$$

Exemple 40

Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ la réflexion par rapport à la droite $(y = x)$ et soit g la rotation du plan d'angle $\frac{\pi}{6}$ centrée à l'origine. Calculons la matrice de l'application $f \circ g$. La base canonique de $\mathbb{R}^2$ est formée des vecteurs $\binom{1}{0}$ et $\binom{0}{1}$.

$$f \circ g \begin{pmatrix} 1 \\ 0 \end{pmatrix} = f \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \qquad f \circ g \begin{pmatrix} 0 \\ 1 \end{pmatrix} = f \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}$$

Donc la matrice de $f \circ g$ est :

$$\operatorname{Mat}(f) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

Voici la preuve du théorème 12.

Démonstration

Supposons $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ linéaire, et soit A sa matrice. On a $f(u+v) = A(u+v) = Au + Av = f(u) + f(v)$ et $f(\lambda u) = A(\lambda u) = \lambda A u = \lambda f(u)$.

Réciproquement, soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ une application qui vérifie (i) et (ii). Nous devons construire une matrice A telle que $f(u) = Au$. Notons d'abord que (i) implique que $f(v_1 + v_2 + \cdots + v_r) = f(v_1) + f(v_2) + \cdots + f(v_r)$. Notons $(e_1, \ldots, e_p)$ les vecteurs de la base canonique de $\mathbb{R}^p$.

Soit A la matrice $n \times p$ dont les colonnes sont

$$f(e_1), f(e_2), \ldots, f(e_p).$$

$$\text{Pour} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{R}^p, \qquad \text{alors} \qquad X = x_1 e_1 + x_2 e_2 + \cdots + x_p e_p$$

et donc

$$\begin{aligned} AX &= A(x_1 e_1 + x_2 e_2 + \cdots + x_p e_p) \\ &= A x_1 e_1 + A x_2 e_2 + \cdots + A x_p e_p \\ &= x_1 A e_1 + x_2 A e_2 + \cdots + x_p A e_p \\ &= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_p f(e_p) \\ &= f(x_1 e_1) + f(x_2 e_2) + \cdots + f(x_p e_p) \\ &= f(x_1 e_1 + x_2 e_2 + \cdots + x_p e_p) = f(X). \end{aligned}$$

On a alors $f(X) = AX$, et f est bien une application linéaire (de matrice A).

Mini-exercices

1. Soit f la réflexion du plan par rapport à l'axe (Ox) et soit g la rotation d'angle $\frac{2\pi}{3}$ centrée à l'origine. Calculer la matrice de $f \circ g$ de deux façons différentes (produit de matrices et image de la base canonique). Cette matrice est-elle inversible ? Si oui, calculer l'inverse. Interprétation géométrique. Même question avec $g \circ f$.
2. Soit f la projection orthogonale de l'espace sur le plan (Oxz) et soit g la rotation d'angle $\frac{\pi}{2}$ d'axe (Oy). Calculer la matrice de $f \circ g$ de deux façons différentes (produit de matrices et image de la base canonique). Cette matrice est-elle inversible ? Si oui, calculer l'inverse. Interprétation géométrique. Même question avec $g \circ f$.

Auteurs

- D'après un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- révisé et reformaté par Arnaud Bodin, relu par Vianney Combet.

Exo7

4 Espaces vectoriels

Vidéo ■ partie 1. Espace vectoriel (début)
Vidéo ■ partie 2. Espace vectoriel (fin)
Vidéo ■ partie 3. Sous-espace vectoriel (début)
Vidéo ■ partie 4. Sous-espace vectoriel (milieu)
Vidéo ■ partie 5. Sous-espace vectoriel (fin)
Vidéo ■ partie 6. Application linéaire (début)
Vidéo ■ partie 7. Application linéaire (milieu)
Vidéo ■ partie 8. Application linéaire (fin)

La notion d'espace vectoriel est une structure fondamentale des mathématiques modernes. Il s'agit de dégager les propriétés communes que partagent des ensembles pourtant très différents. Par exemple, on peut additionner deux vecteurs du plan, et aussi multiplier un vecteur par un réel (pour l'agrandir ou le rétrécir). Mais on peut aussi additionner deux fonctions, ou multiplier une fonction par un réel. Même chose avec les polynômes, les matrices,... Le but est d'obtenir des théorèmes généraux qui s'appliqueront aussi bien aux vecteurs du plan, de l'espace, aux espaces de fonctions, aux polynômes, aux matrices,... La contrepartie de cette grande généralité de situations est que la notion d'espace vectoriel est difficile à appréhender et vous demandera une quantité conséquente de travail ! Il est bon d'avoir d'abord étudié le chapitre « L'espace vectoriel $\mathbb{R}^n$ ».

1. Espace vectoriel (début)

Dans ce chapitre, $\mathbb{K}$ désigne un corps. Dans la plupart des exemples, ce sera le corps des réels $\mathbb{R}$.

1.1. Définition d'un espace vectoriel

Un espace vectoriel est un ensemble formé de vecteurs, de sorte que l'on puisse additionner (et soustraire) deux vecteurs u,v pour en former un troisième $u+v$ (ou $u-v$) et aussi afin que l'on puisse multiplier chaque vecteur u d'un facteur λ pour obtenir un vecteur $\lambda \cdot u$. Voici la définition formelle :

Définition 22

Un ***$\mathbb{K}$-espace vectoriel*** est un ensemble non vide E muni :

– d'une loi de composition interne, c'est-à-dire d'une application de $E \times E$ dans E :

$$\begin{array}{ccc} E\times E & \to & E \\ (u,v) & \mapsto & u+v \end{array}$$

– d'une loi de composition externe, c'est-à-dire d'une application de $\mathbb{K} \times E$ dans E :

$$\begin{array}{ccc} \mathbb{K}\times E & \to & E \\ (\lambda,u) & \mapsto & \lambda\cdot u \end{array}$$

qui vérifient les propriétés suivantes :

1. $u+v=v+u$ (pour tous $u,v\in E$)
2. $u+(v+w)=(u+v)+w$ (pour tous $u,v,w\in E$)
3. Il existe un ***élément neutre*** $0_E\in E$ tel que $u+0_E=u$ (pour tout $u\in E$)
4. Tout $u\in E$ admet un ***symétrique*** u' tel que $u+u'=0_E$. Cet élément u' est noté $-u$.
5. $1\cdot u=u$ (pour tout $u\in E$)
6. $\lambda\cdot(\mu\cdot u)=(\lambda\mu)\cdot u$ (pour tous $\lambda,\mu\in\mathbb{K}$, $u\in E$)
7. $\lambda\cdot(u+v)=\lambda\cdot u+\lambda\cdot v$ (pour tous $\lambda\in\mathbb{K}$, $u,v\in E$)
8. $(\lambda+\mu)\cdot u=\lambda\cdot u+\mu\cdot u$ (pour tous $\lambda,\mu\in\mathbb{K}$, $u\in E$)

Nous reviendrons en détail sur chacune de ces propriétés juste après des exemples.

1.2. Premiers exemples

Exemple 41. Le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$

Posons $\mathbb{K}=\mathbb{R}$ et $E=\mathbb{R}^2$. Un élément $u\in E$ est donc un couple (x,y) avec x élément de $\mathbb{R}$ et y élément de $\mathbb{R}$. Ceci s'écrit

$$\mathbb{R}^2=\big\{(x,y)\mid x\in\mathbb{R},y\in\mathbb{R}\big\}.$$

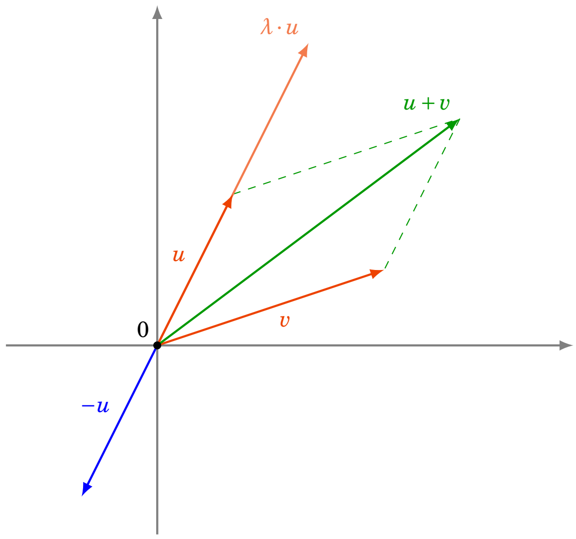
– *Définition de la loi interne.* Si (x,y) et (x',y') sont deux éléments de $\mathbb{R}^2$, alors :

$$(x,y)+(x',y')=(x+x',y+y').$$

– *Définition de la loi externe.* Si λ est un réel et (x,y) est un élément de $\mathbb{R}^2$, alors :

$$\lambda\cdot(x,y)=(\lambda x,\lambda y).$$

L'élément neutre de la loi interne est le vecteur nul $(0,0)$. Le symétrique de (x,y) est $(-x,-y)$, que l'on note aussi $-(x,y)$.



L'exemple suivant généralise le précédent. C'est aussi le bon moment pour lire ou relire le chapitre « L'espace vectoriel $\mathbb{R}^n$ ».

Exemple 42. Le $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$

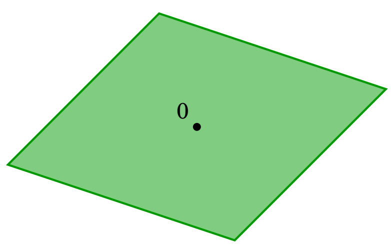
Soit n un entier supérieur ou égal à 1. Posons $\mathbb{K}=\mathbb{R}$ et $E=\mathbb{R}^n$. Un élément $u \in E$ est donc un n-uplet $(x_1, x_2, \ldots, x_n)$ avec $x_1, x_2, \ldots, x_n$ des éléments de $\mathbb{R}$.

- *Définition de la loi interne.* Si $(x_1, \ldots, x_n)$ et $(x'_1, \ldots, x'_n)$ sont deux éléments de $\mathbb{R}^n$, alors :
$$(x_1, \ldots, x_n) + (x'_1, \ldots, x'_n) = (x_1 + x'_1, \ldots, x_n + x'_n).$$
- *Définition de la loi externe.* Si λ est un réel et $(x_1, \ldots, x_n)$ est un élément de $\mathbb{R}^n$, alors :
$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

L'élément neutre de la loi interne est le vecteur nul $(0, 0, \ldots, 0)$. Le symétrique de $(x_1, \ldots, x_n)$ est $(-x_1, \ldots, -x_n)$, que l'on note $-(x_1, \ldots, x_n)$.
De manière analogue, on peut définir le $\mathbb{C}$-espace vectoriel $\mathbb{C}^n$, et plus généralement le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$.

Exemple 43

Tout plan passant par l'origine dans $\mathbb{R}^3$ est un espace vectoriel (par rapport aux opérations habituelles sur les vecteurs). Soient $\mathbb{K}=\mathbb{R}$ et $E=\mathscr{P}$ un plan passant par l'origine. Le plan admet une équation de la forme :
$$ax + by + cz = 0$$
où a, b et c sont des réels non tous nuls.



Un élément $u \in E$ est donc un triplet (noté ici comme un vecteur colonne) $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ tel que $ax + by +$

$cz = 0$.

Soient $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$ deux éléments de $\mathscr{P}$. Autrement dit,

$$\begin{array}{rrcl} & ax + by + cz & = & 0, \\ \text{et} & ax' + by' + cz' & = & 0. \end{array}$$

Alors $\begin{pmatrix} x+x' \\ y+y' \\ z+z' \end{pmatrix}$ est aussi dans $\mathscr{P}$ car on a bien :

$$a(x + x') + b(y + y') + c(z + z') = 0.$$

Les autres propriétés sont aussi faciles à vérifier : par exemple l'élément neutre est $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$; et si $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ appartient à $\mathscr{P}$, alors $ax + by + cz = 0$, que l'on peut réécrire $a(-x) + b(-y) + c(-z) = 0$ et ainsi $-\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ appartient à $\mathscr{P}$.

Attention ! Un plan ne contenant pas l'origine n'est pas un espace vectoriel, car justement il ne contient pas le vecteur nul $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

1.3. Terminologie et notations

Rassemblons les définitions déjà vues.

- On appelle les éléments de E des ***vecteurs***. Au lieu de $\mathbb{K}$-espace vectoriel, on dit aussi espace vectoriel sur $\mathbb{K}$.
- Les éléments de $\mathbb{K}$ seront appelés des ***scalaires***.
- L'***élément neutre*** 0_E s'appelle aussi le ***vecteur nul***. Il ne doit pas être confondu avec l'élément 0 de $\mathbb{K}$. Lorsqu'il n'y aura pas de risque de confusion, 0_E sera aussi noté 0.
- Le ***symétrique*** $-u$ d'un vecteur $u \in E$ s'appelle aussi l'***opposé***.
- La loi de composition interne sur E (notée usuellement +) est appelée couramment l'addition et $u + u'$ est appelée somme des vecteurs u et u'.
- La loi de composition externe sur E est appelée couramment multiplication par un scalaire. La multiplication du vecteur u par le scalaire λ sera souvent notée simplement λu, au lieu de $\lambda \cdot u$.

Somme de n vecteurs. Il est possible de définir, par récurrence, l'addition de n vecteurs, $n \geqslant 2$. La structure d'espace vectoriel permet de définir l'addition de deux vecteurs (et initialise le processus). Si maintenant la somme de $n-1$ vecteurs est définie, alors la somme de n vecteurs $v_1, v_2, \ldots, v_n$ est définie par

$$v_1 + v_2 + \cdots + v_n = (v_1 + v_2 + \cdots + v_{n-1}) + v_n.$$

L'associativité de la loi + nous permet de ne pas mettre de parenthèses dans la somme $v_1 + v_2 + \cdots + v_n$.

On notera $v_1 + v_2 + \cdots + v_n = \sum_{i=1}^{n} v_i$.

1.4. Mini-exercices

1. Vérifier les 8 axiomes qui font de $\mathbb{R}^3$ un $\mathbb{R}$-espace vectoriel.
2. Idem pour une droite $\mathscr{D}$ de $\mathbb{R}^3$ passant par l'origine définie par $\left\{ \begin{array}{rcl} ax + by + cz & = & 0 \\ a'x + b'y + c'z & = & 0 \end{array} \right.$.
3. Justifier que les ensembles suivants *ne sont pas* des espaces vectoriels : $\{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$; $\{(x, y) \in \mathbb{R}^2 \mid x = 1\}$; $\{(x, y) \in \mathbb{R}^2 \mid x \geqslant 0 \text{ et } y \geqslant 0\}$; $\{(x, y) \in \mathbb{R}^2 \mid -1 \leqslant x \leqslant 1 \text{ et } -1 \leqslant y \leqslant 1\}$.

4. Montrer par récurrence que si les v_i sont des éléments d'un $\mathbb{K}$-espace vectoriel E, alors pour tous $\lambda_i \in \mathbb{K}$: $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in E$.

2. Espace vectoriel (fin)

2.1. Détail des axiomes de la définition

Revenons en détail sur la définition d'un espace vectoriel. Soit donc E un $\mathbb{K}$-espace vectoriel. Les éléments de E seront appelés des ***vecteurs***. Les éléments de $\mathbb{K}$ seront appelés des ***scalaires***.

Loi interne.

La loi de composition interne dans E, c'est une application de $E \times E$ dans E :

$$\begin{array}{ccc} E \times E & \to & E \\ (u,v) & \mapsto & u+v \end{array}$$

C'est-à-dire qu'à partir de deux vecteurs u et v de E, on nous en fournit un troisième, qui sera noté $u+v$.

La loi de composition interne dans E et la somme dans $\mathbb{K}$ seront toutes les deux notées $+$, mais le contexte permettra de déterminer aisément de quelle loi il s'agit.

Loi externe.

La loi de composition externe, c'est une application de $\mathbb{K} \times E$ dans E :

$$\begin{array}{ccc} \mathbb{K} \times E & \to & E \\ (\lambda,u) & \mapsto & \lambda \cdot u \end{array}$$

C'est-à-dire qu'à partir d'un scalaire $\lambda \in \mathbb{K}$ et d'un vecteur $u \in E$, on nous fournit un autre vecteur, qui sera noté $\lambda \cdot u$.

Axiomes relatifs à la loi interne.

1. ***Commutativité.*** Pour tous $u, v \in E$, $u+v = v+u$. On peut donc additionner des vecteurs dans l'ordre que l'on souhaite.
2. ***Associativité.*** Pour tous $u, v, w \in E$, on a $u+(v+w) = (u+v)+w$. Conséquence : on peut « oublier » les parenthèses et noter sans ambiguïté $u+v+w$.
3. Il existe un ***élément neutre***, c'est-à-dire qu'il existe un élément de E, noté 0_E, vérifiant : pour tout $u \in E$, $u+0_E = u$ (et on a aussi $0_E + u = u$ par commutativité). Cet élément 0_E s'appelle aussi le ***vecteur nul***.
4. Tout élément u de E admet un ***symétrique*** (ou ***opposé***), c'est-à-dire qu'il existe un élément u' de E tel que $u+u' = 0_E$ (et on a aussi $u'+u = 0_E$ par commutativité). Cet élément u' de E est noté $-u$.

Proposition 11

- S'il existe un élément neutre 0_E vérifiant l'axiome (3) ci-dessus, alors il est unique.
- Soit u un élément de E. S'il existe un élément symétrique u' de E vérifiant l'axiome (4), alors il est unique.

Démonstration

- Soient 0_E et $0'_E$ deux éléments vérifiant la définition de l'élément neutre. On a alors, pour tout élément u de E :

$$u + 0_E = 0_E + u = u \qquad \text{et} \qquad u + 0'_E = 0'_E + u = u$$

  - Alors, la première propriété utilisée avec $u = 0'_E$ donne $0'_E + 0_E = 0_E + 0'_E = 0'_E$.
  - La deuxième propriété utilisée avec $u = 0_E$ donne $0_E + 0'_E = 0'_E + 0_E = 0_E$.
  - En comparant ces deux résultats, il vient $0_E = 0'_E$.
- Supposons qu'il existe deux symétriques de u notés u' et u''. On a :

$$u + u' = u' + u = 0_E \qquad \text{et} \qquad u + u'' = u'' + u = 0_E.$$

  Calculons $u' + (u + u'')$ de deux façons différentes, en utilisant l'associativité de la loi + et les relations précédentes.
  - $u' + (u + u'') = u' + 0_E = u'$
  - $u' + (u + u'') = (u' + u) + u'' = 0_E + u'' = u''$
  - On en déduit $u' = u''$.

Remarque

Les étudiants connaissant la théorie des groupes reconnaîtront, dans les quatre premiers axiomes ci-dessus, les axiomes caractérisant un groupe commutatif.

Axiomes relatifs à la loi externe.

5. Soit 1 l'élément neutre de la multiplication de $\mathbb{K}$. Pour tout élément u de E, on a

$$1 \cdot u = u.$$

6. Pour tous éléments λ et μ de $\mathbb{K}$ et pour tout élément u de E, on a

$$\lambda \cdot (\mu \cdot u) = (\lambda \times \mu) \cdot u.$$

Axiomes liant les deux lois.

7. ***Distributivité*** par rapport à l'addition des vecteurs. Pour tout élément λ de $\mathbb{K}$ et pour tous éléments u et v de E, on a

$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v.$$

8. ***Distributivité*** par rapport à l'addition des scalaires. Pour tous λ et μ de $\mathbb{K}$ et pour tout élément u de E, on a :

$$(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u.$$

La loi interne et la loi externe doivent donc satisfaire ces huit axiomes pour que $(E, +, \cdot)$ soit un espace vectoriel sur $\mathbb{K}$.

2.2. Exemples

Dans tous les exemples qui suivent, la vérification des axiomes se fait simplement et est laissée au soin des étudiants. Seules seront indiquées, dans chaque cas, les valeurs de l'élément neutre de la loi interne et du symétrique d'un élément.

Exemple 44. L'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$

L'ensemble des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ est noté $\mathscr{F}(\mathbb{R},\mathbb{R})$. Nous le munissons d'une structure de $\mathbb{R}$-espace vectoriel de la manière suivante.

– *Loi interne.* Soient f et g deux éléments de $\mathscr{F}(\mathbb{R},\mathbb{R})$. La fonction $f+g$ est définie par :

$$\forall x \in \mathbb{R} \quad (f+g)(x) = f(x)+g(x)$$

(où le signe $+$ désigne la loi interne de $\mathscr{F}(\mathbb{R},\mathbb{R})$ dans le membre de gauche et l'addition dans $\mathbb{R}$ dans le membre de droite).

– *Loi externe.* Si λ est un nombre réel et f une fonction de $\mathscr{F}(\mathbb{R},\mathbb{R})$, la fonction $\lambda \cdot f$ est définie par l'image de tout réel x comme suit :

$$\forall x \in \mathbb{R} \quad (\lambda \cdot f)(x) = \lambda \times f(x).$$

(Nous désignons par $\cdot$ la loi externe de $\mathscr{F}(\mathbb{R},\mathbb{R})$ et par $\times$ la multiplication dans $\mathbb{R}$. Avec l'habitude on oubliera les signes de multiplication : $(\lambda f)(x) = \lambda f(x)$.)

– *Élément neutre.* L'élément neutre pour l'addition est la fonction nulle, définie par :

$$\forall x \in \mathbb{R} \quad f(x) = 0.$$

On peut noter cette fonction $0_{\mathscr{F}(\mathbb{R},\mathbb{R})}$.

– *Symétrique.* Le symétrique de l'élément f de $\mathscr{F}(\mathbb{R},\mathbb{R})$ est l'application g de $\mathbb{R}$ dans $\mathbb{R}$ définie par :

$$\forall x \in \mathbb{R} \quad g(x) = -f(x).$$

Le symétrique de f est noté $-f$.

Exemple 45. Le $\mathbb{R}$-espace vectoriel des suites réelles

On note $\mathscr{S}$ l'ensemble des suites réelles $(u_n)_{n\in\mathbb{N}}$. Cet ensemble peut être vu comme l'ensemble des applications de $\mathbb{N}$ dans $\mathbb{R}$; autrement dit $\mathscr{S} = \mathscr{F}(\mathbb{N},\mathbb{R})$.

– *Loi interne.* Soient $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ deux suites appartenant à $\mathscr{S}$. La suite $u+v$ est la suite $w = (w_n)_{n\in\mathbb{N}}$ dont le terme général est défini par

$$\forall n \in \mathbb{N} \quad w_n = u_n + v_n$$

(où $u_n + v_n$ désigne la somme de u_n et de v_n dans $\mathbb{R}$).

– *Loi externe.* Si λ est un nombre réel et $u = (u_n)_{n\in\mathbb{N}}$ un élément de $\mathscr{S}$, $\lambda \cdot u$ est la suite $v = (v_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N} \quad v_n = \lambda \times u_n$$

où $\times$ désigne la multiplication dans $\mathbb{R}$.

– *Élément neutre.* L'élément neutre de la loi interne est la suite dont tous les termes sont nuls.

– *Symétrique.* Le symétrique de la suite $u = (u_n)_{n\in\mathbb{N}}$ est la suite $u' = (u'_n)_{n\in\mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} \quad u'_n = -u_n.$$

Elle est notée $-u$.

Exemple 46. Les matrices

L'ensemble $M_{n,p}(\mathbb{R})$ des matrices à n lignes et p colonnes à coefficients dans $\mathbb{R}$ est muni d'une structure de $\mathbb{R}$-espace vectoriel. La loi interne est l'addition de deux matrices. La loi externe est la multiplication d'une matrice par un scalaire. L'élément neutre pour la loi interne est la matrice nulle (tous les coefficients sont nuls). Le symétrique de la matrice $A = (a_{i,j})$ est la matrice $(-a_{i,j})$. De même, l'ensemble $M_{n,p}(\mathbb{K})$ des matrices à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.

Autres exemples :

1. L'espace vectoriel $\mathbb{R}[X]$ des polynômes $P(X) = a_nX^n + \cdots + a_2X^2 + a_1X + a_0$. L'addition est l'addition de deux polynômes $P(X)+Q(X)$, la multiplication par un scalaire $\lambda \in \mathbb{R}$ est $\lambda \cdot P(X)$. L'élément neutre est le polynôme nul. L'opposé de $P(X)$ est $-P(X)$.
2. L'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$; l'ensemble des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$,...
3. $\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel : addition $z + z'$ de deux nombres complexes, multiplication λz par un scalaire $\lambda \in \mathbb{R}$. L'élément neutre est le nombre complexe 0 et le symétrique du nombre complexe z est $-z$.

2.3. Règles de calcul

Proposition 12

Soit E un espace vectoriel sur un corps $\mathbb{K}$. Soient $u \in E$ et $\lambda \in \mathbb{K}$. Alors on a :

1. $0 \cdot u = 0_E$
2. $\lambda \cdot 0_E = 0_E$
3. $(-1) \cdot u = -u$
4. $\boxed{\lambda \cdot u = 0_E \iff \lambda = 0 \text{ ou } u = 0_E}$

L'opération qui à (u,v) associe $u + (-v)$ s'appelle la ***soustraction***. Le vecteur $u + (-v)$ est noté $u - v$. Les propriétés suivantes sont satisfaites : $\lambda(u - v) = \lambda u - \lambda v$ et $(\lambda - \mu)u = \lambda u - \mu u$.

Démonstration

Les démonstrations des propriétés sont des manipulations sur les axiomes définissant les espaces vectoriels.

1. – Le point de départ de la démonstration est l'égalité dans $\mathbb{K}$: $0 + 0 = 0$.
   – D'où, pour tout vecteur de E, l'égalité $(0 + 0) \cdot u = 0 \cdot u$.
   – Donc, en utilisant la distributivité de la loi externe par rapport à la loi interne et la définition de l'élément neutre, on obtient $0 \cdot u + 0 \cdot u = 0 \cdot u$. On peut rajouter l'élément neutre dans le terme de droite, pour obtenir : $0 \cdot u + 0 \cdot u = 0 \cdot u + 0_E$.
   – En ajoutant $-(0 \cdot u)$ de chaque côté de l'égalité, on obtient : $0 \cdot u = 0_E$.
2. La preuve est semblable en partant de l'égalité $0_E + 0_E = 0_E$.
3. Montrer $(-1) \cdot u = -u$ signifie exactement que $(-1) \cdot u$ est le symétrique de u, c'est-à-dire vérifie $u + (-1) \cdot u = 0_E$. En effet :

$$u + (-1) \cdot u = 1 \cdot u + (-1) \cdot u = (1 + (-1)) \cdot u = 0 \cdot u = 0_E.$$

4. On sait déjà que si $\lambda = 0$ ou $u = 0_E$, alors les propriétés précédentes impliquent $\lambda \cdot u = 0_E$.

Pour la réciproque, soient $\lambda \in \mathbb{K}$ un scalaire et $u \in E$ un vecteur tels que $\lambda \cdot u = 0_E$.
Supposons λ différent de 0. On doit alors montrer que $u = 0_E$.

- Comme $\lambda \neq 0$, alors λ est inversible pour le produit dans le corps $\mathbb{K}$. Soit λ^{-1} son inverse.
- En multipliant par λ^{-1} les deux membres de l'égalité $\lambda \cdot u = 0_E$, il vient : $\lambda^{-1} \cdot (\lambda \cdot u) = \lambda^{-1} \cdot 0_E$.
- D'où en utilisant les propriétés de la multiplication par un scalaire $(\lambda^{-1} \times \lambda) \cdot u = 0_E$ et donc $1 \cdot u = 0_E$.
- D'où $u = 0_E$.

2.4. Mini-exercices

1. Justifier si les objets suivants sont des espaces vectoriels.
   (a) L'ensemble des fonctions réelles sur $[0,1]$, continues, positives ou nulles, pour l'addition et le produit par un réel.
   (b) L'ensemble des fonctions réelles sur $\mathbb{R}$ vérifiant $\lim_{x \to +\infty} f(x) = 0$ pour les mêmes opérations.
   (c) L'ensemble des fonctions sur $\mathbb{R}$ telles que $f(3) = 7$.
   (d) L'ensemble $\mathbb{R}_+^*$ pour les opérations $x \oplus y = xy$ et $\lambda \cdot x = x^\lambda$ ($\lambda \in \mathbb{R}$).
   (e) L'ensemble des points (x, y) de $\mathbb{R}^2$ vérifiant $\sin(x + y) = 0$.
   (f) L'ensemble des vecteurs (x, y, z) de $\mathbb{R}^3$ orthogonaux au vecteur $(-1, 3, -2)$.
   (g) L'ensemble des fonctions de classe $\mathscr{C}^2$ vérifiant $f'' + f = 0$.
   (h) L'ensemble des fonctions continues sur $[0,1]$ vérifiant $\int_0^1 f(x) \sin x \, dx = 0$.
   (i) L'ensemble des matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{R})$ vérifiant $a + d = 0$.
2. Prouver les propriétés de la soustraction : $\lambda \cdot (u - v) = \lambda \cdot u - \lambda \cdot v$ et $(\lambda - \mu) \cdot u = \lambda \cdot u - \mu \cdot u$.

3. Sous-espace vectoriel (début)

Il est vite fatiguant de vérifier les 8 axiomes qui font d'un ensemble un espace vectoriel. Heureusement, il existe une manière rapide et efficace de prouver qu'un ensemble est un espace vectoriel : grâce à la notion de sous-espace vectoriel.

3.1. Définition d'un sous-espace vectoriel

Définition 23

Soit E un $\mathbb{K}$-espace vectoriel. Une partie F de E est appelée un ***sous-espace vectoriel*** si :

- $0_E \in F$,
- $u + v \in F$ pour tous $u, v \in F$,
- $\lambda \cdot u \in F$ pour tout $\lambda \in \mathbb{K}$ et tout $u \in F$.

Remarque

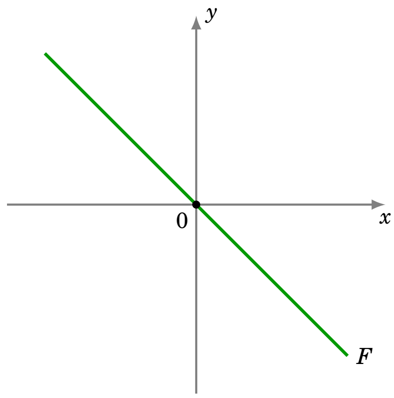
Expliquons chaque condition.

- La première condition signifie que le vecteur nul de E doit aussi être dans F. En fait il suffit même de prouver que F est non vide.
- La deuxième condition, c'est dire que F est stable pour l'addition : la somme $u + v$ de deux vecteurs u, v de F est bien sûr un vecteur de E (car E est un espace vectoriel), mais ici on exige que $u + v$ soit un élément de F.

– La troisième condition, c'est dire que F est stable pour la multiplication par un scalaire.

Exemple 47. Exemples immédiats

1. L'ensemble $F = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$. En effet :
   (a) $(0,0) \in F$,
   (b) si $u = (x_1, y_1)$ et $v = (x_2, y_2)$ appartiennent à F, alors $x_1 + y_1 = 0$ et $x_2 + y_2 = 0$ donc $(x_1 + x_2) + (y_1 + y_2) = 0$ et ainsi $u + v = (x_1 + x_2, y_1 + y_2)$ appartient à F,
   (c) si $u = (x, y) \in F$ et $\lambda \in \mathbb{R}$, alors $x + y = 0$ donc $\lambda x + \lambda y = 0$, d'où $\lambda u \in F$.

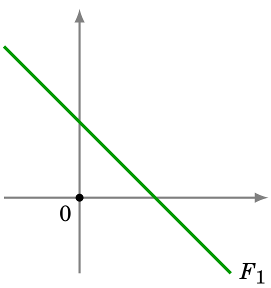


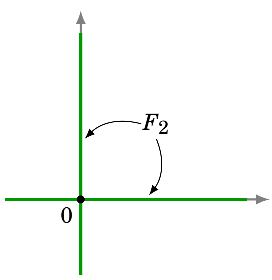
2. L'ensemble des fonctions continues sur $\mathbb{R}$ est un sous-espace vectoriel de l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Preuve : la fonction nulle est continue ; la somme de deux fonctions continues est continue ; une constante fois une fonction continue est une fonction continue.
3. L'ensemble des suites réelles convergentes est un sous-espace vectoriel de l'espace vectoriel des suites réelles.

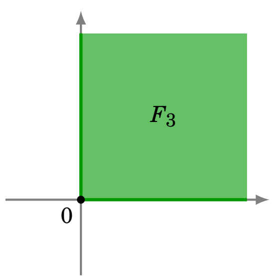
Voici des sous-ensembles qui *ne sont pas* des sous-espaces vectoriels.

Exemple 48

1. L'ensemble $F_1 = \{(x, y) \in \mathbb{R}^2 \mid x + y = 2\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En effet le vecteur nul $(0,0)$ n'appartient pas à F_1.
2. L'ensemble $F_2 = \{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ ou } y = 0\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En effet les vecteurs $u = (1,0)$ et $v = (0,1)$ appartiennent à F_2, mais pas le vecteur $u + v = (1,1)$.
3. L'ensemble $F_3 = \{(x, y) \in \mathbb{R}^2 \mid x \geqslant 0 \text{ et } y \geqslant 0\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En effet le vecteur $u = (1,1)$ appartient à F_3 mais, pour $\lambda = -1$, le vecteur $-u = (-1,-1)$ n'appartient pas à F_3.







3.2. Un sous-espace vectoriel est un espace vectoriel

La notion de sous-espace vectoriel prend tout son intérêt avec le théorème suivant : un sous-espace vectoriel est lui-même un espace vectoriel. C'est ce théorème qui va nous fournir plein d'exemples d'espaces vectoriels.

Théorème 13

Soient E un $\mathbb{K}$-espace vectoriel et F un sous-espace vectoriel de E. Alors F est lui-même un $\mathbb{K}$-espace vectoriel pour les lois induites par E.

Méthodologie. Pour répondre à une question du type « L'ensemble F est-il un espace vectoriel ? », une façon efficace de procéder est de trouver un espace vectoriel E qui contient F, puis prouver que F est un sous-espace vectoriel de E. Il y a seulement trois propriétés à vérifier au lieu de huit !

Exemple 49

1. Est-ce que l'ensemble des fonctions paires (puis des fonctions impaires) forme un espace vectoriel (sur $\mathbb{R}$ avec les lois usuelles sur les fonctions) ?
   Notons $\mathscr{P}$ l'ensemble des fonctions paires et $\mathscr{I}$ l'ensemble des fonctions impaires. Ce sont deux sous-ensembles de l'espace vectoriel $\mathscr{F}(\mathbb{R},\mathbb{R})$ des fonctions.
   $$\mathscr{P} = \left\{ f \in \mathscr{F}(\mathbb{R},\mathbb{R}) \mid \forall x \in \mathbb{R}, f(-x) = f(x) \right\}$$
   $$\mathscr{I} = \left\{ f \in \mathscr{F}(\mathbb{R},\mathbb{R}) \mid \forall x \in \mathbb{R}, f(-x) = -f(x) \right\}$$
   $\mathscr{P}$ et $\mathscr{I}$ sont des sous-espaces vectoriels de $\mathscr{F}(\mathbb{R},\mathbb{R})$. C'est très simple à vérifier, par exemple pour $\mathscr{P}$:
   (a) la fonction nulle est une fonction paire,
   (b) si $f, g \in \mathscr{P}$ alors $f + g \in \mathscr{P}$,
   (c) si $f \in \mathscr{P}$ et si $\lambda \in \mathbb{R}$ alors $\lambda f \in \mathscr{P}$.
   Par le théorème 13, $\mathscr{P}$ est un espace vectoriel (de même pour $\mathscr{I}$).
2. Est-ce que l'ensemble $\mathscr{S}_n$ des matrices symétriques de taille n est un espace vectoriel (sur $\mathbb{R}$ avec les lois usuelles sur les matrices) ?
   $\mathscr{S}_n$ est un sous-ensemble de l'espace vectoriel $M_n(\mathbb{R})$. Et c'est même un sous-espace vectoriel. Il suffit en effet de vérifier que la matrice nulle est symétrique, que la somme de deux matrices symétriques est encore symétrique et finalement que le produit d'une matrice symétrique par un scalaire est une matrice symétrique. Par le théorème 13, $\mathscr{S}_n$ est un espace vectoriel.

Démonstration . Preuve du théorème 13

Soit F un sous-espace vectoriel d'un espace vectoriel $(E, +, \cdot)$. La stabilité de F pour les deux lois permet de munir cet ensemble d'une loi de composition interne et d'une loi de composition externe, en restreignant à F les opérations définies dans E. Les propriétés de commutativité et d'associativité de l'addition, ainsi que les quatre axiomes relatifs à la loi externe sont vérifiés, car ils sont satisfaits dans E donc en particulier dans F, qui est inclus dans E.
L'existence d'un élément neutre découle de la définition de sous-espace vectoriel. Il reste seulement à justifier que si $u \in F$, alors son symétrique $-u$ appartient à F.
Fixons $u \in F$. Comme on a aussi $u \in E$ et que E est un espace vectoriel alors il existe un élément de

E, noté $-u$, tel que $u+(-u)=0_E$. Comme u est élément de F, alors pour $\lambda=-1$, $(-1)u\in F$. Et ainsi $-u$ appartient à F.

Un autre exemple d'espace vectoriel est donné par l'ensemble des solutions d'un système linéaire homogène. Soit $AX=0$ un système de n équations à p inconnues :

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

On a alors

Théorème 14

Soit $A\in M_{n,p}(\mathbb{R})$. Soit $AX=0$ un système d'équations linéaires homogènes à p variables. Alors l'ensemble des vecteurs solutions est un sous-espace vectoriel de $\mathbb{R}^p$.

Démonstration

Soit F l'ensemble des vecteurs $X\in\mathbb{R}^p$ solutions de l'équation $AX=0$. Vérifions que F est un sous-espace vectoriel de $\mathbb{R}^p$.

- Le vecteur 0 est un élément de F.
- F est stable par addition : si X et X' sont des vecteurs solutions, alors $AX=0$ et $AX'=0$, donc $A(X+X')=AX+AX'=0$, et ainsi $X+X'\in F$.
- F est stable par multiplication par un scalaire : si X est un vecteur solution, on a aussi $A(\lambda X)=\lambda(AX)=\lambda 0=0$, ceci pour tout $\lambda\in\mathbb{R}$. Donc $\lambda X\in F$.

Exemple 50

Considérons le système

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

L'ensemble des solutions $F\subset\mathbb{R}^3$ de ce système est :

$$F=\{(x=2s-3t, y=s, z=t)\mid s,t\in\mathbb{R}\}.$$

Par le théorème 14, F est un sous-espace vectoriel de $\mathbb{R}^3$. Donc par le théorème 13, F est un espace vectoriel.

Une autre façon de voir les choses est d'écrire que les éléments de F sont ceux qui vérifient l'équation $(x=2y-3z)$. Autrement dit, F est d'équation $(x-2y+3z=0)$. L'ensemble des solutions F est donc un plan passant par l'origine. Nous avons déjà vu que ceci est un espace vectoriel.

3.3. Mini-exercices

Parmi les ensembles suivants, reconnaître ceux qui sont des sous-espaces vectoriels :

1. $\{(x,y,z)\in\mathbb{R}^3\mid x+y=0\}$
2. $\{(x,y,z,t)\in\mathbb{R}^4\mid x=t \text{ et } y=z\}$
3. $\{(x,y,z)\in\mathbb{R}^3\mid z=1\}$

4. $\{(x,y)\in\mathbb{R}^2 \mid x^2+xy\geqslant 0\}$
5. $\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\geqslant 1\}$
6. $\{f\in\mathscr{F}(\mathbb{R},\mathbb{R}) \mid f(0)=1\}$
7. $\{f\in\mathscr{F}(\mathbb{R},\mathbb{R}) \mid f(1)=0\}$
8. $\{f\in\mathscr{F}(\mathbb{R},\mathbb{R}) \mid f \text{ est croissante }\}$
9. $\{(u_n)_{n\in\mathbb{N}} \mid (u_n) \text{ tend vers } 0\}$

4. Sous-espace vectoriel (milieu)

4.1. Combinaisons linéaires

Définition 24

Soit $n\geqslant 1$ un entier, soient $v_1,v_2,\dots,v_n$, n vecteurs d'un espace vectoriel E. Tout vecteur de la forme

$$u=\lambda_1v_1+\lambda_2v_2+\cdots+\lambda_nv_n$$

(où $\lambda_1,\lambda_2,\dots,\lambda_n$ sont des éléments de $\mathbb{K}$) est appelé ***combinaison linéaire*** des vecteurs $v_1,v_2,\dots,v_n$. Les scalaires $\lambda_1,\lambda_2,\dots,\lambda_n$ sont appelés ***coefficients*** de la combinaison linéaire.

Remarque : Si $n=1$, alors $u=\lambda_1v_1$ et on dit que u est ***colinéaire*** à v_1.

Exemple 51

1. Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$, $(3,3,1)$ est combinaison linéaire des vecteurs $(1,1,0)$ et $(1,1,1)$ car on a l'égalité
$$(3,3,1)=2(1,1,0)+(1,1,1).$$
2. Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$, le vecteur $u=(2,1)$ *n'est pas* colinéaire au vecteur $v_1=(1,1)$ car s'il l'était, il existerait un réel λ tel que $u=\lambda v_1$, ce qui équivaudrait à l'égalité $(2,1)=(\lambda,\lambda)$.
3. Soit $E=\mathscr{F}(\mathbb{R},\mathbb{R})$ l'espace vectoriel des fonctions réelles. Soient f_0, f_1, f_2 et f_3 les fonctions définies par :
$$\forall x\in\mathbb{R}\quad f_0(x)=1,\ f_1(x)=x,\ f_2(x)=x^2,\ f_3(x)=x^3.$$
Alors la fonction f définie par
$$\forall x\in\mathbb{R}\quad f(x)=x^3-2x^2-7x-4$$
est combinaison linéaire des fonctions f_0,f_1,f_2,f_3 puisque l'on a l'égalité
$$f=f_3-2f_2-7f_1-4f_0.$$
4. Dans $M_{2,3}(\mathbb{R})$, on considère $A=\begin{pmatrix}1&1&3\\0&-1&4\end{pmatrix}$. On peut écrire A naturellement sous la forme suivante d'une combinaison linéaire de matrices élémentaires (des zéros partout, sauf un 1) :
$$A=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&1&0\\0&0&0\end{pmatrix}+3\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}-\begin{pmatrix}0&0&0\\0&1&0\end{pmatrix}+4\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}.$$

Voici deux exemples plus compliqués.

Exemple 52

Soient $u = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ et $v = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}$ deux vecteurs de $\mathbb{R}^3$. Montrons que $w = \begin{pmatrix} 9 \\ 2 \\ 7 \end{pmatrix}$ est combinaison linéaire de u et v. On cherche donc λ et μ tels que $w = \lambda u + \mu v$:

$$\begin{pmatrix} 9 \\ 2 \\ 7 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} \lambda \\ 2\lambda \\ -\lambda \end{pmatrix} + \begin{pmatrix} 6\mu \\ 4\mu \\ 2\mu \end{pmatrix} = \begin{pmatrix} \lambda + 6\mu \\ 2\lambda + 4\mu \\ -\lambda + 2\mu \end{pmatrix}.$$

On a donc

$$\begin{cases} 9 &= \lambda + 6\mu \\ 2 &= 2\lambda + 4\mu \\ 7 &= -\lambda + 2\mu. \end{cases}$$

Une solution de ce système est $(\lambda = -3, \mu = 2)$, ce qui implique que w est combinaison linéaire de u et v. On vérifie que l'on a bien

$$\begin{pmatrix} 9 \\ 2 \\ 7 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}.$$

Exemple 53

Soient $u = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ et $v = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}$. Montrons que $w = \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix}$ n'est pas une combinaison linéaire de u et v. L'égalité

$$\begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} \quad \text{équivaut au système} \quad \begin{cases} 4 &= \lambda + 6\mu \\ -1 &= 2\lambda + 4\mu \\ 8 &= -\lambda + 2\mu. \end{cases}$$

Or ce système n'a aucune solution. Donc il n'existe pas $\lambda, \mu \in \mathbb{R}$ tels que $w = \lambda u + \mu v$.

4.2. Caractérisation d'un sous-espace vectoriel

Théorème 15. Caractérisation d'un sous-espace par la notion de combinaison linéaire

Soient E un $\mathbb{K}$-espace vectoriel et F une partie non vide de E. F est un sous-espace vectoriel de E si et seulement si

$$\lambda u + \mu v \in F \qquad \text{pour tous } u, v \in F \qquad \text{et tous } \lambda, \mu \in \mathbb{K}.$$

Autrement dit si et seulement si toute combinaison linéaire de deux éléments de F appartient à F.

Démonstration

- Supposons que F soit un sous-espace vectoriel. Et soient $u, v \in F$, $\lambda, \mu \in \mathbb{K}$. Alors par la définition de sous-espace vectoriel : $\lambda u \in F$ et $\mu v \in F$ et ainsi $\lambda u + \mu v \in F$.
- Réciproquement, supposons que pour chaque $u, v \in F$, $\lambda, \mu \in \mathbb{K}$ on a $\lambda u + \mu v \in F$.
  - Comme F n'est pas vide, soient $u, v \in F$. Posons $\lambda = \mu = 0$. Alors $\lambda u + \mu v = 0_E \in F$.
  - Si $u, v \in F$, alors en posant $\lambda = \mu = 1$ on obtient $u + v \in F$.

– Si $u \in F$ et $\lambda \in \mathbb{K}$ (et pour n'importe quel v, en posant $\mu = 0$), alors $\lambda u \in F$.

4.3. Intersection de deux sous-espaces vectoriels

Proposition 13. Intersection de deux sous-espaces

Soient F, G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. L'intersection $F \cap G$ est un sous-espace vectoriel de E.

On démontrerait de même que l'intersection $F_1 \cap F_2 \cap F_3 \cap \cdots \cap F_n$ d'une famille quelconque de sous-espaces vectoriels de E est un sous-espace vectoriel de E.

Démonstration

Soient F et G deux sous-espaces vectoriels de E.

– $0_E \in F$, $0_E \in G$ car F et G sont des sous-espaces vectoriels de E ; donc $0_E \in F \cap G$.
– Soient u et v deux vecteurs de $F \cap G$. Comme F est un sous-espace vectoriel, alors $u, v \in F$ implique $u + v \in F$. De même $u, v \in G$ implique $u + v \in G$. Donc $u + v \in F \cap G$.
– Soient $u \in F \cap G$ et $\lambda \in \mathbb{K}$. Comme F est un sous-espace vectoriel, alors $u \in F$ implique $\lambda u \in F$. De même $u \in G$ implique $\lambda u \in G$. Donc $\lambda u \in F \cap G$.

Conclusion : $F \cap G$ est un sous-espace vectoriel de E.

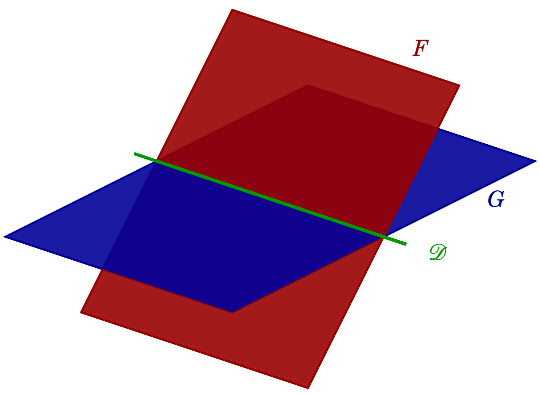
Exemple 54

Soit $\mathscr{D}$ le sous-ensemble de $\mathbb{R}^3$ défini par :

$$\mathscr{D} = \left\{(x, y, z) \in \mathbb{R}^3 \mid x + 3y + z = 0 \quad \text{et} \quad x - y + 2z = 0\right\}.$$

Est-ce que $\mathscr{D}$ est sous-espace vectoriel de $\mathbb{R}^3$? L'ensemble $\mathscr{D}$ est l'intersection de F et G, les sous-ensembles de $\mathbb{R}^3$ définis par :

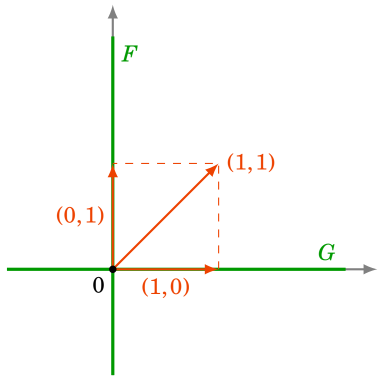
$$F = \left\{(x, y, z) \in \mathbb{R}^3 \mid x + 3y + z = 0\right\}$$
$$G = \left\{(x, y, z) \in \mathbb{R}^3 \mid x - y + 2z = 0\right\}$$



Ce sont deux plans passant par l'origine, donc des sous-espaces vectoriels de $\mathbb{R}^3$. Ainsi $\mathscr{D} = F \cap G$ est un sous-espace vectoriel de $\mathbb{R}^3$, c'est une droite vectorielle.

Remarque

La réunion de deux sous-espaces vectoriels de E n'est pas en général un sous-espace vectoriel de E. Prenons par exemple $E = \mathbb{R}^2$. Considérons les sous-espaces vectoriels $F = \{(x, y) \mid x = 0\}$ et $G = \{(x, y) \mid y = 0\}$. Alors $F \cup G$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. Par exemple, $(0,1)+(1,0) = (1,1)$ est la somme d'un élément de F et d'un élément de G, mais n'est pas dans $F \cup G$.



4.4. Mini-exercices

1. Peut-on trouver $t \in \mathbb{R}$ tel que les vecteurs $\begin{pmatrix} -2 \\ \sqrt{2} \\ t \end{pmatrix}$ et $\begin{pmatrix} -4\sqrt{2} \\ 4t \\ 2\sqrt{2} \end{pmatrix}$ soient colinéaires ?
2. Peut-on trouver $t \in \mathbb{R}$ tel que le vecteur $\begin{pmatrix} 1 \\ 3t \\ t \end{pmatrix}$ soit une combinaison linéaire de $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ et $\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$?

5. Sous-espace vectoriel (fin)

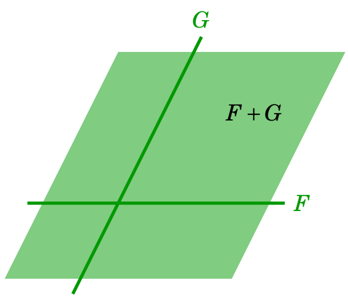
5.1. Somme de deux sous-espaces vectoriels

Comme la réunion de deux sous-espaces vectoriels F et G n'est pas en général un sous-espace vectoriel, il est utile de connaître les sous-espaces vectoriels qui contiennent à la fois les deux sous-espaces vectoriels F et G, et en particulier le plus petit d'entre eux (au sens de l'inclusion).

Définition 25. Définition de la somme de deux sous-espaces

Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. L'ensemble de tous les éléments $u + v$, où u est un élément de F et v un élément de G, est appelé ***somme*** des sous-espaces vectoriels F et G. Cette somme est notée $F + G$. On a donc

$$F + G = \{u + v \mid u \in F, v \in G\}.$$



Proposition 14

Soient F et G deux sous-espaces vectoriels du $\mathbb{K}$-espace vectoriel E.

1. $F+G$ est un sous-espace vectoriel de E.
2. $F+G$ est le plus petit sous-espace vectoriel contenant à la fois F et G.

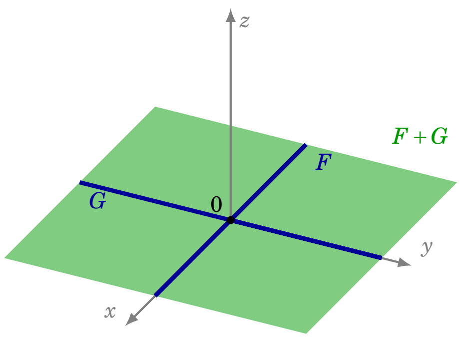
Démonstration

1. Montrons que $F+G$ est un sous-espace vectoriel.
   - $0_E \in F$, $0_E \in G$, donc $0_E = 0_E + 0_E \in F+G$.
   - Soient w et w' des éléments de $F+G$. Comme w est dans $F+G$, il existe u dans F et v dans G tels que $w = u+v$. Comme w' est dans $F+G$, il existe u' dans F et v' dans G tels que $w' = u'+v'$. Alors $w+w' = (u+v)+(u'+v') = (u+u')+(v+v') \in F+G$, car $u+u' \in F$ et $v+v' \in G$.
   - Soit w un élément de $F+G$ et $\lambda \in \mathbb{K}$. Il existe u dans F et v dans G tels que $w = u+v$. Alors $\lambda w = \lambda(u+v) = (\lambda u)+(\lambda v) \in F+G$, car $\lambda u \in F$ et $\lambda v \in G$.
2.
   - L'ensemble $F+G$ contient F et contient G : en effet tout élément u de F s'écrit $u = u+0$ avec u appartenant à F et 0 appartenant à G (puisque G est un sous-espace vectoriel), donc u appartient à $F+G$. De même pour un élément de G.
   - Si H est un sous-espace vectoriel contenant F et G, alors montrons que $F+G \subset H$. C'est clair : si $u \in F$ alors en particulier $u \in H$ (car $F \subset H$), de même si $v \in G$ alors $v \in H$. Comme H est un sous-espace vectoriel, alors $u+v \in H$.

Exemple 55

Déterminons $F+G$ dans le cas où F et G sont les sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\} \qquad \text{et} \qquad G = \{(x,y,z) \in \mathbb{R}^3 \mid x = z = 0\}.$$

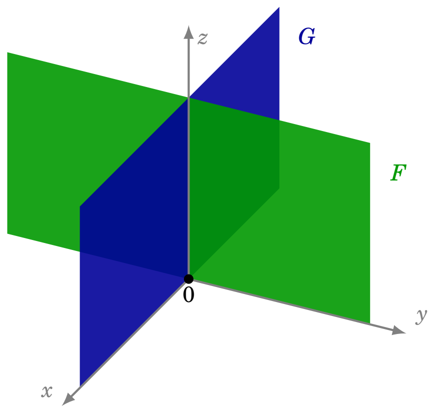


Un élément w de $F+G$ s'écrit $w = u+v$ où u est un élément de F et v un élément de G. Comme $u \in F$ alors il existe $x \in \mathbb{R}$ tel que $u = (x,0,0)$, et comme $v \in G$ il existe $y \in \mathbb{R}$ tel que $v = (0,y,0)$. Donc $w = (x,y,0)$. Réciproquement, un tel élément $w = (x,y,0)$ est la somme de $(x,0,0)$ et de $(0,y,0)$. Donc $F+G = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}$. On voit même que, pour cet exemple, tout élément de $F+G$ s'écrit de façon *unique* comme la somme d'un élément de F et d'un élément de G.

Exemple 56

Soient F et G les deux sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid x = 0\} \qquad \text{et} \qquad G = \{(x,y,z) \in \mathbb{R}^3 \mid y = 0\}.$$



Dans cet exemple, montrons que $F + G = \mathbb{R}^3$. Par définition de $F+G$, tout élément de $F+G$ est dans $\mathbb{R}^3$. Mais réciproquement, si $w = (x,y,z)$ est un élément quelconque de $\mathbb{R}^3$: $w = (x,y,z) = (0,y,z)+(x,0,0)$, avec $(0,y,z) \in F$ et $(x,0,0) \in G$, donc w appartient à $F+G$.
Remarquons que, dans cet exemple, un élément de $\mathbb{R}^3$ ne s'écrit pas forcément de façon unique comme la somme d'un élément de F et d'un élément de G. Par exemple $(1,2,3) = (0,2,3)+(1,0,0) = (0,2,0)+(1,0,3)$.

5.2. Sous-espaces vectoriels supplémentaires

Définition 26. Définition de la somme directe de deux sous-espaces

Soient F et G deux sous-espaces vectoriels de E. F et G sont en ***somme directe*** dans E si

- $F \cap G = \{0_E\}$,
- $F + G = E$.

On note alors $F \oplus G = E$.

Si F et G sont en somme directe, on dit que F et G sont des sous-espaces vectoriels ***supplémentaires*** dans E.

Proposition 15

F et G sont supplémentaires dans E si et seulement si tout élément de E s'écrit d'une manière ***unique*** comme la somme d'un élément de F et d'un élément de G.

Remarque

- Dire qu'un élément w de E s'écrit d'une manière unique comme la somme d'un élément de F et d'un élément de G signifie que si $w = u + v$ avec $u \in F$, $v \in G$ et $w = u' + v'$ avec $u' \in F$, $v' \in G$ alors $u = u'$ et $v = v'$.
- On dit aussi que F est un sous-espace supplémentaire de G (ou que G est un sous-espace supplémentaire de F).
- Il n'y a pas unicité du supplémentaire d'un sous-espace vectoriel donné (voir un exemple ci-dessous).

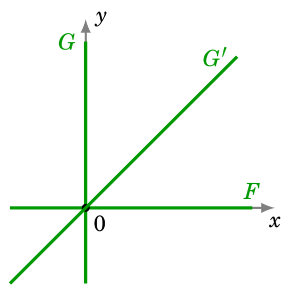
– L'existence d'un supplémentaire d'un sous-espace vectoriel sera prouvée dans le cadre des espaces vectoriels de dimension finie.

Démonstration

- Supposons $E = F \oplus G$ et montrons que tout élément $u \in E$ se décompose de manière unique. Soient donc $u = v + w$ et $u = v' + w'$ avec $v, v' \in F$ et $w, w' \in G$. On a alors $v + w = v' + w'$, donc $v - v' = w' - w$. Comme F est un sous-espace vectoriel alors $v - v' \in F$, mais d'autre part G est aussi un sous-espace vectoriel donc $w' - w \in G$. Conclusion : $v - v' = w' - w \in F \cap G$. Mais par définition d'espaces supplémentaires $F \cap G = \{0_E\}$, donc $v - v' = 0_E$ et aussi $w' - w = 0_E$. On en déduit $v = v'$ et $w = w'$, ce qu'il fallait démontrer.
- Supposons que tout $u \in E$ se décompose de manière unique et montrons $E = F \oplus G$.
  - Montrons $F \cap G = \{0_E\}$. Si $u \in F \cap G$, il peut s'écrire des deux manières suivantes comme somme d'un élément de F et d'un élément de G :
  $$u = 0_E + u \qquad \text{et} \qquad u = u + 0_E.$$
  Par l'unicité de la décomposition, $u = 0_E$.
  - Montrons $F + G = E$. Il n'y rien à prouver, car par hypothèse tout élément u se décompose en $u = v + w$, avec $v \in F$ et $w \in G$.

Exemple 57

1. Soient $F = \{(x,0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ et $G = \{(0,y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}$.
   Montrons que $F \oplus G = \mathbb{R}^2$. La première façon de le voir est que l'on a clairement $F \cap G = \{(0,0)\}$ et que, comme $(x,y) = (x,0) + (0,y)$, alors $F + G = \mathbb{R}^2$. Une autre façon de le voir est d'utiliser la proposition 15, car la décomposition $(x,y) = (x,0) + (0,y)$ est unique.



2. Gardons F et notons $G' = \{(x,x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Montrons que l'on a aussi $F \oplus G' = \mathbb{R}^2$:
   (a) Montrons $F \cap G' = \{(0,0)\}$. Si $(x,y) \in F \cap G'$ alors d'une part $(x,y) \in F$ donc $y = 0$, et aussi $(x,y) \in G'$ donc $x = y$. Ainsi $(x,y) = (0,0)$.
   (b) Montrons $F + G' = \mathbb{R}^2$. Soit $u = (x,y) \in \mathbb{R}^2$. Cherchons $v \in F$ et $w \in G'$ tels que $u = v + w$. Comme $v = (x_1, y_1) \in F$ alors $y_1 = 0$, et comme $w = (x_2, y_2) \in G'$ alors $x_2 = y_2$. Il s'agit donc de trouver x_1 et x_2 tels que
   $$(x,y) = (x_1, 0) + (x_2, x_2).$$
   Donc $(x,y) = (x_1 + x_2, x_2)$. Ainsi $x = x_1 + x_2$ et $y = x_2$, d'où $x_1 = x - y$ et $x_2 = y$. On trouve bien
   $$(x,y) = (x - y, 0) + (y, y),$$
   qui prouve que tout élément de $\mathbb{R}^2$ est somme d'un élément de F et d'un élément de G'.

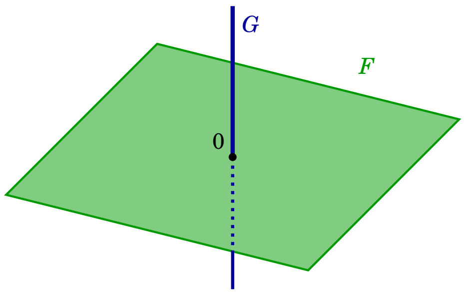
3. De façon plus générale, deux droites distinctes du plan passant par l'origine forment des sous-espaces supplémentaires.

Exemple 58

Est-ce que les sous-espaces vectoriels F et G de $\mathbb{R}^3$ définis par

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\} \qquad \text{et} \qquad G = \{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\}$$

sont supplémentaires dans $\mathbb{R}^3$?



1. Il est facile de vérifier que $F \cap G = \{0\}$. En effet si l'élément $u = (x,y,z)$ appartient à l'intersection de F et de G, alors les coordonnées de u vérifient : $x - y - z = 0$ (car u appartient à F), et $y = z = 0$ (car u appartient à G), donc $u = (0,0,0)$.
2. Il reste à démontrer que $F + G = \mathbb{R}^3$.
   Soit donc $u = (x,y,z)$ un élément quelconque de $\mathbb{R}^3$; il faut déterminer des éléments v de F et w de G tels que $u = v + w$. L'élément v doit être de la forme $v = (y_1 + z_1, y_1, z_1)$ et l'élément w de la forme $w = (x_2, 0, 0)$. On a $u = v + w$ si et seulement si $y_1 = y$, $z_1 = z$, $x_2 = x - y - z$. On a donc

$$(x,y,z) = (y+z, y, z) + (x - y - z, 0, 0)$$

   avec $v = (y+z, y, z)$ dans F et $w = (x - y - z, 0, 0)$ dans G.

Conclusion : $F \oplus G = \mathbb{R}^3$.

Exemple 59

Dans le $\mathbb{R}$-espace vectoriel $\mathscr{F}(\mathbb{R},\mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on considère le sous-espace vectoriel des fonctions paires $\mathscr{P}$ et le sous-espace vectoriel des fonctions impaires $\mathscr{I}$. Montrons que $\mathscr{P} \oplus \mathscr{I} = \mathscr{F}(\mathbb{R},\mathbb{R})$.

1. Montrons $\mathscr{P} \cap \mathscr{I} = \{0_{\mathscr{F}(\mathbb{R},\mathbb{R})}\}$.
   Soit $f \in \mathscr{P} \cap \mathscr{I}$, c'est-à-dire que f est à la fois une fonction paire et impaire. Il s'agit de montrer que f est la fonction identiquement nulle. Soit $x \in \mathbb{R}$. Comme $f(-x) = f(x)$ (car f est paire) et $f(-x) = -f(x)$ (car f est impaire), alors $f(x) = -f(x)$, ce qui implique $f(x) = 0$. Ceci est vrai quel que soit $x \in \mathbb{R}$; donc f est la fonction nulle. Ainsi $\mathscr{P} \cap \mathscr{I} = \{0_{\mathscr{F}(\mathbb{R},\mathbb{R})}\}$.
2. Montrons $\mathscr{P} + \mathscr{I} = \mathscr{F}(\mathbb{R},\mathbb{R})$.
   Soit $f \in \mathscr{F}(\mathbb{R},\mathbb{R})$. Il s'agit de montrer que f peut s'écrire comme la somme d'une fonction paire et d'une fonction impaire.
   Analyse. Si $f = g + h$, avec $g \in \mathscr{P}$, $h \in \mathscr{I}$, alors pour tout x, d'une part, (a) $f(x) = g(x) + h(x)$, et d'autre part, (b) $f(-x) = g(-x) + h(-x) = g(x) - h(x)$. Par somme et différence de (a) et

(b), on tire que

$$g(x)=\frac{f(x)+f(-x)}{2} \qquad \text{et} \qquad h(x)=\frac{f(x)-f(-x)}{2}.$$

Synthèse. Pour $f\in\mathscr{F}(\mathbb{R},\mathbb{R})$, on définit deux fonctions g,h par $g(x)=\frac{f(x)+f(-x)}{2}$ et $h(x)=\frac{f(x)-f(-x)}{2}$. Alors d'une part $f(x)=g(x)+h(x)$ et d'autre part $g\in\mathscr{P}$ (vérifier $g(-x)=g(x)$) et $h\in\mathscr{I}$ (vérifier $h(-x)=-h(x)$). Bilan : $\mathscr{P}+\mathscr{I}=\mathscr{F}(\mathbb{R},\mathbb{R})$.

En conclusion, $\mathscr{P}$ et $\mathscr{I}$ sont en somme directe dans $\mathscr{F}(\mathbb{R},\mathbb{R})$: $\mathscr{P}\oplus\mathscr{I}=\mathscr{F}(\mathbb{R},\mathbb{R})$. Notez que, comme le prouvent nos calculs, les g et h obtenus sont uniques.

5.3. Sous-espace engendré

Théorème 16. Théorème de structure de l'ensemble des combinaisons linéaires

Soit $\{v_1,\ldots,v_n\}$ un ensemble fini de vecteurs d'un $\mathbb{K}$-espace vectoriel E. Alors :

- L'ensemble des combinaisons linéaires des vecteurs $\{v_1,\ldots,v_n\}$ est un sous-espace vectoriel de E.
- C'est le plus petit sous-espace vectoriel de E (au sens de l'inclusion) contenant les vecteurs $v_1,\ldots,v_n$.

Notation. Ce sous-espace vectoriel est appelé ***sous-espace engendré par*** $v_1,\ldots,v_n$ et est noté $\mathrm{Vect}(v_1,\ldots,v_n)$. On a donc

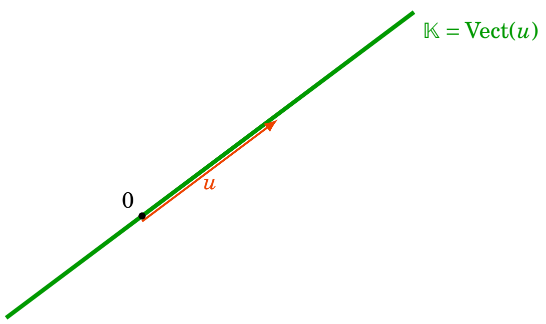
$$\boxed{u\in\mathrm{Vect}(v_1,\ldots,v_n) \quad\Longleftrightarrow\quad \text{il existe } \lambda_1,\ldots,\lambda_n\in\mathbb{K} \quad \text{tels que} \quad u=\lambda_1v_1+\cdots+\lambda_nv_n}$$

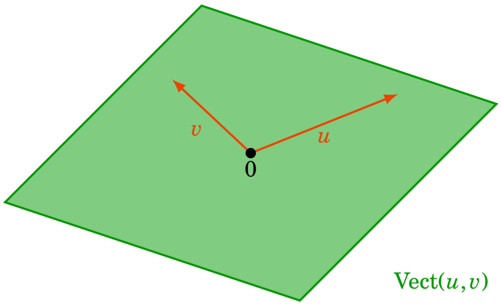
Remarque

- Dire que $\mathrm{Vect}(v_1,\ldots,v_n)$ est le plus petit sous-espace vectoriel de E contenant les vecteurs $v_1,\ldots,v_n$ signifie que si F est un sous-espace vectoriel de E contenant aussi les vecteurs $v_1,\ldots,v_n$ alors $\mathrm{Vect}(v_1,\ldots,v_n)\subset F$.
- Plus généralement, on peut définir le sous-espace vectoriel engendré par une partie $\mathscr{V}$ quelconque (non nécessairement finie) d'un espace vectoriel : $\mathrm{Vect}\,\mathscr{V}$ est le plus petit sous-espace vectoriel contenant $\mathscr{V}$.

Exemple 60

1. E étant un $\mathbb{K}$-espace vectoriel, et u un élément quelconque de E, l'ensemble $\mathrm{Vect}(u)=\{\lambda u\mid\lambda\in\mathbb{K}\}$ est le sous-espace vectoriel de E engendré par u. Il est souvent noté $\mathbb{K}u$. Si u n'est pas le vecteur nul, on parle d'une ***droite vectorielle***.





2. Si u et v sont deux vecteurs de E, alors $\mathrm{Vect}(u,v) = \{\lambda u + \mu v \mid \lambda, \mu \in \mathbb{K}\}$. Si u et v ne sont pas colinéaires, alors $\mathrm{Vect}(u,v)$ est un ***plan vectoriel***.
3. Soient $u = \begin{pmatrix}1\\1\\1\end{pmatrix}$ et $v = \begin{pmatrix}1\\2\\3\end{pmatrix}$ deux vecteurs de $\mathbb{R}^3$. Déterminons $\mathscr{P} = \mathrm{Vect}(u,v)$.

$$\begin{aligned}
\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathrm{Vect}(u,v) \quad &\iff \quad \begin{pmatrix}x\\y\\z\end{pmatrix} = \lambda u + \mu v \quad \text{pour certains } \lambda, \mu \in \mathbb{R} \\
&\iff \quad \begin{pmatrix}x\\y\\z\end{pmatrix} = \lambda \begin{pmatrix}1\\1\\1\end{pmatrix} + \mu \begin{pmatrix}1\\2\\3\end{pmatrix} \\
&\iff \quad \begin{cases} x = \lambda + \mu \\ y = \lambda + 2\mu \\ z = \lambda + 3\mu \end{cases}
\end{aligned}$$

Nous obtenons bien une équation paramétrique du plan $\mathscr{P}$ passant par l'origine et contenant les vecteurs u et v. On sait en trouver une équation cartésienne : $(x - 2y + z = 0)$.

Exemple 61

Soient E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$ et f_0, f_1, f_2 les applications définies par :

$$\forall x \in \mathbb{R} \qquad f_0(x) = 1, \; f_1(x) = x \; \text{ et } \; f_2(x) = x^2.$$

Le sous-espace vectoriel de E engendré par $\{f_0, f_1, f_2\}$ est l'espace vectoriel des fonctions polynômes f de degré inférieur ou égal à 2, c'est-à-dire de la forme $f(x) = ax^2 + bx + c$.

Méthodologie. On peut démontrer qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel de E en montrant que F est égal à l'ensemble des combinaisons linéaires d'un nombre fini de vecteurs de E.

Exemple 62

Est-ce que $F = \{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^3$?
Un triplet de $\mathbb{R}^3$ est élément de F si et seulement si $x = y + z$. Donc u est élément de F si et seulement s'il peut s'écrire $u = (y+z, y, z)$. Or, on a l'égalité

$$(y+z, y, z) = y(1,1,0) + z(1,0,1).$$

Donc F est l'ensemble des combinaisons linéaires de $\{(1,1,0),(1,0,1)\}$. C'est le sous-espace vectoriel engendré par $\{(1,1,0),(1,0,1)\}$: $F = \mathrm{Vect}\{(1,1,0),(1,0,1)\}$. C'est bien un plan vectoriel (un plan passant par l'origine).

Démonstration . Preuve du théorème 16

1. On appelle F l'ensemble des combinaisons linéaires des vecteurs $\{v_1,\dots,v_n\}$.
   (a) $0_E \in F$ car F contient la combinaison linéaire particulière $0v_1+\cdots+0v_n$.
   (b) Si $u,v \in F$ alors il existe $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ tels que $u=\lambda_1 v_1+\cdots+\lambda_n v_n$ et $\mu_1,\dots,\mu_n \in \mathbb{K}$ tels que $v=\mu_1 v_1+\cdots+\mu_n v_n$. On en déduit que $u+v=(\lambda_1+\mu_1)v_1+\cdots+(\lambda_n+\mu_n)v_n$ appartient bien à F.
   (c) De même, $\lambda\cdot u=(\lambda\lambda_1)v_1+\cdots+(\lambda\lambda_n)v_n \in F$.

   Conclusion : F est un sous-espace vectoriel.
2. Si G est un sous-espace vectoriel contenant $\{v_1,\dots,v_n\}$, alors il est stable par combinaison linéaire ; il contient donc toute combinaison linéaire des vecteurs $\{v_1,\dots,v_n\}$. Par conséquent F est inclus dans G : F est le plus petit sous-espace (au sens de l'inclusion) contenant $\{v_1,\dots,v_n\}$.

5.4. Mini-exercices

1. Trouver des sous-espaces vectoriels distincts F et G de $\mathbb{R}^3$ tels que
   (a) $F+G=\mathbb{R}^3$ et $F\cap G\neq\{0\}$;
   (b) $F+G\neq\mathbb{R}^3$ et $F\cap G=\{0\}$;
   (c) $F+G=\mathbb{R}^3$ et $F\cap G=\{0\}$;
   (d) $F+G\neq\mathbb{R}^3$ et $F\cap G\neq\{0\}$.
2. Soient $F=\{(x,y,z)\in\mathbb{R}^3 \mid x+y+z=0\}$ et $G=\mathrm{Vect}\{(1,1,1)\}\subset\mathbb{R}^3$.
   (a) Montrer que F est un espace vectoriel. Trouver deux vecteurs u,v tels que $F=\mathrm{Vect}(u,v)$.
   (b) Calculer $F\cap G$ et montrer que $F+G=\mathbb{R}^3$. Que conclure ?
3. Soient $A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$, $C=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $D=\begin{pmatrix}0&0\\1&0\end{pmatrix}$ des matrices de $M_2(\mathbb{R})$.
   (a) Quel est l'espace vectoriel F engendré par A et B ? Idem avec G engendré par C et D.
   (b) Calculer $F\cap G$. Montrer que $F+G=M_2(\mathbb{R})$. Conclure.

6. Application linéaire (début)

6.1. Définition

Nous avons déjà rencontré la notion d'application linéaire dans le cas $f:\mathbb{R}^p\longrightarrow\mathbb{R}^n$ (voir le chapitre « L'espace vectoriel $\mathbb{R}^n$ »). Cette notion se généralise à des espaces vectoriels quelconques.

Définition 27

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application f de E dans F est une ***application linéaire*** si elle satisfait aux deux conditions suivantes :

1. $f(u+v)=f(u)+f(v)$, pour tous $u,v\in E$;
2. $f(\lambda\cdot u)=\lambda\cdot f(u)$, pour tout $u\in E$ et tout $\lambda\in\mathbb{K}$.

Autrement dit : une application est linéaire si elle « respecte » les deux lois d'un espace vectoriel.

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

6.2. Premiers exemples

Exemple 63

L'application f définie par

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x,y,z) & \mapsto & (-2x, y+3z) \end{array}$$

est une application linéaire. En effet, soient $u=(x,y,z)$ et $v=(x',y',z')$ deux éléments de $\mathbb{R}^3$ et λ un réel.

$$\begin{array}{rcl} f(u+v) & = & f(x+x', y+y', z+z') \\ & = & \big(-2(x+x'), y+y'+3(z+z')\big) \\ & = & (-2x, y+3z)+(-2x', y'+3z') \\ & = & f(u)+f(v) \end{array} \quad \text{et} \quad \begin{array}{rcl} f(\lambda\cdot u) & = & f(\lambda x, \lambda y, \lambda z) \\ & = & (-2\lambda x, \lambda y+3\lambda z) \\ & = & \lambda\cdot(-2x, y+3z) \\ & = & \lambda\cdot f(u) \end{array}$$

Toutes les applications ne sont pas des applications linéaires !

Exemple 64

Soit $f:\mathbb{R}\to\mathbb{R}$ l'application définie par $f(x)=x^2$. On a $f(1)=1$ et $f(2)=4$. Donc $f(2)\neq 2\cdot f(1)$. Ce qui fait que l'on n'a pas l'égalité $f(\lambda x)=\lambda f(x)$ pour un certain choix de λ, x. Donc f n'est pas linéaire. Notez que l'on n'a pas non plus $f(x+x')=f(x)+f(x')$ dès que $xx'\neq 0$.

Voici d'autres exemples d'applications linéaires :

1. Pour une matrice fixée $A\in M_{n,p}(\mathbb{R})$, l'application $f:\mathbb{R}^p\longrightarrow\mathbb{R}^n$ définie par
$$f(X)=AX$$
est une application linéaire.
2. L'***application nulle***, notée $0_{\mathcal{L}(E,F)}$:
$$f:E\longrightarrow F \qquad f(u)=0_F \qquad \text{pour tout } u\in E.$$
3. L'***application identité***, notée id_E :
$$f:E\longrightarrow E \qquad f(u)=u \qquad \text{pour tout } u\in E.$$

6.3. Premières propriétés

Proposition 16

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Si f est une application linéaire de E dans F, alors :

- $f(0_E)=0_F$,
- $f(-u)=-f(u)$, pour tout $u\in E$.

Démonstration

Il suffit d'appliquer la définition de la linéarité avec $\lambda=0$, puis avec $\lambda=-1$.

Pour démontrer qu'une application est linéaire, on peut aussi utiliser une propriété plus « concentrée », donnée par la caractérisation suivante :

Proposition 17. Caractérisation d'une application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels et f une application de E dans F. L'application f est linéaire si et seulement si, pour tous vecteurs u et v de E et pour tous scalaires λ et μ de $\mathbb{K}$,

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

Plus généralement, une application linéaire f préserve les combinaisons linéaires : pour tous $\lambda_1,\ldots,\lambda_n\in\mathbb{K}$ et tous $v_1,\ldots,v_n\in E$, on a

$$f(\lambda_1 v_1+\cdots+\lambda_n v_n)=\lambda_1 f(v_1)+\cdots+\lambda_n f(v_n).$$

Démonstration

- Soit f une application linéaire de E dans F. Soient $u,v\in E$, $\lambda,\mu\in\mathbb{K}$. En utilisant les deux axiomes de la définition, on a

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

- Montrons la réciproque. Soit $f:E\to F$ une application telle que $f(\lambda u+\mu v)=\lambda f(u)+\mu f(v)$ (pour tous $u,v\in E$, $\lambda,\mu\in\mathbb{K}$). Alors, d'une part $f(u+v)=f(u)+f(v)$ (en considérant le cas particulier où $\lambda=\mu=1$), et d'autre part $f(\lambda u)=\lambda f(u)$ (cas particulier où $\mu=0$).

Vocabulaire.
Soient E et F deux $\mathbb{K}$-espaces vectoriels.

- Une application linéaire de E dans F est aussi appelée ***morphisme*** ou ***homomorphisme*** d'espaces vectoriels. L'ensemble des applications linéaires de E dans F est noté $\mathscr{L}(E,F)$.
- Une application linéaire de E dans E est appelée ***endomorphisme*** de E. L'ensemble des endomorphismes de E est noté $\mathscr{L}(E)$.

6.4. Mini-exercices

Montrer que les applications suivantes $f_i:\mathbb{R}^2\to\mathbb{R}^2$ sont linéaires. Caractériser géométriquement ces applications et faire un dessin.

1. $f_1(x,y)=(-x,-y)$;
2. $f_2(x,y)=(3x,3y)$;
3. $f_3(x,y)=(x,-y)$;
4. $f_4(x,y)=(-x,y)$;
5. $f_5(x,y)=\left(\frac{\sqrt{3}}{2}x-\frac{1}{2}y,\frac{1}{2}x+\frac{\sqrt{3}}{2}y\right)$.

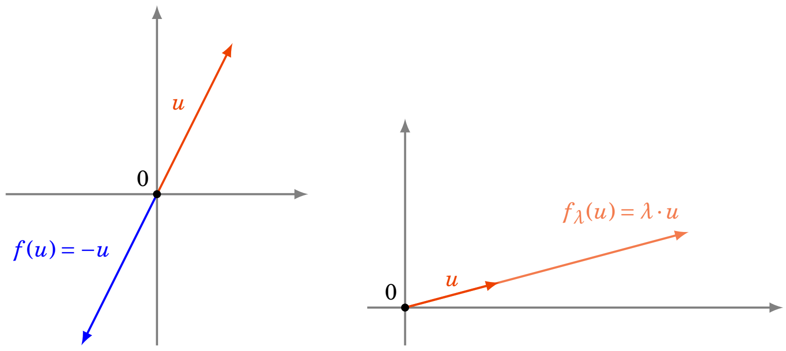
7. Application linéaire (milieu)

7.1. Exemples géométriques

Symétrie centrale.
Soient E un $\mathbb{K}$-espace vectoriel. On définit l'application f par :

$$\begin{array}{rcl} f:E & \to & E \\ u & \mapsto & -u \end{array}$$

f est linéaire et s'appelle la ***symétrie centrale*** par rapport à l'origine 0_E.



Homothétie.
Soient E un $\mathbb{K}$-espace vectoriel et $\lambda \in \mathbb{K}$. On définit l'application f_λ par :

$$\begin{array}{rcl} f_\lambda : E & \to & E \\ u & \mapsto & \lambda u \end{array}$$

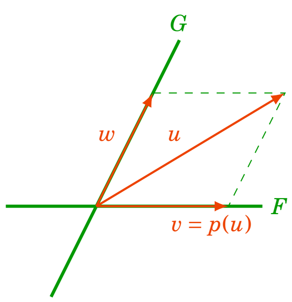
f_λ est linéaire. f_λ est appelée ***homothétie*** de rapport λ.
Cas particuliers notables :

- $\lambda = 1$, f_λ est l'application identité ;
- $\lambda = 0$, f_λ est l'application nulle ;
- $\lambda = -1$, on retrouve la symétrie centrale.

Preuve que f_λ est une application linéaire :

$$f_\lambda(\alpha u + \beta v) = \lambda(\alpha u + \beta v) = \alpha(\lambda u) + \beta(\lambda v) = \alpha f_\lambda(u) + \beta f_\lambda(v).$$

Projection.
Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels supplémentaires dans E, c'est-à-dire $E = F \oplus G$. Tout vecteur u de E s'écrit de façon unique $u = v + w$ avec $v \in F$ et $w \in G$. La ***projection*** sur F parallèlement à G est l'application $p : E \to E$ définie par $p(u) = v$.



- Une projection est une application linéaire.
  En effet, soient $u, u' \in E$, $\lambda, \mu \in \mathbb{K}$. On décompose u et u' en utilisant que $E = F \oplus G$: $u = v + w$, $u' = v' + w'$ avec $v, v' \in F$, $w, w' \in G$. Commençons par écrire

$$\lambda u + \mu u' = \lambda(v + w) + \mu(v' + w') = (\lambda v + \mu v') + (\lambda w + \mu w').$$

  Comme F et G sont des un sous-espaces vectoriels de E, alors $\lambda v + \mu v' \in F$ et $\lambda w + \mu w' \in G$. Ainsi :

$$p(\lambda u + \mu u') = \lambda v + \mu v' = \lambda p(u) + \mu p(u').$$

– Une projection p vérifie l'égalité $p^2 = p$.
Note : $p^2 = p$ signifie $p \circ p = p$, c'est-à-dire pour tout $u \in E$: $p\big(p(u)\big) = p(u)$. Il s'agit juste de remarquer que si $v \in F$ alors $p(v) = v$ (car $v = v + 0$, avec $v \in F$ et $0 \in G$). Maintenant, pour $u \in E$, on a $u = v + w$ avec $v \in F$ et $w \in G$. Par définition $p(u) = v$. Mais alors $p\big(p(u)\big) = p(v) = v$. Bilan : $p \circ p(u) = v = p(u)$. Donc $p \circ p = p$.

Exemple 65

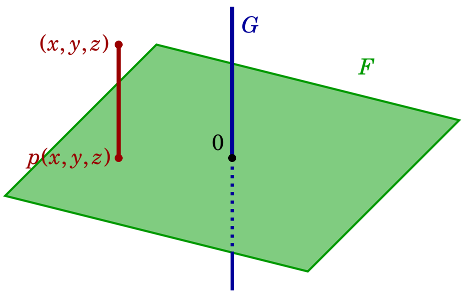
Nous avons vu que les sous-espaces vectoriels F et G de $\mathbb{R}^3$ définis par

$$F = \big\{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\big\} \qquad \text{et} \qquad G = \big\{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\big\}$$

sont supplémentaires dans $\mathbb{R}^3$: $\mathbb{R}^3 = F \oplus G$ (exemple 58). Nous avions vu que la décomposition s'écrivait :

$$(x,y,z) = (y+z,y,z) + (x-y-z,0,0).$$

Si p est la projection sur F parallèlement à G, alors on a $p(x,y,z) = (y+z,y,z)$.



Exemple 66

Nous avons vu dans l'exemple 59 que l'ensemble des fonctions paires $\mathscr{P}$ et l'ensemble des fonctions impaires $\mathscr{I}$ sont des sous-espaces vectoriels supplémentaires dans $\mathscr{F}(\mathbb{R},\mathbb{R})$. Notons p la projection sur $\mathscr{P}$ parallèlement à $\mathscr{I}$. Si f est un élément de $\mathscr{F}(\mathbb{R},\mathbb{R})$, on a $p(f) = g$ où

$$\begin{array}{rcl} g : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x)+f(-x)}{2}. \end{array}$$

7.2. Autres exemples

1. La **dérivation**. Soient $E = \mathscr{C}^1(\mathbb{R},\mathbb{R})$ l'espace vectoriel des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ dérivables avec f' continue et $F = \mathscr{C}^0(\mathbb{R},\mathbb{R})$ l'espace vectoriel des fonctions continues. Soit

$$\begin{array}{rcl} d : \mathscr{C}^1(\mathbb{R},\mathbb{R}) & \longrightarrow & \mathscr{C}^0(\mathbb{R},\mathbb{R}) \\ f & \longmapsto & f' \end{array}$$

Alors d est une application linéaire, car $(\lambda f + \mu g)' = \lambda f' + \mu g'$ et donc $d(\lambda f + \mu g) = \lambda d(f) + \mu d(g)$.

2. L'**intégration**. Soient $E = \mathscr{C}^0(\mathbb{R},\mathbb{R})$ et $F = \mathscr{C}^1(\mathbb{R},\mathbb{R})$. Soit

$$\begin{array}{rcl} I : \mathscr{C}^0(\mathbb{R},\mathbb{R}) & \longrightarrow & \mathscr{C}^1(\mathbb{R},\mathbb{R}) \\ f(x) & \longmapsto & \int_0^x f(t)\,dt \end{array}$$

L'application I est linéaire car $\int_0^x \big(\lambda f(t) + \mu g(t)\big)\,dt = \lambda \int_0^x f(t)\,dt + \mu \int_0^x g(t)\,dt$ pour toutes fonctions f et g et pour tous $\lambda, \mu \in \mathbb{R}$.

3. Avec les **polynômes**.
Soit $E = \mathbb{R}_n[X]$ l'espace vectoriel des polynômes de degré $\leqslant n$. Soit $F = \mathbb{R}_{n+1}[X]$ et soit

$$\begin{array}{rccc} f: & E & \longrightarrow & F \\ & P(X) & \longmapsto & XP(X) \end{array}$$

Autrement dit, si $P(X) = a_nX^n + \cdots + a_1X + a_0$, alors $f(P(X)) = a_nX^{n+1} + \cdots + a_1X^2 + a_0X$.
C'est une application linéaire : $f(\lambda P(X) + \mu Q(X)) = \lambda XP(X) + \mu XQ(X) = \lambda f(P(X)) + \mu f(Q(X))$.

4. La **transposition**.
Considérons l'application T de $M_n(\mathbb{K})$ dans $M_n(\mathbb{K})$ donnée par la transposition :

$$\begin{array}{rcc} T : M_n(\mathbb{K}) & \longrightarrow & M_n(\mathbb{K}) \\ A & \longmapsto & A^T \end{array}$$

T est linéaire, car on sait que pour toutes matrices $A, B \in M_n(\mathbb{K})$ et tous scalaires $\lambda, \mu \in \mathbb{K}$:

$$(\lambda A + \mu B)^T = (\lambda A)^T + (\mu B)^T = \lambda A^T + \mu B^T.$$

5. La **trace**.

$$\begin{array}{rcc} \mathrm{tr} : M_n(\mathbb{K}) & \longrightarrow & \mathbb{K} \\ A & \longmapsto & \mathrm{tr}\, A \end{array}$$

est une application linéaire car $\mathrm{tr}(\lambda A + \mu B) = \lambda\, \mathrm{tr}\, A + \mu\, \mathrm{tr}\, B$.

7.3. Mini-exercices

1. Les applications suivantes sont-elles linéaires ?
   (a) $\mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto 3x - 2$
   (b) $\mathbb{R}^4 \longrightarrow \mathbb{R}, \quad (x, y, x', y') \longmapsto x \cdot x' + y \cdot y'$
   (c) $\mathscr{C}^0(\mathbb{R}, \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \longmapsto f(1)$
   (d) $\mathscr{C}^1(\mathbb{R}, \mathbb{R}) \longrightarrow \mathscr{C}^0(\mathbb{R}, \mathbb{R}), \quad f \longmapsto f' + f$
   (e) $\mathscr{C}^0([0,1], \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \longmapsto \int_0^1 |f(t)|\, dt$
   (f) $\mathscr{C}^0([0,1], \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \longmapsto \max_{x \in [0,1]} f(x)$
   (g) $\mathbb{R}_3[X] \longrightarrow \mathbb{R}_3[X], \quad P(X) \longmapsto P(X+1) - P(0)$
2. Soient $f, g : M_n(\mathbb{R}) \longrightarrow M_n(\mathbb{R})$ définies par $A \longmapsto \frac{A + A^T}{2}$ et $A \longmapsto \frac{A - A^T}{2}$. Montrer que f et g sont des applications linéaires. Montrer que $f(A)$ est une matrice symétrique, $g(A)$ une matrice antisymétrique et que $A = f(A) + g(A)$. En déduire que les matrices symétriques et les matrices antisymétriques sont en somme directe dans $M_n(\mathbb{R})$. Caractériser géométriquement f et g.

8. Application linéaire (fin)

8.1. Image d'une application linéaire

Commençons par des rappels. Soient E et F deux ensembles et f une application de E dans F. Soit A un sous-ensemble de E. L'ensemble des images par f des éléments de A, appelé ***image directe*** de A par f, est noté $f(A)$. C'est un sous-ensemble de F. On a par définition :

$$f(A) = \big\{ f(x) \mid x \in A \big\}.$$

Dans toute la suite, E et F désigneront des $\mathbb{K}$-espaces vectoriels et $f : E \to F$ sera une application linéaire.
$f(E)$ s'appelle l'***image*** de l'application linéaire f et est noté $\operatorname{Im} f$.

Proposition 18. Structure de l'image d'un sous-espace vectoriel

1. Si E' est un sous-espace vectoriel de E, alors $f(E')$ est un sous-espace vectoriel de F.
2. En particulier, $\operatorname{Im} f$ est un sous-espace vectoriel de F.

Remarque

On a par définition de l'image directe $f(E)$:

$$f \text{ est surjective si et seulement si } \operatorname{Im} f = F.$$

Démonstration

Tout d'abord, comme $0_E \in E'$ alors $0_F = f(0_E) \in f(E')$. Ensuite on montre que pour tout couple (y_1, y_2) d'éléments de $f(E')$ et pour tous scalaires λ, μ, l'élément $\lambda y_1 + \mu y_2$ appartient à $f(E')$. En effet :

$$y_1 \in f(E') \Longleftrightarrow \exists x_1 \in E', f(x_1) = y_1$$
$$y_2 \in f(E') \Longleftrightarrow \exists x_2 \in E', f(x_2) = y_2.$$

Comme f est linéaire, on a

$$\lambda y_1 + \mu y_2 = \lambda f(x_1) + \mu f(x_2) = f(\lambda x_1 + \mu x_2).$$

Or $\lambda x_1 + \mu x_2$ est un élément de E', car E' est un sous-espace vectoriel de E, donc $\lambda y_1 + \mu y_2$ est bien un élément de $f(E')$.

8.2. Noyau d'une application linéaire

Définition 28. Définition du noyau

Soient E et F deux $\mathbb{K}$-espaces vectoriels et f une application linéaire de E dans F. Le ***noyau*** de f, noté $\operatorname{Ker}(f)$, est l'ensemble des éléments de E dont l'image est 0_F :

$$\boxed{\operatorname{Ker}(f) = \{x \in E \mid f(x) = 0_F\}}$$

Autrement dit, le noyau est l'image réciproque du vecteur nul de l'espace d'arrivée : $\operatorname{Ker}(f) = f^{-1}\{0_F\}$.

Proposition 19

Soient E et F deux $\mathbb{K}$-espaces vectoriels et f une application linéaire de E dans F. Le noyau de f est un sous-espace vectoriel de E.

Démonstration

$\mathrm{Ker}(f)$ est non vide car $f(0_E) = 0_F$ donc $0_E \in \mathrm{Ker}(f)$. Soient $x_1, x_2 \in \mathrm{Ker}(f)$ et $\lambda, \mu \in \mathbb{K}$. Montrons que $\lambda x_1 + \mu x_2$ est un élément de $\mathrm{Ker}(f)$. On a, en utilisant la linéarité de f et le fait que x_1 et x_2 sont des éléments de $\mathrm{Ker}(f)$: $f(\lambda x_1 + \mu x_2) = \lambda f(x_1) + \mu f(x_2) = \lambda 0_F + \mu 0_F = 0_F$.

Exemple 67

Reprenons l'exemple de l'application linéaire f définie par

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x,y,z) & \mapsto & (-2x, y+3z) \end{array}$$

- Calculons le noyau $\mathrm{Ker}(f)$.

$$\begin{array}{rcl} (x,y,z) \in \mathrm{Ker}(f) & \iff & f(x,y,z) = (0,0) \\ & \iff & (-2x, y+3z) = (0,0) \\ & \iff & \left\{ \begin{array}{rcl} -2x & = & 0 \\ y+3z & = & 0 \end{array} \right. \\ & \iff & (x,y,z) = (0,-3z,z), \quad z \in \mathbb{R} \end{array}$$

  Donc $\mathrm{Ker}(f) = \big\{(0,-3z,z) \mid z \in \mathbb{R}\big\}$. Autrement dit, $\mathrm{Ker}(f) = \mathrm{Vect}\big\{(0,-3,1)\big\}$: c'est une droite vectorielle.
- Calculons l'image de f. Fixons $(x', y') \in \mathbb{R}^2$.

$$\begin{array}{rcl} (x',y') = f(x,y,z) & \iff & (-2x, y+3z) = (x',y') \\ & \iff & \left\{ \begin{array}{rcl} -2x & = & x' \\ y+3z & = & y' \end{array} \right. \end{array}$$

  On peut prendre par exemple $x = -\frac{x'}{2}$, $y' = y$, $z = 0$. Conclusion : pour n'importe quel $(x',y') \in \mathbb{R}^2$, on a $f(-\frac{x'}{2}, y', 0) = (x', y')$. Donc $\mathrm{Im}(f) = \mathbb{R}^2$, et f est surjective.

Exemple 68

Soit $A \in M_{n,p}(\mathbb{R})$. Soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ l'application linéaire définie par $f(X) = AX$. Alors $\mathrm{Ker}(f) = \big\{X \in \mathbb{R}^p \mid AX = 0\big\}$: c'est donc l'ensemble des $X \in \mathbb{R}^p$ solutions du système linéaire homogène $AX = 0$. On verra plus tard que $\mathrm{Im}(f)$ est l'espace engendré par les colonnes de la matrice A.

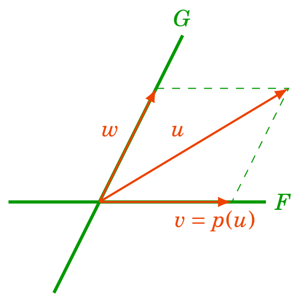
Le noyau fournit une nouvelle façon d'obtenir des sous-espaces vectoriels.

Exemple 69

Un plan $\mathscr{P}$ passant par l'origine, d'équation $(ax + by + cz = 0)$, est un sous-espace vectoriel de $\mathbb{R}^3$. En effet, soit $f : \mathbb{R}^3 \to \mathbb{R}$ l'application définie par $f(x,y,z) = ax + by + cz$. Il est facile de vérifier que f est linéaire, de sorte que $\mathrm{Ker}\, f = \big\{(x,y,z) \in \mathbb{R}^3 \mid ax + by + cz = 0\big\} = \mathscr{P}$ est un sous-espace vectoriel.

Exemple 70

Soient E un $\mathbb{K}$-espace vectoriel, F et G deux sous-espaces vectoriels de E, supplémentaires : $E = F \oplus G$. Soit p la projection sur F parallèlement à G. Déterminons le noyau et l'image de p.



Un vecteur u de E s'écrit d'une manière unique $u = v + w$ avec $v \in F$ et $w \in G$ et par définition $p(u) = v$.

- $\mathrm{Ker}(p) = G$: le noyau de p est l'ensemble des vecteurs u de E tels que $v = 0$, c'est donc G.
- $\mathrm{Im}(p) = F$. Il est immédiat que $\mathrm{Im}(p) \subset F$. Réciproquement, si $u \in F$ alors $p(u) = u$, donc $F \subset \mathrm{Im}(p)$.

Conclusion :

$$\mathrm{Ker}(p) = G \qquad \text{et} \qquad \mathrm{Im}(p) = F.$$

Théorème 17. Caractérisation des applications linéaires injectives

Soient E et F deux $\mathbb{K}$-espaces vectoriels et f une application linéaire de E dans F. Alors :

$$\boxed{f \text{ injective} \quad \iff \quad \mathrm{Ker}(f) = \{0_E\}}$$

Autrement dit, f est injective si et seulement si son noyau ne contient que le vecteur nul. En particulier, pour montrer que f est injective, il suffit de vérifier que :

$$\text{si } f(x) = 0_F \text{ alors } x = 0_E.$$

Démonstration

- Supposons que f soit injective et montrons que $\mathrm{Ker}(f) = \{0_E\}$. Soit x un élément de $\mathrm{Ker}(f)$. On a $f(x) = 0_F$. Or, comme f est linéaire, on a aussi $f(0_E) = 0_F$. De l'égalité $f(x) = f(0_E)$, on déduit $x = 0_E$ car f est injective. Donc $\mathrm{Ker}(f) = \{0_E\}$.
- Réciproquement, supposons maintenant que $\mathrm{Ker}(f) = \{0_E\}$. Soient x et y deux éléments de E tels que $f(x) = f(y)$. On a donc $f(x) - f(y) = 0_F$. Comme f est linéaire, on en déduit $f(x-y) = 0_F$, c'est-à-dire $x - y$ est un élément de $\mathrm{Ker}(f)$. Donc $x - y = 0_E$, soit $x = y$.

Exemple 71

Considérons, pour $n \geqslant 1$, l'application linéaire

$$\begin{aligned} f : \mathbb{R}_n[X] &\longrightarrow \mathbb{R}_{n+1}[X] \\ P(X) &\longmapsto X \cdot P(X). \end{aligned}$$

Étudions d'abord le noyau de f : soit $P(X) = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{R}_n[X]$ tel que $X \cdot P(X) = 0$.

Alors
$$a_n X^{n+1} + \cdots + a_1 X^2 + a_0 X = 0.$$
Ainsi, $a_i = 0$ pour tout $i \in \{0, \ldots, n\}$ et donc $P(X) = 0$. Le noyau de f est donc nul : $\mathrm{Ker}(f) = \{0\}$. L'espace $\mathrm{Im}(f)$ est l'ensemble des polynômes de $\mathbb{R}_{n+1}[X]$ sans terme constant : $\mathrm{Im}(f) = \mathrm{Vect}\{X, X^2, \ldots, X^{n+1}\}$.
Conclusion : f est injective, mais n'est pas surjective.

8.3. L'espace vectoriel $\mathscr{L}(E,F)$

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Remarquons tout d'abord que, similairement à l'exemple 44, l'ensemble des applications de E dans F, noté $\mathscr{F}(E,F)$, peut être muni d'une loi de composition interne + et d'une loi de composition externe, définies de la façon suivante : f, g étant deux éléments de $\mathscr{F}(E,F)$, et λ étant un élément de $\mathbb{K}$, pour tout vecteur u de E,
$$(f+g)(u) = f(u) + g(u) \qquad \text{et} \qquad (\lambda \cdot f)(u) = \lambda f(u).$$

Proposition 20

L'ensemble des applications linéaires entre deux $\mathbb{K}$-espaces vectoriels E et F, noté $\mathscr{L}(E,F)$, muni des deux lois définies précédemment, est un $\mathbb{K}$-espace vectoriel.

Démonstration

L'ensemble $\mathscr{L}(E,F)$ est inclus dans le $\mathbb{K}$-espace vectoriel $\mathscr{F}(E,F)$. Pour montrer que $\mathscr{L}(E,F)$ est un $\mathbb{K}$-espace vectoriel, il suffit donc de montrer que $\mathscr{L}(E,F)$ est un sous-espace vectoriel de $\mathscr{F}(E,F)$:

- Tout d'abord, l'application nulle appartient à $\mathscr{L}(E,F)$.
- Soient $f, g \in \mathscr{L}(E,F)$, et montrons que $f+g$ est linéaire. Pour tous vecteurs u et v de E et pour tous scalaires α, β de $\mathbb{K}$,
$$\begin{array}{rcll} (f+g)(\alpha u+\beta v) & = & f(\alpha u+\beta v)+g(\alpha u+\beta v) & \text{(définition de } f+g\text{)} \\ & = & \alpha f(u)+\beta f(v)+\alpha g(u)+\beta g(v) & \text{(linéarité de } f \text{ et } g\text{)} \\ & = & \alpha\big(f(u)+g(u)\big)+\beta\big(f(v)+g(v)\big) & \text{(propriétés des lois de } F\text{)} \\ & = & \alpha(f+g)(u)+\beta(f+g)(v) & \text{(définition de } f+g\text{)} \end{array}$$
  $f+g$ est donc linéaire et $\mathscr{L}(E,F)$ est stable pour l'addition.
- Soient $f \in \mathscr{L}(E,F)$, $\lambda \in \mathbb{K}$, et montrons que λf est linéaire.
$$\begin{array}{rcll} (\lambda f)(\alpha u+\beta v) & = & \lambda f(\alpha u+\beta v) & \text{(définition de } \lambda f\text{)} \\ & = & \lambda\big(\alpha f(u)+\beta f(v)\big) & \text{(linéarité de } f\text{)} \\ & = & \alpha\lambda f(u)+\beta\lambda f(v) & \text{(propriétés des lois de } F\text{)} \\ & = & \alpha(\lambda f)(u)+\beta(\lambda f)(v) & \text{(définition de } \lambda f\text{)} \end{array}$$
  λf est donc linéaire et $\mathscr{L}(E,F)$ est stable pour la loi externe.

$\mathscr{L}(E,F)$ est donc un sous-espace vectoriel de $\mathscr{F}(E,F)$.

En particulier, $\mathscr{L}(E)$ est un sous-espace vectoriel de $\mathscr{F}(E,E)$.

8.4. Composition et inverse d'applications linéaires

Proposition 21. Composée de deux applications linéaires

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels, f une application linéaire de E dans F et g une application linéaire de F dans G. Alors $g \circ f$ est une application linéaire de E dans G.

Remarque

En particulier, le composé de deux endomorphismes de E est un endomorphisme de E. Autrement dit, $\circ$ est une loi de composition interne sur $\mathscr{L}(E)$.

Démonstration

Soient u et v deux vecteurs de E, et α et β deux éléments de $\mathbb{K}$. Alors :

$$\begin{aligned}
(g \circ f)(\alpha u + \beta v) &= g\big(f(\alpha u + \beta v)\big) && \text{(définition de } g \circ f\text{)} \\
&= g\big(\alpha f(u) + \beta f(v)\big) && \text{(linéarité de } f\text{)} \\
&= \alpha g(f(u)) + \beta g(f(v)) && \text{(linéarité de } g\text{)} \\
&= \alpha (g \circ f)(u) + \beta (g \circ f)(v) && \text{(définition de } g \circ f\text{)}
\end{aligned}$$

La composition des applications linéaires se comporte bien :

$$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2 \qquad (g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f \qquad (\lambda g) \circ f = g \circ (\lambda f) = \lambda (g \circ f)$$

Vocabulaire.
Soient E et F deux $\mathbb{K}$-espaces vectoriels.

- Une application linéaire ***bijective*** de E sur F est appelée ***isomorphisme*** d'espaces vectoriels. Les deux espaces vectoriels E et F sont alors dits ***isomorphes***.
- Un endomorphisme bijectif de E (c'est-à-dire une application linéaire bijective de E dans E) est appelé ***automorphisme*** de E. L'ensemble des automorphismes de E est noté $GL(E)$.

Proposition 22. Linéarité de l'application réciproque d'un isomorphisme

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Si f est un isomorphisme de E sur F, alors f^{-1} est un isomorphisme de F sur E.

Démonstration

Comme f est une application bijective de E sur F, alors f^{-1} est une application bijective de F sur E. Il reste donc à prouver que f^{-1} est bien linéaire. Soient u' et v' deux vecteurs de F et soient α et β deux éléments de $\mathbb{K}$. On pose $f^{-1}(u') = u$ et $f^{-1}(v') = v$, et on a alors $f(u) = u'$ et $f(v) = v'$. Comme f est linéaire, on a

$$f^{-1}(\alpha u' + \beta v') = f^{-1}\big(\alpha f(u) + \beta f(v)\big) = f^{-1}\big(f(\alpha u + \beta v)\big) = \alpha u + \beta v$$

car $f^{-1} \circ f = \mathrm{id}_E$ (où id_E désigne l'application identité de E dans E). Ainsi

$$f^{-1}(\alpha u' + \beta v') = \alpha f^{-1}(u') + \beta f^{-1}(v'),$$

et f^{-1} est donc linéaire.

Exemple 72

Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (2x+3y, x+y)$. Il est facile de prouver que f est linéaire. Pour prouver que f est bijective, on pourrait calculer son noyau et son image. Mais ici nous allons calculer directement son inverse : on cherche à résoudre $f(x,y) = (x',y')$. Cela correspond à l'équation $(2x+3y, x+y) = (x',y')$ qui est un système linéaire à deux équations et deux inconnues. On trouve $(x,y) = (-x'+3y', x'-2y')$. On pose donc $f^{-1}(x',y') = (-x'+3y', x'-2y')$. On vérifie aisément que f^{-1} est l'inverse de f, et on remarque que f^{-1} est une application linéaire.

Exemple 73

Plus généralement, soit $f : \mathbb{R}^n \to \mathbb{R}^n$ l'application linéaire définie par $f(X) = AX$ (où A est une matrice de $M_n(\mathbb{R})$). Si la matrice A est inversible, alors f^{-1} est une application linéaire bijective et est définie par $f^{-1}(X) = A^{-1}X$.
Dans l'exemple précédent,

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \qquad A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix}.$$

8.5. Mini-exercices

1. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (-x, y+z, 2z)$. Montrer que f est une application linéaire. Calculer $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$. f admet-elle un inverse ? Même question avec $f(x,y,z) = (x-y, x+y, y)$.
2. Soient E un espace vectoriel, et F, G deux sous-espaces tels que $E = F \oplus G$. Chaque $u \in E$ se décompose de manière unique $u = v + w$ avec $v \in F$, $w \in G$. La ***symétrie*** par rapport à F parallèlement à G est l'application $s : E \to E$ définie par $s(u) = v - w$. Faire un dessin. Montrer que s est une application linéaire. Montrer que $s^2 = \mathrm{id}_E$. Calculer $\mathrm{Ker}(s)$ et $\mathrm{Im}(s)$. s admet-elle un inverse ?
3. Soit $f : \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ définie par $P(X) \mapsto P''(X)$ (où P'' désigne la dérivée seconde). Montrer que f est une application linéaire. Calculer $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$. f admet-elle un inverse ?

Auteurs

- D'après un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties d'un cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- et un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- mixés et révisés par Arnaud Bodin, relu par Vianney Combet.

Exo7

5 Dimension finie

Les espaces vectoriels qui sont engendrés par un nombre fini de vecteurs sont appelés espaces vectoriels de dimension finie. Pour ces espaces, nous allons voir comment calculer une base, c'est-à-dire une famille minimale de vecteurs qui engendrent tout l'espace. Le nombre de vecteurs dans une base s'appelle la dimension et nous verrons comment calculer la dimension des espaces et des sous-espaces.

1. Famille libre

1.1. Combinaison linéaire (rappel)

Soit E un $\mathbb{K}$-espace vectoriel.

Définition 29

Soient $v_1, v_2, \ldots, v_p$, $p \geqslant 1$ vecteurs d'un espace vectoriel E. Tout vecteur de la forme

$$u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p$$

(où $\lambda_1, \lambda_2, \ldots, \lambda_p$ sont des éléments de $\mathbb{K}$) est appelé ***combinaison linéaire*** des vecteurs $v_1, v_2, \ldots, v_p$. Les scalaires $\lambda_1, \lambda_2, \ldots, \lambda_p$ sont appelés ***coefficients*** de la combinaison linéaire.

1.2. Définitions

Définition 30

Une famille $\{v_1, v_2, \ldots, v_p\}$ de E est une ***famille libre*** ou ***linéairement indépendante*** si toute combinaison linéaire nulle

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = 0$$

est telle que tous ses coefficients sont nuls, c'est-à-dire $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$.

Dans le cas contraire, c'est-à-dire s'il existe une combinaison linéaire nulle à coefficients non tous nuls, on dit que la famille est ***liée*** ou ***linéairement dépendante***. Une telle combinaison linéaire s'appelle alors une ***relation de dépendance linéaire*** entre les v_j.

1.3. Premiers exemples

Pour des vecteurs de $\mathbb{R}^n$, décider si une famille $\{v_1,\dots,v_p\}$ est libre ou liée revient à résoudre un système linéaire.

Exemple 74

Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$, considérons la famille

$$\left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix},\begin{pmatrix}2\\1\\0\end{pmatrix}\right\}.$$

On souhaite déterminer si elle est libre ou liée. On cherche des scalaires $(\lambda_1,\lambda_2,\lambda_3)$ tels que

$$\lambda_1\left(\begin{smallmatrix}1\\2\\3\end{smallmatrix}\right)+\lambda_2\left(\begin{smallmatrix}4\\5\\6\end{smallmatrix}\right)+\lambda_3\left(\begin{smallmatrix}2\\1\\0\end{smallmatrix}\right)=\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)$$

ce qui équivaut au système :

$$\begin{cases}\lambda_1 + 4\lambda_2 + 2\lambda_3 = 0\\ 2\lambda_1 + 5\lambda_2 + \lambda_3 = 0\\ 3\lambda_1 + 6\lambda_2 = 0\end{cases}$$

On calcule (voir un peu plus bas) que ce système est équivalent à :

$$\begin{cases}\lambda_1 - 2\lambda_3 = 0\\ \lambda_2 + \lambda_3 = 0\end{cases}$$

Ce système a une infinité de solutions et en prenant par exemple $\lambda_3=1$ on obtient $\lambda_1=2$ et $\lambda_2=-1$, ce qui fait que

$$2\begin{pmatrix}1\\2\\3\end{pmatrix}-\begin{pmatrix}4\\5\\6\end{pmatrix}+\begin{pmatrix}2\\1\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

La famille

$$\left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix},\begin{pmatrix}2\\1\\0\end{pmatrix}\right\}$$

est donc une famille liée.

Voici les calculs de la réduction de Gauss sur la matrice associée au système :

$$\begin{pmatrix}1&4&2\\2&5&1\\3&6&0\end{pmatrix}\sim\begin{pmatrix}1&4&2\\0&-3&-3\\0&-6&-6\end{pmatrix}\sim\begin{pmatrix}1&4&2\\0&-3&-3\\0&0&0\end{pmatrix}\sim\begin{pmatrix}1&4&2\\0&1&1\\0&0&0\end{pmatrix}\sim\begin{pmatrix}1&0&-2\\0&1&1\\0&0&0\end{pmatrix}$$

Exemple 75

Soient $v_1=\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)$, $v_2=\left(\begin{smallmatrix}2\\-1\\0\end{smallmatrix}\right)$, $v_3=\left(\begin{smallmatrix}2\\1\\1\end{smallmatrix}\right)$. Est-ce que la famille $\{v_1,v_2,v_3\}$ est libre ou liée ? Résolvons le système linéaire correspondant à l'équation $\lambda_1v_1+\lambda_2v_2+\lambda_3v_3=0$:

$$\begin{cases}\lambda_1 + 2\lambda_2 + 2\lambda_3 = 0\\ \lambda_1 - \lambda_2 + \lambda_3 = 0\\ \lambda_1 + \lambda_3 = 0\end{cases}$$

On résout ce système et on trouve comme seule solution $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. La famille $\{v_1, v_2, v_3\}$ est donc une famille libre.

Exemple 76

Soient $v_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 2 \\ 5 \\ -1 \end{pmatrix}$ et $v_3 = \begin{pmatrix} 7 \\ -1 \\ 5 \\ 8 \end{pmatrix}$. Alors $\{v_1, v_2, v_3\}$ forme une famille liée, car

$$3v_1 + v_2 - v_3 = 0.$$

1.4. Autres exemples

Exemple 77

Les polynômes $P_1(X) = 1 - X$, $P_2(X) = 5 + 3X - 2X^2$ et $P_3(X) = 1 + 3X - X^2$ forment une famille liée dans l'espace vectoriel $\mathbb{R}[X]$, car

$$3P_1(X) - P_2(X) + 2P_3(X) = 0.$$

Exemple 78

Dans le $\mathbb{R}$-espace vectoriel $\mathscr{F}(\mathbb{R}, \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on considère la famille $\{\cos, \sin\}$. Montrons que c'est une famille libre. Supposons que l'on ait $\lambda \cos + \mu \sin = 0$. Cela équivaut à

$$\forall x \in \mathbb{R} \qquad \lambda \cos(x) + \mu \sin(x) = 0.$$

En particulier, pour $x = 0$, cette égalité donne $\lambda = 0$. Et pour $x = \frac{\pi}{2}$, elle donne $\mu = 0$. Donc la famille $\{\cos, \sin\}$ est libre. En revanche la famille $\{\cos^2, \sin^2, 1\}$ est liée car on a la relation de dépendance linéaire $\cos^2 + \sin^2 - 1 = 0$. Les coefficients de dépendance linéaire sont $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = -1$.

1.5. Famille liée

Soit E un $\mathbb{K}$-espace vectoriel. Si $v \neq 0$, la famille à un seul vecteur $\{v\}$ est libre (et liée si $v = 0$). Considérons le cas particulier d'une famille de deux vecteurs.

Proposition 23

La famille $\{v_1, v_2\}$ est liée si et seulement si v_1 est un multiple de v_2 ou v_2 est un multiple de v_1.

Ce qui se reformule ainsi par contraposition : « La famille $\{v_1, v_2\}$ est libre si et seulement si v_1 n'est pas un multiple de v_2 et v_2 n'est pas un multiple de v_1. »

Démonstration

- Supposons la famille $\{v_1, v_2\}$ liée, alors il existe λ_1, λ_2 non tous les deux nuls tels que $\lambda_1 v_1 + \lambda_2 v_2 = 0$. Si c'est λ_1 qui n'est pas nul, on peut diviser par λ_1, ce qui donne $v_1 = -\frac{\lambda_2}{\lambda_1} v_2$ et v_1 est un multiple de v_2. Si c'est λ_2 qui n'est pas nul, alors de même v_2 est un multiple de v_1.
- Réciproquement, si v_1 est un multiple de v_2, alors il existe un scalaire μ tel que $v_1 = \mu v_2$, soit $1 v_1 + (-\mu) v_2 = 0$, ce qui est une relation de dépendance linéaire entre v_1 et v_2 puisque $1 \neq 0$: la famille $\{v_1, v_2\}$ est alors liée. Même conclusion si c'est v_2 qui est un multiple de v_1.

Généralisons tout de suite cette proposition à une famille d'un nombre quelconque de vecteurs.

Théorème 18

Soit E un $\mathbb{K}$-espace vectoriel. Une famille $\mathscr{F} = \{v_1, v_2, \dots, v_p\}$ de $p \geqslant 2$ vecteurs de E est une famille liée si et seulement si au moins un des vecteurs de $\mathscr{F}$ est combinaison linéaire des autres vecteurs de $\mathscr{F}$.

Démonstration

C'est essentiellement la même démonstration que ci-dessus.

- Supposons d'abord $\mathscr{F}$ liée. Il existe donc une relation de dépendance linéaire
$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = 0,$$
avec $\lambda_k \neq 0$ pour au moins un indice k. Passons tous les autres termes à droite du signe égal. Il vient
$$\lambda_k v_k = -\lambda_1 v_1 - \lambda_2 v_2 - \cdots - \lambda_p v_p,$$
où v_k ne figure pas au second membre. Comme $\lambda_k \neq 0$, on peut diviser cette égalité par λ_k et l'on obtient
$$v_k = -\frac{\lambda_1}{\lambda_k} v_1 - \frac{\lambda_2}{\lambda_k} v_2 - \cdots - \frac{\lambda_p}{\lambda_k} v_p,$$
c'est-à-dire que v_k est combinaison linéaire des autres vecteurs de $\mathscr{F}$, ce qui peut encore s'écrire $v_k \in \operatorname{Vect}\big(\mathscr{F} \setminus \{v_k\}\big)$ (avec la notation ensembliste $A \setminus B$ pour l'ensemble des éléments de A qui n'appartiennent pas à B).
- Réciproquement, supposons que pour un certain k, on ait $v_k \in \operatorname{Vect}\big(\mathscr{F} \setminus \{v_k\}\big)$. Ceci signifie que l'on peut écrire
$$v_k = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_p v_p,$$
où v_k ne figure pas au second membre. Passant v_k au second membre, il vient
$$0 = \mu_1 v_1 + \mu_2 v_2 + \cdots - v_k + \cdots + \mu_p v_p,$$
ce qui est une relation de dépendance linéaire pour $\mathscr{F}$ (puisque $-1 \neq 0$) et ainsi la famille $\mathscr{F}$ est liée.

1.6. Interprétation géométrique de la dépendance linéaire

- Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, deux vecteurs sont linéairement dépendants si et seulement s'ils sont colinéaires. Ils sont donc sur une même droite vectorielle.
- Dans $\mathbb{R}^3$, trois vecteurs sont linéairement dépendants si et seulement s'ils sont coplanaires. Ils sont donc dans un même plan vectoriel.

Proposition 24

Soit $\mathscr{F} = \{v_1, v_2, \dots, v_p\}$ une famille de vecteurs de $\mathbb{R}^n$. Si $\mathscr{F}$ contient plus de n éléments (c'est-à-dire $p > n$), alors $\mathscr{F}$ est une famille liée.

Démonstration

Supposons que

$$v_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} \qquad v_2 = \begin{pmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{n2} \end{pmatrix} \quad \dots \quad v_p = \begin{pmatrix} v_{1p} \\ v_{2p} \\ \vdots \\ v_{np} \end{pmatrix}.$$

L'équation

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0$$

donne alors le système suivant

$$\begin{cases} v_{11}x_1 + v_{12}x_2 + \cdots + v_{1p}x_p & = & 0 \\ v_{21}x_1 + v_{22}x_2 + \cdots + v_{2p}x_p & = & 0 \\ \vdots \\ v_{n1}x_1 + v_{n2}x_2 + \cdots + v_{np}x_p & = & 0 \end{cases}$$

C'est un système homogène de n équations à p inconnues. Lorsque $p > n$, ce système a des solutions non triviales (voir le chapitre « Systèmes linéaires », dernier théorème) ce qui montre que la famille $\mathscr{F}$ est une famille liée.

Mini-exercices

1. Pour quelles valeurs de $t \in \mathbb{R}$, $\left\{\binom{-1}{t}, \binom{t^2}{-t}\right\}$ est une famille libre de $\mathbb{R}^2$? Même question avec la famille $\left\{\begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} \begin{pmatrix} t^2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ t \\ 1 \end{pmatrix}\right\}$ de $\mathbb{R}^3$.
2. Montrer que toute famille contenant une famille liée est liée.
3. Montrer que toute famille inclue dans une famille libre est libre.
4. Montrer que si $f : E \to F$ est une application linéaire et que $\{v_1, \dots, v_p\}$ est une famille liée de E, alors $\{f(v_1), \dots, f(v_p)\}$ est une famille liée de F.
5. Montrer que si $f : E \to F$ est une application linéaire *injective* et que $\{v_1, \dots, v_p\}$ est une famille libre de E, alors $\{f(v_1), \dots, f(v_p)\}$ est une famille libre de F.

2. Famille génératrice

Soit E un espace vectoriel sur un corps $\mathbb{K}$.

2.1. Définition

Définition 31

Soient $v_1, \dots, v_p$ des vecteurs de E. La famille $\{v_1, \dots, v_p\}$ est une ***famille génératrice*** de l'espace vectoriel E si tout vecteur de E est une combinaison linéaire des vecteurs $v_1, \dots, v_p$. Ce qui peut s'écrire aussi :

$$\forall v \in E \qquad \exists \lambda_1, \dots, \lambda_p \in \mathbb{K} \qquad v = \lambda_1 v_1 + \cdots + \lambda_p v_p$$

On dit aussi que la famille $\{v_1, \dots, v_p\}$ ***engendre*** l'espace vectoriel E.

Cette notion est bien sûr liée à la notion de sous-espace vectoriel engendré : les vecteurs $\{v_1,\dots,v_p\}$ forment une famille génératrice de E si et seulement si $E = \mathrm{Vect}(v_1,\dots,v_p)$.

2.2. Exemples

Exemple 79

Considérons par exemple les vecteurs $v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $v_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$ et $v_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ de $E = \mathbb{R}^3$. La famille $\{v_1,v_2,v_3\}$ est génératrice car tout vecteur $v = \begin{pmatrix}x\\y\\z\end{pmatrix}$ de $\mathbb{R}^3$ peut s'écrire

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

Les coefficients sont ici $\lambda_1 = x$, $\lambda_2 = y$, $\lambda_3 = z$.

Exemple 80

Soient maintenant les vecteurs $v_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $v_2 = \begin{pmatrix}1\\2\\3\end{pmatrix}$ de $E = \mathbb{R}^3$. Les vecteurs $\{v_1,v_2\}$ *ne forment pas* une famille génératrice de $\mathbb{R}^3$. Par exemple, le vecteur $v = \begin{pmatrix}0\\1\\0\end{pmatrix}$ n'est pas dans $\mathrm{Vect}(v_1,v_2)$. En effet, si c'était le cas, alors il existerait $\lambda_1,\lambda_2 \in \mathbb{R}$ tels que $v = \lambda_1 v_1 + \lambda_2 v_2$. Ce qui s'écrirait aussi $\begin{pmatrix}0\\1\\0\end{pmatrix} = \lambda_1\begin{pmatrix}1\\1\\1\end{pmatrix} + \lambda_2\begin{pmatrix}1\\2\\3\end{pmatrix}$, d'où le système linéaire :

$$\begin{cases} \lambda_1 + \lambda_2 & = & 0 \\ \lambda_1 + 2\lambda_2 & = & 1 \\ \lambda_1 + 3\lambda_2 & = & 0 \end{cases}$$

Ce système n'a pas de solution. (La première et la dernière ligne impliquent $\lambda_1 = 0, \lambda_2 = 0$, ce qui est incompatible avec la deuxième.)

Exemple 81

Soit $E = \mathbb{R}^2$.

- Soient $v_1 = \begin{pmatrix}1\\0\end{pmatrix}$ et $v_2 = \begin{pmatrix}0\\1\end{pmatrix}$. La famille $\{v_1,v_2\}$ est génératrice de $\mathbb{R}^2$ car tout vecteur de $\mathbb{R}^2$ se décompose comme $\begin{pmatrix}x\\y\end{pmatrix} = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}$.
- Soient maintenant $v_1' = \begin{pmatrix}2\\1\end{pmatrix}$ et $v_2' = \begin{pmatrix}1\\1\end{pmatrix}$. Alors $\{v_1',v_2'\}$ est aussi une famille génératrice. En effet, soit $v = \begin{pmatrix}x\\y\end{pmatrix}$ un élément quelconque de $\mathbb{R}^2$. Montrer que v est combinaison linéaire de v_1' et v_2' revient à démontrer l'existence de deux réels λ et μ tels que $u = \lambda v_1' + \mu v_2'$. Il s'agit donc d'étudier l'existence de solutions au système :

$$\begin{cases} 2\lambda + \mu & = & x \\ \lambda + \mu & = & y \end{cases}$$

Il a pour solution $\lambda = x - y$ et $\mu = -x + 2y$, et ceci, quels que soient les réels x et y.

Ceci prouve qu'il peut exister plusieurs familles finies différentes, non incluses les unes dans les autres, engendrant le même espace vectoriel.

Exemple 82

Soit $\mathbb{R}_n[X]$ l'espace vectoriel des polynômes de degré $\leqslant n$. Alors les polynômes $\{1, X, \ldots, X^n\}$ forment une famille génératrice. Par contre, l'espace vectoriel $\mathbb{R}[X]$ de tous les polynômes ne possède pas de famille finie génératrice.

2.3. Liens entre familles génératrices

La proposition suivante est souvent utile :

Proposition 25

Soit $\mathscr{F} = \{v_1, v_2, \ldots, v_p\}$ une famille génératrice de E. Alors $\mathscr{F}' = \left\{v'_1, v'_2, \ldots, v'_q\right\}$ est aussi une famille génératrice de E si et seulement si tout vecteur de $\mathscr{F}$ est une combinaison linéaire de vecteurs de $\mathscr{F}'$.

Démonstration

C'est une conséquence immédiate de la définition de $\operatorname{Vect}\mathscr{F}$ et de $\operatorname{Vect}\mathscr{F}'$.

Nous chercherons bientôt à avoir un nombre minimal de générateurs. Voici une proposition sur la réduction d'une famille génératrice.

Proposition 26

Si la famille de vecteurs $\{v_1, \ldots, v_p\}$ engendre E et si l'un des vecteurs, par exemple v_p, est combinaison linéaire des autres, alors la famille $\{v_1, \ldots, v_p\} \setminus \{v_p\} = \{v_1, \ldots, v_{p-1}\}$ est encore une famille génératrice de E.

Démonstration

En effet, comme les vecteurs $v_1, \ldots, v_p$ engendrent E, alors pour tout élément x de E, il existe des scalaires $\lambda_1, \ldots, \lambda_p$ tels que

$$x = \lambda_1 v_1 + \cdots + \lambda_p v_p.$$

Or l'hypothèse v_p est combinaison linéaire des vecteurs $v_1, \ldots, v_{p-1}$ se traduit par l'existence de scalaires $\alpha_1, \ldots, \alpha_{p-1}$ tels que

$$v_p = \alpha_1 v_1 + \cdots + \alpha_{p-1} v_{p-1}.$$

Alors, le vecteur x s'écrit :

$$x = \lambda_1 v_1 + \cdots + \lambda_{p-1} v_{p-1} + \lambda_p \left(\alpha_1 v_1 + \cdots + \alpha_{p-1} v_{p-1}\right).$$

Donc

$$x = \left(\lambda_1 + \lambda_p \alpha_1\right) v_1 + \cdots + \left(\lambda_{p-1} + \lambda_p \alpha_{p-1}\right) v_{p-1},$$

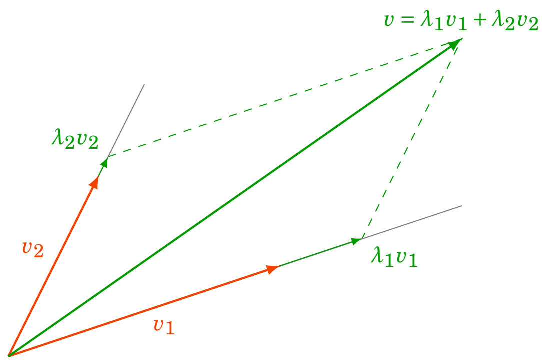
ce qui prouve que x est combinaison linéaire des vecteurs $v_1, \ldots, v_{p-1}$. Ceci achève la démonstration. Il est clair que si l'on remplace v_p par n'importe lequel des vecteurs v_i, la démonstration est la même.

Mini-exercices

1. À quelle condition sur $t \in \mathbb{R}$ la famille $\left\{\binom{0}{t-1}, \binom{t}{-t} \binom{t^2-t}{t-1}\right\}$ est une famille génératrice de $\mathbb{R}^2$?
2. Même question avec la famille $\left\{\begin{pmatrix}1\\0\\t\end{pmatrix} \begin{pmatrix}1\\t\\t^2\end{pmatrix} \begin{pmatrix}1\\t^2\\1\end{pmatrix}\right\}$ de $\mathbb{R}^3$.
3. Montrer qu'une famille de vecteurs contenant une famille génératrice est encore une famille génératrice de E.
4. Montrer que si $f : E \to F$ est une application linéaire *surjective* et que $\{v_1, \dots, v_p\}$ est une famille génératrice de E , alors $\{f(v_1), \dots, f(v_p)\}$ est une famille génératrice de F.

3. Base

La notion de base généralise la notion de repère. Dans $\mathbb{R}^2$, un repère est donné par un couple de vecteurs non colinéaires. Dans $\mathbb{R}^3$, un repère est donné par un triplet de vecteurs non coplanaires. Dans un repère, un vecteur se décompose suivant les vecteurs d'une base. Il en sera de même pour une base d'un espace vectoriel.



3.1. Définition

Définition 32. Base d'un espace vectoriel

Soit E un $\mathbb{K}$-espace vectoriel. Une famille $\mathscr{B} = (v_1, v_2, \dots, v_n)$ de vecteurs de E est une ***base*** de E si $\mathscr{B}$ est une famille libre ***et*** génératrice.

Théorème 19

Soit $\mathscr{B} = (v_1, v_2, \dots, v_n)$ une base de l'espace vectoriel E. Tout vecteur $v \in E$ s'exprime de façon unique comme combinaison linéaire d'éléments de $\mathscr{B}$. Autrement dit, il ***existe*** des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ ***uniques*** tels que :

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Remarque

1. $(\lambda_1,\ldots,\lambda_n)$ s'appellent les ***coordonnées*** du vecteur v dans la base $\mathscr{B}$.
2. Il faut observer que pour une base $\mathscr{B}=(v_1,v_2,\ldots,v_n)$ on introduit un ***ordre*** sur les vecteurs. Bien sûr, si on permutait les vecteurs on obtiendrait toujours une base, mais il faudrait aussi permuter les coordonnées.
3. Notez que l'application

$$\begin{array}{rcl}\phi & : & \mathbb{K}^n \to E \\ (\lambda_1,\lambda_2,\ldots,\lambda_n) & \mapsto & \lambda_1 v_1+\lambda_2 v_2+\cdots+\lambda_n v_n\end{array}$$

est un isomorphisme de l'espace vectoriel $\mathbb{K}^n$ vers l'espace vectoriel E.

Démonstration . Preuve du théorème 19

- Par définition, $\mathscr{B}$ est une famille génératrice de E, donc pour tout $v\in E$ il existe $\lambda_1,\ldots,\lambda_n\in\mathbb{K}$ tels que

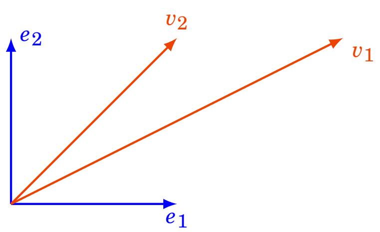
$$v=\lambda_1 v_1+\lambda_2 v_2+\cdots+\lambda_n v_n.$$

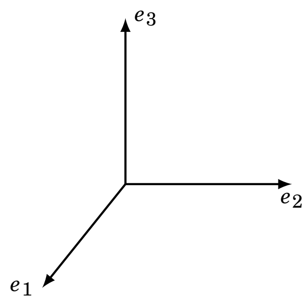
Cela prouve la partie existence.
- Il reste à montrer l'unicité des $\lambda_1,\lambda_2,\ldots,\lambda_n$. Soient $\mu_1,\mu_2,\ldots,\mu_n\in\mathbb{K}$ d'autres scalaires tels que $v=\mu_1 v_1+\mu_2 v_2+\cdots+\mu_n v_n$. Alors, par différence on a : $(\lambda_1-\mu_1)v_1+(\lambda_2-\mu_2)v_2+\cdots+(\lambda_n-\mu_n)v_n=0$. Comme $\mathscr{B}=\{v_1,\ldots,v_n\}$ est une famille libre, ceci implique $\lambda_1-\mu_1=0,\quad \lambda_2-\mu_2=0,\quad \ldots,\quad \lambda_n-\mu_n=0$ et donc $\lambda_1=\mu_1,\quad \lambda_2=\mu_2,\quad \ldots,\lambda_n=\mu_n$.

3.2. Exemples

Exemple 83

1. Soient les vecteurs $e_1=\binom{1}{0}$ et $e_2=\binom{0}{1}$. Alors (e_1,e_2) est une base de $\mathbb{R}^2$, appelée ***base canonique*** de $\mathbb{R}^2$.
2. Soient les vecteurs $v_1=\binom{2}{1}$ et $v_2=\binom{1}{1}$. Alors (v_1,v_2) forment aussi une base de $\mathbb{R}^2$.





3. De même dans $\mathbb{R}^3$, si $e_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$, $e_2=\begin{pmatrix}0\\1\\0\end{pmatrix}$, $e_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$, alors (e_1,e_2,e_3) forment la ***base canonique*** de $\mathbb{R}^3$.

Exemple 84

Soient $v_1=\begin{pmatrix}1\\2\\1\end{pmatrix}$, $v_2=\begin{pmatrix}2\\9\\0\end{pmatrix}$ et $v_3=\begin{pmatrix}3\\3\\4\end{pmatrix}$. Montrons que la famille $\mathscr{B}=(v_1,v_2,v_3)$ est une base de $\mathbb{R}^3$.

Dans les deux premiers points, nous ramenons le problème à l'étude d'un système linéaire.

1. Montrons d'abord que $\mathscr{B}$ est une famille génératrice de $\mathbb{R}^3$. Soit $v=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}$ un vecteur

quelconque de $\mathbb{R}^3$. On cherche $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3.$$

Ceci se reformule comme suit :

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 9 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \lambda_1 + 2\lambda_2 + 3\lambda_3 \\ 2\lambda_1 + 9\lambda_2 + 3\lambda_3 \\ \lambda_1 + 4\lambda_3 \end{pmatrix}.$$

Ceci conduit au système suivant :

$$\left\{\begin{array}{ccccccc} \lambda_1 & + & 2\lambda_2 & + & 3\lambda_3 & = & a_1 \\ 2\lambda_1 & + & 9\lambda_2 & + & 3\lambda_3 & = & a_2 \\ \lambda_1 & & & + & 4\lambda_3 & = & a_3. \end{array}\right. \tag{S}$$

Il nous restera à montrer que ce système a une solution $\lambda_1, \lambda_2, \lambda_3$.

2. Pour montrer que $\mathscr{B}$ est une famille libre, il faut montrer que l'unique solution de

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$$

est

$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Ceci équivaut à montrer que le système

$$\left\{\begin{array}{ccccccc} \lambda_1 & + & 2\lambda_2 & + & 3\lambda_3 & = & 0 \\ 2\lambda_1 & + & 9\lambda_2 & + & 3\lambda_3 & = & 0 \\ \lambda_1 & & & + & 4\lambda_3 & = & 0 \end{array}\right. \tag{S'}$$

a une unique solution

$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

3. Nous pouvons maintenant répondre à la question sans explicitement résoudre les systèmes.

Remarquons que les deux systèmes ont la même matrice de coefficients. On peut donc montrer simultanément que $\mathscr{B}$ est une famille génératrice et une famille libre de $\mathbb{R}^3$ en montrant que la matrice des coefficients est inversible. En effet, si la matrice des coefficients est inversible, alors (S) admet une solution $(\lambda_1, \lambda_2, \lambda_3)$ quel que soit (a_1, a_2, a_3) et d'autre part (S') admet la seule solution $(0,0,0)$.

Cette matrice est

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}.$$

Pour montrer qu'elle est inversible, on peut calculer son inverse ou seulement son déterminant qui vaut $\det A = -1$ (le déterminant étant non nul la matrice est inversible).

Conclusion : $\mathscr{B}$ est une famille libre et génératrice ; c'est une base de $\mathbb{R}^3$.

Exemple 85

Les vecteurs de $\mathbb{K}^n$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \qquad \dots \qquad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

forment une base de $\mathbb{K}^n$, appelée la ***base canonique*** de $\mathbb{K}^n$.

Remarque

L'exemple 84 se généralise de la façon suivante. Pour montrer que n vecteurs de $\mathbb{R}^n$ forment une base de $\mathbb{R}^n$, il suffit de montrer la chose suivante : la matrice A constituée des composantes de ces vecteurs (chaque vecteur formant une colonne de A) est inversible.

Application : montrer que les vecteurs

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \dots \qquad v_n = \begin{pmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{pmatrix}$$

forment aussi une base de $\mathbb{R}^n$.

Voici quelques autres exemples :

Exemple 86

1. La base canonique de $\mathbb{R}_n[X]$ est $\mathcal{B} = (1, X, X^2, \dots, X^n)$. Attention, il y a $n+1$ vecteurs !
2. Voici une autre base de $\mathbb{R}_n[X]$: $(1, 1+X, 1+X+X^2, \dots, 1+X+X^2+\dots+X^n)$.
3. L'espace vectoriel $\mathcal{M}_2(\mathbb{R})$ des matrices 2×2 admet une base formée des vecteurs :

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

   En effet, n'importe quelle matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de $\mathcal{M}_2(\mathbb{R})$ se décompose de manière unique en

$$M = aM_1 + bM_2 + cM_3 + dM_4.$$

4. C'est un bon exercice de prouver que les quatre matrices suivantes forment aussi une base de $\mathcal{M}_2(\mathbb{R})$:

$$M'_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \qquad M'_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad M'_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad M'_4 = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}.$$

3.3. Existence d'une base

Voyons maintenant un théorème d'existence d'une base finie. Dans la suite, les espaces vectoriels sont supposés non réduits à $\{0\}$.

Théorème 20. Théorème d'existence d'une base

Tout espace vectoriel admettant une famille finie de générateurs admet une base.

3.4. Théorème de la base incomplète

Une version importante et plus générale de ce qui précède est le théorème suivant :

Théorème 21. Théorème de la base incomplète

Soit E un $\mathbb{K}$-espace vectoriel admettant une famille génératrice.

1. *Toute famille libre $\mathscr{L}$ peut être complétée en une base.* C'est-à-dire qu'il existe une famille $\mathscr{F}$ telle que $\mathscr{L} \cup \mathscr{F}$ soit une famille libre et génératrice de E.
2. *De toute famille génératrice $\mathscr{G}$ on peut extraire une base de E.* C'est-à-dire qu'il existe une famille $\mathscr{B} \subset \mathscr{G}$ telle que $\mathscr{B}$ soit une famille libre et génératrice de E.

3.5. Preuves

Les deux théorèmes précédents sont la conséquence d'un résultat encore plus général :

Théorème 22

Soit $\mathscr{G}$ une partie génératrice finie de E et $\mathscr{L}$ une partie libre de E. Alors il existe une partie $\mathscr{F}$ de $\mathscr{G}$ telle que $\mathscr{L} \cup \mathscr{F}$ soit une base de E.

Le théorème 21 de la base incomplète se déduit du théorème 22 ainsi :

1. On sait qu'il existe une famille génératrice de E : notons-la $\mathscr{G}$. On applique le théorème 22 avec ce $\mathscr{L}$ et ce $\mathscr{G}$.
2. On applique le théorème 22 avec $\mathscr{L} = \varnothing$ et la famille $\mathscr{G}$ de l'énoncé.

En particulier, le théorème 20 d'existence d'une base se démontre comme le point (2) ci-dessus avec $\mathscr{L} = \varnothing$ et $\mathscr{G}$ une famille génératrice de E.

3.6. Preuves (suite)

Nous avons : Théorème 22 $\implies$ Théorème 21 $\implies$ Théorème 20.
Il nous reste donc à prouver le théorème 22. La démonstration que nous en donnons est un algorithme.

Démonstration

- Étape 0. Si $\mathscr{L}$ est une famille génératrice de E, on pose $\mathscr{F} = \varnothing$ et c'est fini puisque $\mathscr{L}$ est une famille génératrice et libre, donc une base. Sinon on passe à l'étape suivante.
- Étape 1. Comme $\mathscr{L}$ n'est pas une famille génératrice, alors il existe au moins un élément g_1 de $\mathscr{G}$ qui n'est pas combinaison linéaire des éléments de $\mathscr{L}$. (En effet, par l'absurde, si tous les éléments de $\mathscr{G}$ sont dans $\operatorname{Vect}\mathscr{L}$, alors $\mathscr{L}$ serait aussi une famille génératrice.) On pose $\mathscr{L}_1 = \mathscr{L} \cup \{g_1\}$. Alors la famille $\mathscr{L}_1$ vérifie les propriétés suivantes :
  (i) $\mathscr{L} \subsetneq \mathscr{L}_1 \subset E$: la famille $\mathscr{L}_1$ est strictement plus grande que $\mathscr{L}$.
  (ii) $\mathscr{L}_1$ est une famille libre. (En effet, si $\mathscr{L}_1$ n'était pas une famille libre, alors une combinaison linéaire nulle impliquerait que $g_1 \in \operatorname{Vect}\mathscr{L}$.)
  On recommence le même raisonnement à partir de $\mathscr{L}_1$: si $\mathscr{L}_1$ est une famille génératrice de

E, alors on pose $\mathscr{F} = \{g_1\}$ et on s'arrête. Sinon on passe à l'étape suivante.

- Étape 2. Il existe au moins un élément g_2 de $\mathscr{G}$ qui n'est pas combinaison linéaire des éléments de $\mathscr{L}_1$. Alors la partie $\mathscr{L}_2 = \mathscr{L}_1 \cup \{g_2\} = \mathscr{L} \cup \{g_1, g_2\}$ est strictement plus grande que $\mathscr{L}_1$ et est encore une famille libre.
  Si $\mathscr{L}_2$ est une famille génératrice, on pose $\mathscr{F} = \{g_1, g_2\}$ et c'est fini. Sinon on passe à l'étape d'après.
- ...

L'algorithme consiste donc à construire une suite, strictement croissante pour l'inclusion, de parties libres, où, si $\mathscr{L}_{k-1}$ n'engendre pas E, alors $\mathscr{L}_k$ est construite partir de $\mathscr{L}_{k-1}$ en lui ajoutant un vecteur g_k de $\mathscr{G}$, de sorte que $\mathscr{L}_k = \mathscr{L}_{k-1} \cup \{g_k\}$ reste une famille libre.

- L'algorithme se termine. Comme la partie $\mathscr{G}$ est finie, le processus s'arrête en moins d'étapes qu'il y a d'éléments dans $\mathscr{G}$. Notez que, comme $\mathscr{G}$ est une famille génératrice, dans le pire des cas on peut être amené à prendre $\mathscr{F} = \mathscr{G}$.
- L'algorithme est correct. Lorsque l'algorithme s'arrête, disons à l'étape s : on a $\mathscr{L}_s = \mathscr{L} \cup \mathscr{F}$ où $\mathscr{F} = \{g_1, \ldots, g_s\}$. Par construction, $\mathscr{L}_s$ est une famille finie, libre et aussi génératrice (car c'est la condition d'arrêt). Donc $\mathscr{L} \cup \mathscr{F}$ est une base de E.

Exemple 87

Soit $\mathbb{R}[X]$ le $\mathbb{R}$-espace vectoriel des polynômes réels et E le sous-espace de $\mathbb{R}[X]$ engendré par la famille $\mathscr{G} = \{P_1, P_2, P_3, P_4, P_5\}$ définie par :

$$P_1(X) = 1 \qquad P_2(X) = X \qquad P_3(X) = X + 1 \qquad P_4(X) = 1 + X^3 \qquad P_5(X) = X - X^3$$

Partons de $\mathscr{L} = \varnothing$ et cherchons $\mathscr{F} \subset \mathscr{G}$ telle que $\mathscr{F}$ soit une base de E.

- Étape 0. Comme $\mathscr{L}$ n'est pas génératrice (vu que $\mathscr{L} = \varnothing$), on passe à l'étape suivante.
- Étape 1. On pose $\mathscr{L}_1 = \mathscr{L} \cup \{P_1\} = \{P_1\}$. Comme P_1 est non nul, $\mathscr{L}_1$ est une famille libre.
- Étape 2. Considérons P_2. Comme les éléments P_1 et P_2 sont linéairement indépendants, $\mathscr{L}_2 = \{P_1, P_2\}$ est une famille libre.
- Étape 3. Considérons P_3 : ce vecteur est combinaison linéaire des vecteurs P_1 et P_2 car $P_3(X) = X + 1 = P_1(X) + P_2(X)$ donc $\{P_1, P_2, P_3\}$ est une famille liée. Considérons alors P_4. Un calcul rapide prouve que les vecteurs P_1, P_2 et P_4 sont linéairement indépendants. Alors $\mathscr{L}_3 = \{P_1, P_2, P_4\}$ est une famille libre.
  Il ne reste que le vecteur P_5 à considérer. Il s'agit, pour pouvoir conclure, d'étudier l'indépendance linéaire des vecteurs P_1, P_2, P_4, P_5. Or un calcul rapide montre l'égalité

$$P_1 + P_2 - P_4 - P_5 = 0,$$

  ce qui prouve que la famille $\{P_1, P_2, P_4, P_5\}$ est liée. Donc avec les notations de l'algorithme, $s = 3$ et $\mathscr{L}_3 = \{P_1, P_2, P_4\}$ est une base de E.

Mini-exercices

1. Trouver toutes les façons d'obtenir une base de $\mathbb{R}^2$ avec les vecteurs suivants : $v_1 = \binom{-1}{-3}$, $v_2 = \binom{3}{3}$, $v_3 = \binom{0}{0}$, $v_4 = \binom{2}{0}$, $v_5 = \binom{2}{6}$.
2. Montrer que la famille $\{v_1, v_2, v_3, v_4\}$ des vecteurs $v_1 = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$, $v_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ est une famille génératrice du sous-espace vectoriel d'équation $2x - y + z = 0$ de $\mathbb{R}^3$. En extraire une base.
3. Déterminer une base du sous-espace vectoriel E_1 de $\mathbb{R}^3$ d'équation $x + 3y - 2z = 0$. Complé-

ter cette base en une base de $\mathbb{R}^3$. Idem avec E_2 vérifiant les deux équations $x+3y-2z=0$ et $y=z$.

4. Donner une base de l'espace vectoriel des matrices 3×3 ayant une diagonale nulle. Idem avec l'espace vectoriel des polynômes $P\in\mathbb{R}_n[X]$ vérifiant $P(0)=0$, $P'(0)=0$.

4. Dimension d'un espace vectoriel

4.1. Définition

Définition 33

Un $\mathbb{K}$-espace vectoriel E admettant une base ayant un nombre fini d'éléments est dit de ***dimension finie***.

Par le théorème 20 d'existence d'une base, c'est équivalent à l'existence d'une famille finie génératrice.
On va pouvoir parler de ***la*** dimension d'un espace vectoriel grâce au théorème suivant :

Théorème 23. Théorème de la dimension

Toutes les bases d'un espace vectoriel E de dimension finie ont le même nombre d'éléments.

Nous détaillerons la preuve un peu plus loin.

Définition 34

La ***dimension*** d'un espace vectoriel de dimension finie E, notée $\dim E$, est par définition le nombre d'éléments d'une base de E.

Méthodologie. Pour déterminer la dimension d'un espace vectoriel, il suffit de trouver une base de E (une partie à la fois libre et génératrice) : le cardinal (nombre d'éléments) de cette partie donne la dimension de E. Le théorème 23 de la dimension prouve que même si on choisissait une base différente alors ces deux bases auraient le même nombre d'éléments.
Convention. On convient d'attribuer à l'espace vectoriel $\{0\}$ la dimension 0.

4.2. Exemples

Exemple 88

1. La base canonique de $\mathbb{R}^2$ est $\left(\binom{1}{0},\binom{0}{1}\right)$. La dimension de $\mathbb{R}^2$ est donc 2.
2. Les vecteurs $\left(\binom{2}{1},\binom{1}{1}\right)$ forment aussi une base de $\mathbb{R}^2$, et illustrent qu'une autre base contient le même nombre d'éléments.
3. Plus généralement, $\mathbb{K}^n$ est de dimension n, car par exemple sa base canonique $(e_1,e_2,\ldots,e_n)$ contient n éléments.
4. $\dim\mathbb{R}_n[X]=n+1$ car une base de $\mathbb{R}_n[X]$ est $(1,X,X^2,\ldots,X^n)$, qui contient $n+1$ éléments.

Exemple 89

Les espaces vectoriels suivants ne sont pas de dimension finie :

- $\mathbb{R}[X]$: l'espace vectoriel de tous les polynômes,
- $\mathscr{F}(\mathbb{R},\mathbb{R})$: l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$,
- $\mathscr{S} = \mathscr{F}(\mathbb{N},\mathbb{R})$: l'espace vectoriel des suites réelles.

Exemple 90

Nous avons vu que l'ensemble des solutions d'un système d'équations linéaires ***homogène*** est un espace vectoriel. On considère par exemple le système

$$\left\{\begin{array}{rcrcrcrcrcl} 2x_1 & + & 2x_2 & - & x_3 & & & + & x_5 & = & 0 \\ -x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & + & x_5 & = & 0 \\ x_1 & + & x_2 & - & 2x_3 & & & - & x_5 & = & 0 \\ & & & & x_3 & + & x_4 & + & x_5 & = & 0. \end{array}\right.$$

On vérifie que la solution générale de ce système est

$$x_1 = -s-t \qquad x_2 = s \qquad x_3 = -t \qquad x_4 = 0 \qquad x_5 = t\,.$$

Donc les vecteurs solutions s'écrivent sous la forme

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ -t \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} -s \\ s \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -t \\ 0 \\ -t \\ 0 \\ t \end{pmatrix} = s\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Ceci montre que les vecteurs

$$v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

engendrent l'espace des solutions du système. D'autre part, on vérifie que v_1 et v_2 sont linéairement indépendants. Donc (v_1, v_2) est une base de l'espace des solutions du système. Ceci montre que cet espace vectoriel est de dimension 2.

4.3. Compléments

Lorsqu'un espace vectoriel est de dimension finie, le fait de connaître sa dimension est une information très riche ; les propriétés suivantes montrent comment exploiter cette information.
Le schéma de preuve sera : Lemme 1 $\Longrightarrow$ Proposition 27 $\Longrightarrow$ Théorème 23.

Lemme 1

Soit E un espace vectoriel. Soit $\mathscr{L}$ une famille libre et soit $\mathscr{G}$ une famille génératrice finie de E. Alors $\operatorname{Card}\mathscr{L} \leqslant \operatorname{Card}\mathscr{G}$.

Ce lemme implique le résultat important :

Proposition 27

Soit E un $\mathbb{K}$-espace vectoriel admettant une base ayant n éléments. Alors :

1. Toute partie libre de E a au plus n éléments.
2. Toute partie génératrice de E a au moins n éléments.

En effet, soit $\mathscr{B}$ une base de E telle que $\operatorname{Card}\mathscr{B} = n$.

1. On applique le lemme 1 à la famille $\mathscr{B}$ considérée génératrice ; alors une famille libre $\mathscr{L}$ vérifie $\operatorname{Card}\mathscr{L} \leqslant \operatorname{Card}\mathscr{B} = n$.
2. On applique le lemme 1 à la famille $\mathscr{B}$ considérée maintenant comme une famille libre, alors une famille génératrice $\mathscr{G}$ vérifie $n = \operatorname{Card}\mathscr{B} \leqslant \operatorname{Card}\mathscr{G}$.

Cette proposition impliquera bien le théorème 23 de la dimension :

Corollaire 4

Si E est un espace vectoriel admettant une base ayant n éléments, alors toute base de E possède n éléments.

La preuve du corollaire (et donc du théorème 23 de la dimension) est la suivante : par la proposition 27, si $\mathscr{B}$ est une base quelconque de E, alors $\mathscr{B}$ est à la fois une famille libre et génératrice, donc possède à la fois au plus n éléments et au moins n éléments, donc exactement n éléments.

Il reste à énoncer un résultat important et très utile :

Théorème 24

Soient E un $\mathbb{K}$-espace vectoriel de dimension n, et $\mathscr{F} = (v_1, \dots, v_n)$ une famille de n vecteurs de E. Il y a équivalence entre :

(i) $\mathscr{F}$ est une base de E,
(ii) $\mathscr{F}$ est une famille libre de E,
(iii) $\mathscr{F}$ est une famille génératrice de E.

La preuve sera une conséquence du théorème 23 de la dimension et du théorème 21 de la base incomplète.
Autrement dit, lorsque le nombre de vecteurs considéré est exactement égal à la dimension de l'espace vectoriel, l'une des deux conditions – être libre ou bien génératrice – suffit pour que ces vecteurs déterminent une base de E.

Démonstration

- Les implications (i) $\implies$ (ii) et (i) $\implies$ (iii) découlent de la définition d'une base.
- Voici la preuve de (ii) $\implies$ (i).
  Si $\mathscr{F}$ est une famille libre ayant n éléments, alors par le théorème de la base incomplète (théorème 21) il existe une famille $\mathscr{F}'$ telle que $\mathscr{F} \cup \mathscr{F}'$ soit une base de E. D'une part $\mathscr{F} \cup \mathscr{F}'$ est une base de E qui est de dimension n, donc par le théorème 23, $\operatorname{Card}\left(\mathscr{F} \cup \mathscr{F}'\right) = n$. Mais d'autre part $\operatorname{Card}\left(\mathscr{F} \cup \mathscr{F}'\right) = \operatorname{Card}\mathscr{F} + \operatorname{Card}\mathscr{F}'$ (par l'algorithme du théorème 21) et par hypothèse $\operatorname{Card}\mathscr{F} = n$. Donc $\operatorname{Card}\mathscr{F}' = 0$, ce qui implique que $\mathscr{F}' = \varnothing$ et donc que $\mathscr{F}$ est déjà une base de E.
- Voici la preuve de (iii) $\implies$ (i).
  Par hypothèse, $\mathscr{F}$ est cette fois une famille génératrice. Toujours par le théorème 21, on peut extraire de cette famille une base $\mathscr{B} \subset \mathscr{F}$. Puis par le théorème 23, $\operatorname{Card}\mathscr{B} = n$, donc

$n = \operatorname{Card} \mathscr{B} \leqslant \operatorname{Card} \mathscr{F} = n$. Donc $\mathscr{B} = \mathscr{F}$ et $\mathscr{F}$ est bien une base.

Exemple 91

Pour quelles valeurs de $t \in \mathbb{R}$ les vecteurs (v_1, v_2, v_3) suivants forment une base de $\mathbb{R}^3$?

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 3 \\ t \end{pmatrix} \qquad v_3 = \begin{pmatrix} 1 \\ 1 \\ t \end{pmatrix}$$

- Nous avons une famille de 3 vecteurs dans l'espace $\mathbb{R}^3$ de dimension 3. Donc pour montrer que la famille (v_1, v_2, v_3) est une base, par le théorème 24, il suffit de montrer que la famille est libre ou bien de montrer qu'elle est génératrice. Dans la pratique, il est souvent plus facile de vérifier qu'une famille est libre.
- À quelle condition la famille $\{v_1, v_2, v_3\}$ est libre ? Soient $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$. Cela implique le système

$$\left\{ \begin{array}{rcl} \lambda_1 + \lambda_2 + \lambda_3 & = & 0 \\ \lambda_1 + 3\lambda_2 + \lambda_3 & = & 0 \\ 4\lambda_1 + t\lambda_2 + t\lambda_3 & = & 0 \end{array} \right. .$$

Ce système est équivalent à :

$$\left\{ \begin{array}{rcl} \lambda_1 + \lambda_2 + \lambda_3 & = & 0 \\ 2\lambda_2 & = & 0 \\ (t-4)\lambda_2 + (t-4)\lambda_3 & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} \lambda_1 + \lambda_3 & = & 0 \\ \lambda_2 & = & 0 \\ (t-4)\lambda_3 & = & 0 \end{array} \right.$$

- Il est clair que si $t \neq 4$, alors la seule solution est $(\lambda_1, \lambda_2, \lambda_3) = (0,0,0)$ et donc $\{v_1, v_2, v_3\}$ est une famille libre. Si $t = 4$, alors par exemple $(\lambda_1, \lambda_2, \lambda_3) = (1, 0, -1)$ est une solution non nulle, donc la famille n'est pas libre.
- Conclusion : si $t \neq 4$ la famille est libre, donc par le théorème 24 la famille (v_1, v_2, v_3) est en plus génératrice, donc c'est une base de $\mathbb{R}^3$. Si $t = 4$, la famille n'est pas libre et n'est donc pas une base.

4.4. Preuve

Il nous reste la preuve du lemme 1. La démonstration est délicate et hors-programme.

Démonstration

La preuve de ce lemme se fait en raisonnant par récurrence.
On démontre par récurrence que, pour tout $n \geqslant 1$, la propriété suivante est vraie : « Dans un espace vectoriel engendré par n vecteurs, toute famille ayant $n+1$ éléments est liée. »
Initialisation. On vérifie que la propriété est vraie pour $n = 1$. Soit E un espace vectoriel engendré par un vecteur noté g_1, et soit $\{v_1, v_2\}$ une partie de E ayant deux éléments. Les vecteurs v_1 et v_2 peuvent s'écrire comme combinaisons linéaires du vecteur g_1 ; autrement dit, il existe des scalaires α_1, α_2 tels que $v_1 = \alpha_1 g_1$ et $v_2 = \alpha_2 g_1$, ce qui donne la relation : $\alpha_2 v_1 - \alpha_1 v_2 = 0_E$. En supposant v_2 non nul (sinon il est évident que $\{v_1, v_2\}$ est liée), le scalaire α_2 est donc non nul. On a trouvé une combinaison linéaire nulle des vecteurs v_1, v_2, avec des coefficients non tous nuls. Donc la famille $\{v_1, v_2\}$ est liée.

Hérédité. On démontre maintenant que si la propriété est vraie au rang $n-1$ ($n \geqslant 2$), alors elle vraie au rang n. Soit E un espace vectoriel engendré par n vecteurs notés $g_1, g_2, \ldots, g_n$, et soit

$\{v_1, v_2, \ldots, v_n, v_{n+1}\}$ une famille de E ayant $n+1$ éléments. Tout vecteur v_j, pour $j = 1, 2, \ldots, n+1$, est combinaison linéaire de $g_1, g_2, \ldots, g_n$, donc il existe des scalaires $\alpha_1^j, \alpha_2^j, \ldots, \alpha_n^j$ tels que :

$$v_j = \alpha_1^j g_1 + \alpha_2^j g_2 + \cdots + \alpha_n^j g_n.$$

Remarque. On est contraint d'utiliser ici deux indices i, j pour les scalaires (attention ! j n'est pas un exposant) car deux informations sont nécessaires : l'indice j indique qu'il s'agit de la décomposition du vecteur v_j, et i indique à quel vecteur de la partie génératrice est associé ce coefficient.

En particulier, pour $j = n+1$, le vecteur v_{n+1} s'écrit :

$$v_{n+1} = \alpha_1^{n+1} g_1 + \alpha_2^{n+1} g_2 + \cdots + \alpha_n^{n+1} g_n.$$

Si v_{n+1} est nul, c'est terminé, la partie est liée ; sinon, v_{n+1} est non nul, et au moins un des coefficients α_j^{n+1} est non nul. On suppose, pour alléger l'écriture, que α_n^{n+1} est non nul (sinon il suffit de changer l'ordre des vecteurs). On construit une nouvelle famille de n vecteurs de E de telle sorte que ces vecteurs soient combinaisons linéaires de $g_1, g_2, \ldots, g_{n-1}$, c'est-à-dire appartiennent au sous-espace engendré par $\{g_1, g_2, \ldots, g_{n-1}\}$. Pour $j = 1, 2, \ldots, n$, on définit w_j par :

$$w_j = \alpha_n^{n+1} v_j - \alpha_n^j v_{n+1} = \sum_{k=1}^{n} (\alpha_n^{n+1} \alpha_k^j - \alpha_n^j \alpha_k^{n+1}) g_k.$$

Le coefficient de g_n est nul. Donc w_j est bien combinaison linéaire de $g_1, g_2, \ldots, g_{n-1}$. On a n vecteurs qui appartiennent à un espace vectoriel engendré par $n-1$ vecteurs ; on peut appliquer l'hypothèse de récurrence : la famille $\{w_1, w_2, \ldots, w_n\}$ est liée. Par conséquent, il existe des scalaires non tous nuls λ_1, λ_2, ..., λ_n tels que

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n = 0.$$

En remplaçant les w_j par leur expression en fonction des vecteurs v_i, on obtient :

$$\alpha_n^{n+1} \lambda_1 v_1 + \alpha_n^{n+1} \lambda_2 v_2 + \cdots + \alpha_n^{n+1} \lambda_n v_n - (\lambda_1 \alpha_n^1 + \cdots + \lambda_n \alpha_n^n) v_{n+1} = 0_E$$

Le coefficient α_n^{n+1} a été supposé non nul et au moins un des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_n$ est non nul ; on a donc une combinaison linéaire nulle des vecteurs $v_1, v_2, \ldots, v_n, v_{n+1}$ avec des coefficients qui ne sont pas tous nuls, ce qui prouve que ces vecteurs forment une famille liée.

Conclusion. La démonstration par récurrence est ainsi achevée.

Mini-exercices

Dire si les assertions suivantes sont vraies ou fausses. Justifier votre réponse par un résultat du cours ou un contre-exemple :

1. Une famille de $p \geqslant n$ vecteurs dans un espace vectoriel de dimension n est génératrice.
2. Une famille de $p > n$ vecteurs dans un espace vectoriel de dimension n est liée.
3. Une famille de $p < n$ vecteurs dans un espace vectoriel de dimension n est libre.
4. Une famille génératrice de $p \leqslant n$ vecteurs dans un espace vectoriel de dimension n est libre.
5. Une famille de $p \neq n$ vecteurs dans un espace vectoriel de dimension n n'est pas une base.

6. Toute famille libre à p éléments d'un espace vectoriel de dimension n se complète par une famille ayant exactement $n-p$ éléments en une base de E.

5. Dimension des sous-espaces vectoriels

Tout sous-espace vectoriel F d'un $\mathbb{K}$-espace vectoriel E étant lui même un $\mathbb{K}$-espace vectoriel, la question est de savoir s'il est de dimension finie ou s'il ne l'est pas.
Prenons l'exemple de l'espace vectoriel $E = \mathscr{F}(\mathbb{R},\mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$:

- il contient le sous-espace vectoriel $F_1 = \mathbb{R}_n[X]$ des (fonctions) polynômes de degré $\leqslant n$, qui est de dimension finie ;
- et aussi le sous-espace vectoriel $F_2 = \mathbb{R}[X]$ de l'ensemble des (fonctions) polynômes, qui lui est de dimension infinie.

5.1. Dimension d'un sous-espace vectoriel

Nous allons voir par contre que lorsque E est de dimension finie alors F aussi.

Théorème 25

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.

1. Alors tout sous-espace vectoriel F de E est de dimension finie ;
2. $\dim F \leqslant \dim E$;
3. $F = E \iff \dim F = \dim E$.

Démonstration

- Soit E un espace vectoriel de dimension n et soit F un sous-espace vectoriel de E. Si $F = \{0\}$ il n'y a rien à montrer. On suppose donc $F \neq \{0\}$ et soit v un élément non nul de F. La famille $\{v\}$ est une famille libre de F, donc F contient des familles libres. Toute famille libre d'éléments de F étant une famille libre d'éléments de E (voir la définition des familles libres), alors comme E est de dimension n, toutes les familles libres de F ont au plus n éléments.
- On considère l'ensemble K des entiers k tels qu'il existe une famille libre de F ayant k éléments :

$$K = \Big\{k \in \mathbb{N} \mid \exists \{v_1, v_2, \dots, v_k\} \subset F \ \text{ et } \ \{v_1, v_2, \dots, v_k\} \text{ est une famille libre de } F\Big\}$$

Cet ensemble K est non vide (car $1 \in K$) ; K est un sous-ensemble borné de $\mathbb{N}$ (puisque tout élément de K est compris entre 1 et n) donc K admet un maximum. Notons p ce maximum et soit $\{v_1, v_2, \dots, v_p\}$ une famille libre de F ayant p éléments.
- Montrons que $\{v_1, v_2, \dots, v_p\}$ est aussi génératrice de F. Par l'absurde, s'il existe w un élément de F qui n'est pas dans $\mathrm{Vect}(v_1, \dots, v_p)$, alors la famille $\{v_1, \dots, v_p, w\}$ ne peut pas être libre (sinon p ne serait pas le maximum de K). La famille $\{v_1, \dots, v_p, w\}$ est donc liée, mais alors la relation de dépendance linéaire implique que $w \in \mathrm{Vect}(v_1, \dots, v_p)$, ce qui est une contradiction.
Conclusion : $(v_1, \dots, v_p)$ est une famille libre et génératrice, donc est une base de F.
- On a ainsi démontré simultanément que :
  - F est de dimension finie (puisque $(v_1, v_2, \dots, v_p)$ est une base de F).
  - Ainsi $\dim F = p$, donc $\dim F \leqslant \dim E$ (puisque toute famille libre de F a au plus n éléments).
  - De plus, lorsque $p = n$, le p-uplet $(v_1, v_2, \dots, v_p)$, qui est une base de F, est aussi une base de E (car $\{v_1, v_2, \dots, v_p\}$ est alors une famille libre de E ayant exactement n éléments, donc est une base de E). Tout élément de E s'écrit comme une combinaison linéaire de $v_1, v_2, \dots, v_p$, d'où $E = F$.

5.2. Exemples

Exemple 92

Si E est un $\mathbb{K}$-espace vectoriel de dimension 2, les sous-espaces vectoriels de E sont :

- soit de dimension 0 : c'est alors le sous-espace $\{0\}$;
- soit de dimension 1 : ce sont les droites vectorielles, c'est-à-dire les sous-espaces $\mathbb{K}u = \mathrm{Vect}\{u\}$ engendrés par les vecteurs non nuls u de E ;
- soit de dimension 2 : c'est alors l'espace E tout entier.

Vocabulaire. Plus généralement, dans un $\mathbb{K}$-espace vectoriel E de dimension n ($n \geqslant 2$), tout sous-espace vectoriel de E de dimension 1 est appelé ***droite vectorielle*** de E et tout sous-espace vectoriel de E de dimension 2 est appelé ***plan vectoriel*** de E. Tout sous-espace vectoriel de E de dimension $n-1$ est appelé ***hyperplan*** de E. Pour $n = 3$, un hyperplan est un plan vectoriel ; pour $n = 2$, un hyperplan est une droite vectorielle.

Le théorème 25 précédent permet de déduire le corollaire suivant :

Corollaire 5

Soit E un $\mathbb{K}$-espace vectoriel. Soient F et G deux sous-espaces vectoriels de E. On suppose que F est de dimension finie et que $G \subset F$. Alors :

$$F = G \iff \dim F = \dim G$$

Autrement dit, sachant qu'un sous-espace est inclus dans un autre, alors pour montrer qu'ils sont égaux il suffit de montrer l'égalité des dimensions.

Exemple 93

Deux droites vectorielles F et G sont soit égales, soit d'intersection réduite au vecteur nul.

Exemple 94

Soient les sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F = \left\{ \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right) \in \mathbb{R}^3 \mid 2x - 3y + z = 0 \right\} \quad \text{et} \quad G = \mathrm{Vect}(u, v) \quad \text{où } u = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right) \text{ et } v = \left(\begin{smallmatrix} 2 \\ 1 \\ -1 \end{smallmatrix}\right).$$

Est-ce que $F = G$?

1. On remarque que les vecteurs u et v ne sont pas colinéaires, donc G est de dimension 2, et de plus ils appartiennent à F, donc G est contenu dans F.
2. Pour trouver la dimension de F, on pourrait déterminer une base de F et on montrerait alors que la dimension de F est 2. Mais il est plus judicieux ici de remarquer que F est contenu strictement dans $\mathbb{R}^3$ (par exemple le vecteur $\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$ de $\mathbb{R}^3$ n'est pas dans F), donc $\dim F < \dim \mathbb{R}^3 = 3$; mais puisque F contient G alors $\dim F \geqslant \dim G = 2$, donc la dimension de F ne peut être que 2.
3. On a donc démontré que $G \subset F$ et que $\dim G = \dim F$, ce qui entraîne $G = F$.

5.3. Théorème des quatre dimensions

Théorème 26. Théorème des quatre dimensions

Soient E un espace vectoriel de dimension finie et F, G des sous-espaces vectoriels de E. Alors :

$$\boxed{\dim(F+G) = \dim F + \dim G - \dim(F \cap G)}$$

Corollaire 6

Si $E = F \oplus G$, alors $\dim E = \dim F + \dim G$.

Exemple 95

Dans un espace vectoriel E de dimension 6, on considère deux sous-espaces F et G avec $\dim F = 3$ et $\dim G = 4$. Que peut-on dire de $F \cap G$? de $F+G$? Peut-on avoir $F \oplus G = E$?

- $F \cap G$ est un sous-espace vectoriel inclus dans F, donc $\dim(F \cap G) \leqslant \dim F = 3$. Donc les dimensions possibles pour $F \cap G$ sont pour l'instant $0, 1, 2, 3$.
- $F+G$ est un sous-espace vectoriel contenant G et inclus dans E, donc $4 = \dim G \leqslant \dim(F+G) \leqslant \dim E = 6$. Donc les dimensions possibles pour $F+G$ sont $4, 5, 6$.
- Le théorème 26 des quatre dimensions nous donne la relation : $\dim(F \cap G) = \dim F + \dim G - \dim(F+G) = 3+4-\dim(F+G) = 7 - \dim(F+G)$. Comme $F+G$ est de dimension 4, 5 ou 6, alors la dimension de $F \cap G$ est 3, 2 ou 1.
- Conclusion : les dimensions possibles pour $F+G$ sont 4, 5 ou 6 ; les dimensions correspondantes pour $F \cap G$ sont alors 3, 2 ou 1. Dans tous les cas, $F \cap G \neq \{0\}$ et en particulier F et G ne sont jamais en somme directe dans E.

La méthode de la preuve du théorème 26 des quatre dimensions implique aussi :

Corollaire 7

Tout sous-espace vectoriel F d'un espace vectoriel E de dimension finie admet un supplémentaire.

Démonstration . Preuve du théorème 26

- Notez l'analogie de la formule avec la formule pour les ensembles finis :

$$\operatorname{Card}(A \cup B) = \operatorname{Card} A + \operatorname{Card} B - \operatorname{Card}(A \cap B).$$

- Nous allons partir d'une base $\mathscr{B}_{F\cap G} = \{u_1, \ldots, u_p\}$ de $F \cap G$. On commence par compléter $\mathscr{B}_{F\cap G}$ en une base $\mathscr{B}_F = \{u_1, \ldots, u_p, v_{p+1}, \ldots, v_q\}$ de F. On complète ensuite $\mathscr{B}_{F\cap G}$ en une base $\mathscr{B}_G = \{u_1, \ldots, u_p, w_{p+1}, \ldots, w_r\}$ de G.
- Nous allons maintenant montrer que la famille $\{u_1, \ldots, u_p, v_{p+1}, \ldots, v_q, w_{p+1}, \ldots, w_r\}$ est une base de $F+G$. Il est tout d'abord clair que c'est une famille génératrice de $F+G$ (car $\mathscr{B}_F$ est une famille génératrice de F et $\mathscr{B}_G$ est une famille génératrice de G).
- Montrons que cette famille est libre. Soit une combinaison linéaire nulle :

$$\sum_{i=1}^{p} \alpha_i u_i + \sum_{j=p+1}^{q} \beta_j v_j + \sum_{k=p+1}^{r} \gamma_k w_k = 0 \tag{5.1}$$

On pose $u = \sum_{i=1}^{p} \alpha_i u_i$, $v = \sum_{j=p+1}^{q} \beta_j v_j$, $w = \sum_{k=p+1}^{r} \gamma_k w_k$. Alors d'une part $u+v \in F$ (car $\mathscr{B}_F$ est une base de F) mais comme l'équation (5.1) équivaut à $u+v+w=0$, alors $u+v = -w \in G$

(car $w \in G$). Maintenant $u+v \in F \cap G$ et aussi bien sûr $u \in F \cap G$, donc $v = \sum_{j=p+1}^{q} \beta_j v_j \in F \cap G$. Cela implique $\beta_j = 0$ pour tout j (car les $\{v_j\}$ complètent la base de $F \cap G$).
La combinaison linéaire nulle (5.1) devient $\sum_{i=1}^{p} \alpha_i u_i + \sum_{k=p+1}^{r} \gamma_k w_k = 0$. Or $\mathscr{B}_G$ est une base de G, donc $\alpha_i = 0$ et $\gamma_k = 0$ pour tout i,k. Ainsi $\mathscr{B}_{F+G} = \{u_1,\dots,u_p,v_{p+1},\dots,v_q,w_{p+1},\dots,w_r\}$ est une base de $F+G$.

– Il ne reste plus qu'à compter le nombre de vecteurs de chaque base : $\dim F \cap G = \operatorname{Card} \mathscr{B}_{F\cap G} = p$, $\dim F = \operatorname{Card} \mathscr{B}_F = q$, $\dim G = \operatorname{Card} \mathscr{B}_G = r$, $\dim(F+G) = \operatorname{Card} \mathscr{B}_{F+G} = q+r-p$. Ce qui prouve bien $\dim(F+G) = \dim F + \dim G - \dim(F \cap G)$.

Mini-exercices

1. Soient $F = \operatorname{Vect}\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}3\\-1\\2\end{pmatrix}\right)$ et $G = \operatorname{Vect}\left(\begin{pmatrix}-7\\7\\0\end{pmatrix}, \begin{pmatrix}6\\5\\11\end{pmatrix}\right)$. Montrer que $F = G$.
2. Dans $\mathbb{R}^3$, on considère $F = \operatorname{Vect}\left(\begin{pmatrix}1\\t\\-1\end{pmatrix}, \begin{pmatrix}t\\1\\1\end{pmatrix}\right)$, $G = \operatorname{Vect}\begin{pmatrix}1\\1\\1\end{pmatrix}$. Calculer les dimensions de $F, G, F \cap G, F+G$ en fonction de $t \in \mathbb{R}$.
3. Dans un espace vectoriel de dimension 7, on considère des sous-espaces F et G vérifiant $\dim F = 3$ et $\dim G \leqslant 2$. Que peut-on dire pour $\dim(F \cap G)$? Et pour $\dim(F+G)$?
4. Dans un espace vectoriel E de dimension finie, montrer l'équivalence entre : (i) $F \oplus G = E$; (ii) $F+G=E$ et $\dim F + \dim G = \dim E$; (iii) $F \cap G = \{0_E\}$ et $\dim F + \dim G = \dim E$.
5. Soit H un hyperplan dans un espace vectoriel de dimension finie E. Soit $v \in E \setminus H$. Montrer que H et $\operatorname{Vect}(v)$ sont des sous-espaces supplémentaires dans E.

Auteurs

- D'après un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties d'un cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- et un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- mixé, révisé et complété par Arnaud Bodin. Relu par Vianney Combet.

6 Matrices et applications linéaires

Ce chapitre est l'aboutissement de toutes les notions d'algèbre linéaire vues jusqu'ici : espaces vectoriels, dimension, applications linéaires, matrices. Nous allons voir que dans le cas des espaces vectoriels de dimension finie, l'étude des applications linéaires se ramène à l'étude des matrices, ce qui facilite les calculs.

1. Rang d'une famille de vecteurs

Le rang d'une famille de vecteurs est la dimension du plus petit sous-espace vectoriel contenant tous ces vecteurs.

1.1. Définition

Soient E un $\mathbb{K}$-espace vectoriel et $\{v_1,\dots,v_p\}$ une famille finie de vecteurs de E. Le sous-espace vectoriel $\operatorname{Vect}(v_1,\dots,v_p)$ engendré par $\{v_1,\dots,v_p\}$ étant de dimension finie, on peut donc donner la définition suivante :

Définition 35. Rang d'une famille finie de vecteurs

Soit E un $\mathbb{K}$-espace vectoriel et soit $\{v_1,\dots,v_p\}$ une famille finie de vecteurs de E. Le ***rang*** de la famille $\{v_1,\dots,v_p\}$ est la dimension du sous-espace vectoriel $\operatorname{Vect}(v_1,\dots,v_p)$ engendré par les vecteurs $v_1,\dots,v_p$. Autrement dit :

$$\boxed{\operatorname{rg}(v_1,\dots,v_p) = \dim \operatorname{Vect}(v_1,\dots,v_p)}$$

Calculer le rang d'une famille de vecteurs n'est pas toujours évident, cependant il y a des inégalités qui découlent directement de la définition.

Proposition 28

Soient E un $\mathbb{K}$-espace vectoriel et $\{v_1,\dots,v_p\}$ une famille de p vecteurs de E. Alors :

1. $0 \leqslant \operatorname{rg}(v_1,\dots,v_p) \leqslant p$: le rang est inférieur ou égal au nombre d'éléments dans la famille.
2. Si E est de dimension finie alors $\operatorname{rg}(v_1,\dots,v_p) \leqslant \dim E$: le rang est inférieur ou égal à la dimension de l'espace ambiant E.

Remarque

- Le rang d'une famille vaut 0 si et seulement si tous les vecteurs sont nuls.
- Le rang d'une famille $\{v_1,\dots,v_p\}$ vaut p si et seulement si la famille $\{v_1,\dots,v_p\}$ est libre.

Exemple 96

Quel est le rang de la famille $\{v_1,v_2,v_3\}$ suivante dans l'espace vectoriel $\mathbb{R}^4$?

$$v_1=\begin{pmatrix}1\\0\\1\\0\end{pmatrix}\qquad v_2=\begin{pmatrix}0\\1\\1\\1\end{pmatrix}\qquad v_3=\begin{pmatrix}-1\\1\\0\\1\end{pmatrix}$$

- Ce sont des vecteurs de $\mathbb{R}^4$ donc $\operatorname{rg}(v_1,v_2,v_3)\leqslant 4$.
- Mais comme il n'y a que 3 vecteurs alors $\operatorname{rg}(v_1,v_2,v_3)\leqslant 3$.
- Le vecteur v_1 est non nul donc $\operatorname{rg}(v_1,v_2,v_3)\geqslant 1$.
- Il est clair que v_1 et v_2 sont linéairement indépendants donc $\operatorname{rg}(v_1,v_2,v_3)\geqslant\operatorname{rg}(v_1,v_2)=2$.

Il reste donc à déterminer si le rang vaut 2 ou 3. On cherche si la famille $\{v_1,v_2,v_3\}$ est libre ou liée en résolvant le système linéaire $\lambda_1v_1+\lambda_2v_2+\lambda_3v_3=0$. On trouve $v_1-v_2+v_3=0$. La famille est donc liée. Ainsi $\operatorname{Vect}(v_1,v_2,v_3)=\operatorname{Vect}(v_1,v_2)$, donc $\operatorname{rg}(v_1,v_2,v_3)=\dim\operatorname{Vect}(v_1,v_2,v_3)=2$.

1.2. Rang d'une matrice

Une matrice peut être vue comme une juxtaposition de vecteurs colonnes.

Définition 36

On définit le rang d'une matrice comme étant le rang de ses vecteurs colonnes.

Exemple 97

Le rang de la matrice

$$A=\begin{pmatrix}1&2&-\frac12&0\\2&4&-1&0\end{pmatrix}\in\mathscr{M}_{2,4}$$

est par définition le rang de la famille de vecteurs de $\mathbb{K}^2$: $\Big\{v_1=\binom{1}{2},v_2=\binom{2}{4},v_3=\binom{-\frac12}{-1},v_4=\binom{0}{0}\Big\}$. Tous ces vecteurs sont colinéaires à v_1, donc le rang de la famille $\{v_1,v_2,v_3,v_4\}$ est 1 et ainsi $\operatorname{rg}A=1$.

Réciproquement, on se donne une famille de p vecteurs $\{v_1,\dots,v_p\}$ d'un espace vectoriel E de dimension n. Fixons une base $\mathscr{B}=\{e_1,\dots,e_n\}$ de E. Chaque vecteur v_j se décompose dans la base $\mathscr{B}$: $v_j=a_{1j}e_1+\cdots+a_{ij}e_i+\cdots+a_{nj}e_n$, ce que l'on note $v_j=\begin{pmatrix}a_{1j}\\\vdots\\a_{ij}\\\vdots\\a_{nj}\end{pmatrix}_{\mathscr{B}}$. En juxtaposant ces vecteurs colonnes, on obtient une matrice $A\in\mathscr{M}_{n,p}(\mathbb{K})$. Le rang de la famille $\{v_1,\dots,v_p\}$ est égal au rang de la matrice A.

Définition 37

On dit qu'une matrice est échelonnée par rapport aux colonnes si le nombre de zéros commençant une colonne croît strictement colonne après colonne, jusqu'à ce qu'il ne reste plus que des zéros. Autrement dit, la matrice transposée est échelonnée par rapport aux lignes.

Voici un exemple d'une matrice échelonnée par colonnes ; les $*$ désignent des coefficients quelconques, les $+$ des coefficients non nuls :

$$\begin{pmatrix} + & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ * & + & 0 & 0 & 0 & 0 \\ * & * & + & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & + & 0 & 0 \end{pmatrix}$$

Le rang d'une matrice échelonnée est très simple à calculer.

Proposition 29

Le rang d'une matrice échelonnée par colonnes est égal au nombre de colonnes non nulles.

Par exemple, dans la matrice échelonnée donnée en exemple ci-dessus, 4 colonnes sur 6 sont non nulles, donc le rang de cette matrice est 4.
La preuve de cette proposition consiste à remarquer que les vecteurs colonnes non nuls sont linéairement indépendants, ce qui au vu de la forme échelonnée de la matrice est facile.

1.3. Opérations conservant le rang

Proposition 30

Le rang d'une matrice ayant les colonnes $C_1, C_2, \dots, C_p$ n'est pas modifié par les trois opérations élémentaires suivantes sur les vecteurs :

1. $C_i \leftarrow \lambda C_i$ avec $\lambda \neq 0$: on peut multiplier une colonne par un scalaire non nul.
2. $C_i \leftarrow C_i + \lambda C_j$ avec $\lambda \in \mathbb{K}$ (et $j \neq i$) : on peut ajouter à la colonne C_i un multiple d'une autre colonne C_j.
3. $C_i \leftrightarrow C_j$: on peut échanger deux colonnes.

Plus généralement, l'opération $C_i \leftarrow C_i + \lambda_1 C_1 + \lambda_2 C_2 + \cdots + \lambda_p C_p$ conserve le rang de la matrice.
On a même un résultat plus fort, comme vous le verrez dans la preuve : l'espace vectoriel engendré par les vecteurs colonnes est conservé par ces opérations.

Démonstration

Le premier et troisième point de la proposition sont faciles.
Pour simplifier l'écriture de la démonstration du deuxième point, montrons que l'opération $C_1 \leftarrow C_1 + \lambda C_2$ ne change pas le rang. Notons v_i le vecteur correspondant à la colonne C_i d'une matrice A.
L'opération sur les colonnes $C_1 \leftarrow C_1 + \lambda C_2$ change la matrice A en une matrice A' dont les vecteurs colonnes sont : $v_1 + \lambda v_2, v_2, v_3, \dots, v_p$.
Il s'agit de montrer que les sous-espaces $F = \mathrm{Vect}(v_1, v_2, \dots, v_p)$ et $G = \mathrm{Vect}(v_1 + \lambda v_2, v_2, v_3, \dots, v_p)$ ont la même dimension. Nous allons montrer qu'ils sont égaux !

- Tout générateur de G est une combinaison linéaire des v_i, donc $G \subset F$.

– Pour montrer que $F \subset G$, il suffit de montrer v_1 est combinaison linéaire des générateurs de G, ce qui s'écrit : $v_1 = (v_1 + \lambda v_2) - \lambda v_2$.

Conclusion : $F = G$ et donc $\dim F = \dim G$.

Méthodologie. Comment calculer le rang d'une matrice ou d'un système de vecteurs ?
Il s'agit d'appliquer la méthode de Gauss sur les colonnes de la matrice A (considérée comme une juxtaposition de vecteurs colonnes). Le principe de la méthode de Gauss affirme que par les opérations élémentaires $C_i \leftarrow \lambda C_i$, $C_i \leftarrow C_i + \lambda C_j$, $C_i \leftrightarrow C_j$, on transforme la matrice A en une matrice échelonnée par rapport aux colonnes. Le rang de la matrice est alors le nombre de colonnes non nulles.
Remarque : la méthode de Gauss classique concerne les opérations sur les lignes et aboutit à une matrice échelonnée par rapport aux lignes. Les opérations sur les colonnes de A correspondent aux opérations sur les lignes de la matrice transposée A^T.

1.4. Exemples

Exemple 98

Quel est le rang de la famille des 5 vecteurs suivants de $\mathbb{R}^4$?

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \quad v_3 = \begin{pmatrix} 3 \\ 2 \\ -1 \\ -3 \end{pmatrix} \quad v_4 = \begin{pmatrix} 3 \\ 5 \\ 0 \\ -1 \end{pmatrix} \quad v_5 = \begin{pmatrix} 3 \\ 8 \\ 1 \\ 1 \end{pmatrix}$$

On est ramené à calculer le rang de la matrice :

$$\begin{pmatrix} 1 & -1 & 3 & 3 & 3 \\ 1 & 2 & 2 & 5 & 8 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & 1 & -3 & -1 & 1 \end{pmatrix}$$

En faisant les opérations $C_2 \leftarrow C_2 + C_1$, $C_3 \leftarrow C_3 - 3C_1$, $C_4 \leftarrow C_4 - 3C_1$, $C_5 \leftarrow C_5 - 3C_1$, on obtient des zéros sur la première ligne à droite du premier pivot :

$$\begin{pmatrix} 1 & -1 & 3 & 3 & 3 \\ 1 & 2 & 2 & 5 & 8 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & 1 & -3 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & -1 & 2 & 5 \\ 1 & 1 & -4 & -3 & -2 \\ 1 & 2 & -6 & -4 & -2 \end{pmatrix}$$

On échange C_2 et C_3 par l'opération $C_2 \leftrightarrow C_3$ pour avoir le coefficient -1 en position de pivot et ainsi éviter d'introduire des fractions.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & -1 & 2 & 5 \\ 1 & 1 & -4 & -3 & -2 \\ 1 & 2 & -6 & -4 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & 2 & 5 \\ 1 & -4 & 1 & -3 & -2 \\ 1 & -6 & 2 & -4 & -2 \end{pmatrix}$$

En faisant les opérations $C_3 \leftarrow C_3 + 3C_2$, $C_4 \leftarrow C_4 + 2C_2$ et $C_5 \leftarrow C_5 + 5C_2$, on obtient des zéros à droite de ce deuxième pivot :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & 2 & 5 \\ 1 & -4 & 1 & -3 & -2 \\ 1 & -6 & 2 & -4 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & -11 & -22 \\ 1 & -6 & -16 & -16 & -32 \end{pmatrix}$$

Enfin, en faisant les opérations $C_4 \leftarrow C_4 - C_3$ et $C_5 \leftarrow C_5 - 2C_3$, on obtient une matrice échelonnée par colonnes :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & -11 & -22 \\ 1 & -6 & -16 & -16 & -32 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & 0 & 0 \\ 1 & -6 & -16 & 0 & 0 \end{pmatrix}$$

Il y a 3 colonnes non nulles : on en déduit que le rang de la famille de vecteurs $\{v_1, v_2, v_3, v_4, v_5\}$ est 3.

En fait, nous avons même démontré que

$$\mathrm{Vect}(v_1, v_2, v_3, v_4, v_5) = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -4 \\ -6 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -11 \\ -16 \end{pmatrix} \right).$$

Exemple 99

Considérons les trois vecteurs suivants dans $\mathbb{R}^5$: $v_1 = (1,2,1,2,0)$, $v_2 = (1,0,1,4,4)$ et $v_3 = (1,1,1,0,0)$. Montrons que la famille $\{v_1, v_2, v_3\}$ est libre dans $\mathbb{R}^5$. Pour cela, calculons le rang de cette famille de vecteurs ou, ce qui revient au même, celui de la matrice suivante :

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

Par des opérations élémentaires sur les colonnes, on obtient :

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \\ 2 & 4 & 0 \\ 0 & 4 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & -2 & -1 \\ 1 & 0 & 0 \\ 2 & 2 & -2 \\ 0 & 4 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & -1 \\ 1 & 0 & 0 \\ 2 & 1 & -2 \\ 0 & 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & -3 \\ 0 & 2 & -2 \end{pmatrix}$$

Comme la dernière matrice est échelonnée par colonnes et que ses 3 colonnes sont non nulles, on en déduit que la famille $\{v_1, v_2, v_3\}$ constituée de 3 vecteurs est de rang 3, et donc qu'elle est libre dans $\mathbb{R}^5$.

Exemple 100

Considérons les quatre vecteurs suivants dans $\mathbb{R}^3$: $v_1 = (1,2,3)$, $v_2 = (2,0,6)$, $v_3 = (3,2,1)$ et $v_4 = (-1,2,2)$. Montrons que la famille $\{v_1, v_2, v_3, v_4\}$ engendre $\mathbb{R}^3$. Pour cela, calculons le rang de cette famille de vecteurs ou, ce qui revient au même, celui de la matrice suivante :

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 2 & 2 \\ 3 & 6 & 1 & 2 \end{pmatrix}.$$

Par des opérations élémentaires sur les colonnes, on obtient :

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 2 & 2 \\ 3 & 6 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & -4 & 4 \\ 3 & 0 & -8 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & 0 & -8 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & 0 & -8 & 0 \end{pmatrix}.$$

La famille $\{v_1, v_2, v_3, v_4\}$ est donc de rang 3. Cela signifie que $\operatorname{Vect}(v_1, v_2, v_3, v_4)$ est un sous-espace vectoriel de dimension 3 de $\mathbb{R}^3$. On a donc $\operatorname{Vect}(v_1, v_2, v_3, v_4) = \mathbb{R}^3$. Autrement dit, la famille $\{v_1, v_2, v_3, v_4\}$ engendre $\mathbb{R}^3$.

1.5. Rang et matrice inversible

Nous anticipons sur la suite, pour énoncer un résultat important :

Théorème 27. Matrice inversible et rang

Une matrice carrée de taille n est inversible si et seulement si elle est de rang n.

La preuve repose sur plusieurs résultats qui seront vus au fil de ce chapitre.

Démonstration

Soit A une matrice carrée d'ordre n. Soit f l'endomorphisme de $\mathbb{K}^n$ dont la matrice dans la base canonique est A. On a les équivalences suivantes :

$$\begin{array}{rcl} A \text{ de rang } n & \iff & f \text{ de rang } n \\ & \iff & f \text{ surjective} \\ & \iff & f \text{ bijective} \\ & \iff & A \text{ inversible.} \end{array}$$

Nous avons utilisé le fait qu'un endomorphisme d'un espace vectoriel de dimension finie est bijectif si et seulement s'il est surjectif et le théorème sur la caractérisation de la matrice d'un isomorphisme.

1.6. Rang engendré par les vecteurs lignes

On a considéré jusqu'ici une matrice $A \in \mathcal{M}_{n,p}$ comme une juxtaposition de vecteurs colonnes $(v_1, \dots, v_p)$ et défini $\operatorname{rg} A = \dim \operatorname{Vect}(v_1, \dots, v_p)$. Considérons maintenant que A est aussi une superposition de vecteurs lignes $(w_1, \dots, w_n)$.

Proposition 31

$\operatorname{rg} A = \dim \operatorname{Vect}(w_1,\ldots,w_n)$

Nous admettrons ce résultat. Autrement dit : *l'espace vectoriel engendré par les vecteurs colonnes et l'espace vectoriel engendré par les vecteurs lignes sont de même dimension.*
Une formulation plus théorique est que *le rang d'une matrice égale le rang de sa transposée* :

$$\boxed{\operatorname{rg} A = \operatorname{rg} A^T}$$

Attention ! Les dimensions $\dim \operatorname{Vect}(v_1,\ldots,v_p)$ et $\dim \operatorname{Vect}(w_1,\ldots,w_n)$ sont égales, mais les espaces vectoriels $\operatorname{Vect}(v_1,\ldots,v_p)$ et $\operatorname{Vect}(w_1,\ldots,w_n)$ ne sont pas les mêmes.

Mini-exercices

1. Quel est le rang de la famille de vecteurs $\left(\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}3\\4\\2\end{pmatrix},\begin{pmatrix}0\\-2\\-1\end{pmatrix},\begin{pmatrix}2\\2\\1\end{pmatrix}\right)$?
   Même question pour $\left(\begin{pmatrix}1\\t\\1\end{pmatrix},\begin{pmatrix}t\\1\\t\end{pmatrix},\begin{pmatrix}1\\1\\t\end{pmatrix}\right)$ en fonction du paramètre $t \in \mathbb{R}$.
2. Mettre sous forme échelonnée par rapport aux colonnes la matrice $\begin{pmatrix}1 & 2 & -4 & -2 & -1\\ 0 & -2 & 4 & 2 & 0\\ 1 & 1 & -2 & -1 & 1\end{pmatrix}$.
   Calculer son rang. Idem avec $\begin{pmatrix}1 & 7 & 2 & 5\\ -2 & 1 & 1 & 5\\ -1 & 2 & 1 & 4\\ 1 & 4 & 1 & 2\end{pmatrix}$.
3. Calculer le rang de $\begin{pmatrix}2 & 4 & -5 & -7\\ -1 & 3 & 1 & 2\\ 1 & a & -2 & b\end{pmatrix}$ en fonction de a et b.
4. Calculer les rangs précédents en utilisant les vecteurs lignes.
5. Soit $f : E \to F$ une application linéaire. Quelle inégalité relie $\operatorname{rg}(f(v_1),\ldots,f(v_p))$ et $\operatorname{rg}(v_1,\ldots,v_p)$? Que se passe-t-il si f est injective ?

2. Applications linéaires en dimension finie

Lorsque $f : E \to F$ est une application linéaire et que E est de dimension finie, la théorie de la dimension fournit de nouvelles propriétés très riches pour l'application linéaire f.

2.1. Construction et caractérisation

Une application linéaire $f : E \to F$, d'un espace vectoriel de dimension finie dans un espace vectoriel quelconque, est entièrement déterminée par les images des vecteurs d'une base de l'espace vectoriel E de départ. C'est ce qu'affirme le théorème suivant :

Théorème 28. Construction d'une application linéaire

Soient E et F deux espaces vectoriels sur un même corps $\mathbb{K}$. On suppose que l'espace vectoriel E est de dimension finie n et que $(e_1,\dots,e_n)$ est une base de E. Alors pour tout choix $(v_1,\dots,v_n)$ de n vecteurs de F, il existe une et une seule application linéaire $f : E \to F$ telle que, pour tout $i = 1,\dots,n$:

$$f(e_i) = v_i.$$

Le théorème ne fait aucune hypothèse sur la dimension de l'espace vectoriel d'arrivée F.

Exemple 101

Il existe une unique application linéaire $f : \mathbb{R}^n \to \mathbb{R}[X]$ telle que $f(e_i) = (X+1)^i$ pour $i = 1,\dots,n$ (où $(e_1,\dots,e_n)$ est la base canonique de $\mathbb{R}^n$).
Pour un vecteur $x = (x_1,\dots,x_n)$, on a

$$f(x_1,\dots,x_n) = f(x_1e_1+\cdots+x_ne_n) = x_1f(e_1)+\cdots+x_nf(e_n) = \sum_{i=1}^n x_i(X+1)^i.$$

Démonstration

- *Unicité.* Supposons qu'il existe une application linéaire $f : E \to F$ telle que $f(e_i) = v_i$, pour tout $i = 1,\dots,n$. Pour $x \in E$, il existe des scalaires $x_1, x_2,\dots,x_n$ uniques tels que $x = \sum_{i=1}^n x_ie_i$. Comme f est linéaire, on a

$$f(x) = f\left(\sum_{i=1}^n x_ie_i\right) = \sum_{i=1}^n x_if(e_i) = \sum_{i=1}^n x_iv_i. \qquad (*)$$

  Donc, si elle existe, f est unique.
- *Existence.* Nous venons de voir que s'il existe une solution c'est nécessairement l'application définie par l'équation $(*)$. Montrons qu'une application définie par l'équation $(*)$ est linéaire et vérifie $f(e_i) = v_i$. Si $(x_1,\dots,x_n)$ (resp. $y = (y_1,\dots,y_n)$) sont les coordonnées de x (resp. y) dans la base $(e_1,\dots,e_n)$, alors

$$f(\lambda x+\mu y) = f\left(\sum_{i=1}^n(\lambda x_i+\mu y_i)e_i\right) = \sum_{i=1}^n(\lambda x_i+\mu y_i)f(e_i) = \lambda\sum_{i=1}^n x_if(e_i)+\mu\sum_{i=1}^n y_if(e_i) = \lambda f(x)+\mu f(y).$$

  Enfin les coordonnées de e_i sont $(0,\dots,0,1,0,\dots,0)$ (avec un 1 en i-ème position), donc $f(e_i) = 1\cdot v_i = v_i$. Ce qui termine la preuve du théorème.

2.2. Rang d'une application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels et soit $f : E \to F$ une application linéaire. On rappelle que l'on note $f(E)$ par $\operatorname{Im} f$, c'est-à-dire $\operatorname{Im} f = \{f(x) | x \in E\}$. $\operatorname{Im} f$ est un sous-espace vectoriel de F.

Proposition 32

Si E est de dimension finie, alors :

- $\operatorname{Im} f = f(E)$ est un espace vectoriel de dimension finie.
- Si $(e_1,\dots,e_n)$ est une base de E, alors $\operatorname{Im} f = \operatorname{Vect}\big(f(e_1),\dots,f(e_n)\big)$.

La dimension de cet espace vectoriel $\operatorname{Im} f$ est appelée ***rang de f*** :

$$\boxed{\operatorname{rg}(f) = \dim \operatorname{Im} f = \dim \operatorname{Vect}\big(f(e_1),\dots,f(e_n)\big)}$$

Démonstration

Il suffit de démontrer que tout élément de $\operatorname{Im} f$ est combinaison linéaire des vecteurs $f(e_1),\dots,f(e_n)$. Soit y un élément quelconque de $\operatorname{Im} f$. Il existe donc un élément x de E tel que $y = f(x)$. Comme $(e_1,\dots,e_n)$ est une base de E, il existe des scalaires $(x_1,\dots,x_n)$ tels que $x = \sum_{i=1}^n x_i e_i$. En utilisant la linéarité de f, on en déduit que $y = f(x) = \sum_{i=1}^n x_i f(e_i)$, ce qui achève la démonstration.

Le rang est plus petit que la dimension de E et aussi plus petit que la dimension de F, si F est de dimension finie :

Proposition 33

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f : E \to F$ une application linéaire. On a

$$\operatorname{rg}(f) \leqslant \min(\dim E, \dim F).$$

Exemple 102

Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ l'application linéaire définie par $f(x,y,z) = (3x-4y+2z, 2x-3y-z)$. Quel est le rang de f ?

Si on note $e_1 = \left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, $e_2 = \left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$ et $e_3 = \left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right)$, alors (e_1,e_2,e_3) est la base canonique de $\mathbb{R}^3$.

Il s'agit de trouver le rang de la famille $\{v_1,v_2,v_3\}$:

$$v_1 = f(e_1) = f\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right) = \left(\begin{smallmatrix}3\\2\end{smallmatrix}\right) \qquad v_2 = f(e_2) = f\left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right) = \left(\begin{smallmatrix}-4\\-3\end{smallmatrix}\right) \qquad v_3 = f(e_3) = f\left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right) = \left(\begin{smallmatrix}2\\-1\end{smallmatrix}\right)$$

ou, ce qui revient au même, trouver le rang de la matrice

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

Commençons par estimer le rang sans faire de calculs.

- Nous avons une famille de 3 vecteurs donc $\operatorname{rg} f \leqslant 3$.
- Mais en fait les vecteurs v_1,v_2,v_3 vivent dans un espace de dimension 2 donc $\operatorname{rg} f \leqslant 2$.
- f n'est pas l'application linéaire nulle (autrement dit v_1,v_2,v_3 ne sont pas tous nuls) donc $\operatorname{rg} f \geqslant 1$.

Donc le rang de f vaut 1 ou 2. Il est facile de voir que v_1 et v_2 sont linéairement indépendants, donc le rang est 2 :

$$\operatorname{rg} f = \operatorname{rg}\big(f(e_1), f(e_2), f(e_3)\big) = \dim \operatorname{Vect}(v_1,v_2,v_3) = 2$$

Remarque : il est encore plus facile de voir que le rang de la matrice A est 2 en remarquant que ses deux seules lignes ne sont pas colinéaires.

2.3. Théorème du rang

Le théorème du rang est un résultat fondamental dans la théorie des applications linéaires en dimension finie.
On se place toujours dans la même situation :

- $f : E \to F$ est une application linéaire entre deux $\mathbb{K}$-espaces vectoriels,
- E est un espace vectoriel de dimension finie,
- le ***noyau*** de f est $\operatorname{Ker} f = \{x \in E \mid f(x) = 0_F\}$; c'est un sous-espace vectoriel de E, donc $\operatorname{Ker} f$ est de dimension finie,
- l'***image*** de f est $\operatorname{Im} f = f(E) = \{f(x) \mid x \in E\}$; c'est un sous-espace vectoriel de F et est de dimension finie.

Le théorème du rang donne une relation entre la dimension du noyau et la dimension de l'image de f.

Théorème 29. Théorème du rang

Soit $f : E \to F$ une application linéaire entre deux $\mathbb{K}$-espaces vectoriels, E étant de dimension finie. Alors

$$\boxed{\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f}$$

Autrement dit : $\boxed{\dim E = \dim \operatorname{Ker} f + \operatorname{rg} f}$
Dans la pratique, cette formule sert à déterminer la dimension du noyau connaissant le rang, ou bien le rang connaissant la dimension du noyau.

Exemple 103

Soit l'application linéaire

$$\begin{array}{rccc} f & : & \mathbb{R}^4 & \longrightarrow & \mathbb{R}^3 \\ & & (x_1,x_2,x_3,x_4) & \longmapsto & (x_1-x_2+x_3, 2x_1+2x_2+6x_3+4x_4, -x_1-2x_3-x_4) \end{array}$$

Calculons le rang de f et la dimension du noyau de f.

- **Première méthode.** On calcule d'abord le noyau :

$$(x_1,x_2,x_3,x_4) \in \operatorname{Ker} f \iff f(x_1,x_2,x_3,x_4) = (0,0,0) \iff \left\{\begin{array}{rcrcrcrcl} x_1 & - & x_2 & + & x_3 & & & = & 0 \\ 2x_1 & + & 2x_2 & + & 6x_3 & + & 4x_4 & = & 0 \\ -x_1 & & & - & 2x_3 & - & x_4 & = & 0 \end{array}\right.$$

On résout ce système et on trouve qu'il est équivalent à

$$\left\{\begin{array}{rcrcrcrcl} x_1 & - & x_2 & + & x_3 & & & = & 0 \\ & & x_2 & + & x_3 & + & x_4 & = & 0 \end{array}\right.$$

On choisit x_3 et x_4 comme paramètres et on trouve :

$$\operatorname{Ker} f = \left\{(-2x_3-x_4, -x_3-x_4, x_3, x_4) \mid x_3,x_4 \in \mathbb{R}\right\} = \left\{x_3\begin{pmatrix}-2\\-1\\1\\0\end{pmatrix} + x_4\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} \mid x_3,x_4 \in \mathbb{R}\right\} = \operatorname{Vect}\left(\begin{pmatrix}-2\\-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}\right)$$

Les deux vecteurs définissant le noyau sont linéairement indépendants, donc $\dim \operatorname{Ker} f = 2$.
On applique maintenant le théorème du rang pour en déduire sans calculs la dimension de l'image : $\dim \operatorname{Im} f = \dim \mathbb{R}^4 - \dim \operatorname{Ker} f = 4 - 2 = 2$. Donc le rang de f est 2.

– **Deuxième méthode.** On calcule d'abord l'image. On note (e_1, e_2, e_3, e_4) la base canonique de $\mathbb{R}^4$. Calculons $v_i = f(e_i)$:

$$v_1 = f(e_1) = f\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \begin{pmatrix}1\\2\\-1\end{pmatrix} \quad v_2 = f(e_2) = f\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \begin{pmatrix}-1\\2\\0\end{pmatrix}$$

$$v_3 = f(e_3) = f\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\6\\-2\end{pmatrix} \quad v_4 = f(e_4) = f\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\4\\-1\end{pmatrix}$$

On réduit la matrice A, formée des vecteurs colonnes, sous une forme échelonnée :

$$A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 2 & 2 & 6 & 4 \\ -1 & 0 & -2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}$$

Donc le rang de A est 2, ainsi $\operatorname{rg} f = \dim \operatorname{Im} f = \dim \operatorname{Vect}\big(f(e_1), f(e_2), f(e_3), f(e_4)\big) = 2$.
Maintenant, par le théorème du rang, $\dim \operatorname{Ker} f = \dim \mathbb{R}^4 - \operatorname{rg} f = 4 - 2 = 2$.

On trouve bien sûr le même résultat par les deux méthodes.

Exemple 104

Soit l'application linéaire

$$\begin{array}{rcl} f : \mathbb{R}_n[X] & \longrightarrow & \mathbb{R}_n[X] \\ P(X) & \longmapsto & P''(X) \end{array}$$

où $P''(X)$ est la dérivée seconde de $P(X)$. Quel est le rang et la dimension du noyau de f ?

– **Première méthode.** On calcule d'abord le noyau :

$$P(X) \in \operatorname{Ker} f \iff f\big(P(X)\big) = 0 \iff P''(X) = 0 \iff P'(X) = a \iff P(X) = aX + b$$

où $a, b \in \mathbb{R}$ sont des constantes. Cela prouve que $\operatorname{Ker} f$ est engendré par les deux polynômes : 1 (le polynôme constant) et X. Ainsi $\operatorname{Ker} f = \operatorname{Vect}(1, X)$. Donc $\dim \operatorname{Ker} f = 2$. Par le théorème du rang, $\operatorname{rg} f = \dim \operatorname{Im} f = \dim \mathbb{R}_n[X] - \dim \operatorname{Ker} f = (n+1) - 2 = n - 1$.

– **Deuxième méthode.** On calcule d'abord l'image : $(1, X, X^2, \ldots, X^n)$ est une base de l'espace de départ $\mathbb{R}_n[X]$, donc $\operatorname{rg} f = \dim \operatorname{Im} f = \dim \operatorname{Vect}\big(f(1), f(X), \ldots, f(X^n)\big)$. Tout d'abord, $f(1) = 0$ et $f(X) = 0$. Pour $k \geqslant 2$, $f(X^k) = k(k-1)X^{k-2}$. Comme les degrés sont échelonnés, il est clair que $\big\{f(X^2), f(X^3), \ldots, f(X^n)\big\} = \big\{2, 6X, 12X^2, \ldots, n(n-1)X^{n-2}\big\}$ engendre un espace de dimension $n-1$, donc $\operatorname{rg} f = n - 1$. Par le théorème du rang, $\dim \operatorname{Ker} f = \dim \mathbb{R}_n[X] - \operatorname{rg} f = (n+1) - (n-1) = 2$.

Démonstration . Preuve du théorème du rang

– Premier cas : f est injective.
En désignant par $(e_1, \ldots, e_n)$ une base de E, nous avons vu que la famille à n éléments $\big(f(e_1), \ldots, f(e_n)\big)$ est une famille libre de F (car f est injective), donc une famille libre de $\operatorname{Im} f$. De plus, $\big\{f(e_1), \ldots, f(e_n)\big\}$ est une partie génératrice de $\operatorname{Im} f$. Donc $(f(e_1), \ldots, f(e_n))$ est une base de $\operatorname{Im} f$. Ainsi $\dim \operatorname{Im} f = n$, et comme f est injective, $\dim \operatorname{Ker} f = 0$, et ainsi le théorème du rang est vrai.

– Deuxième cas : f n'est pas injective.

Dans ce cas le noyau de f est un sous-espace de E de dimension p avec $1 \leqslant p \leqslant n$. Soit $(\varepsilon_1,\dots,\varepsilon_p)$ une base de $\operatorname{Ker} f$. D'après le théorème de la base incomplète, il existe $n-p$ vecteurs $\varepsilon_{p+1},\dots,\varepsilon_n$ de E tels que $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)$ soit une base de E.
Alors $\operatorname{Im} f$ est engendrée par les vecteurs $f(\varepsilon_1), f(\varepsilon_2),\dots,f(\varepsilon_n)$. Mais, comme pour tout i vérifiant $1 \leqslant i \leqslant p$ on a $f(\varepsilon_i)=0$, $\operatorname{Im} f$ est engendrée par les vecteurs $f(\varepsilon_{p+1}),\dots,f(\varepsilon_n)$.
Montrons que ces vecteurs forment une famille libre. Soient $\alpha_{p+1},\dots,\alpha_n$ des scalaires tels que

$$\alpha_{p+1} f(\varepsilon_{p+1}) + \cdots + \alpha_n f(\varepsilon_n) = 0.$$

Puisque f est linéaire, cette égalité équivaut à l'égalité $f\left(\alpha_{p+1}\varepsilon_{p+1} + \cdots + \alpha_n\varepsilon_n\right) = 0$, qui prouve que le vecteur $\alpha_{p+1}\varepsilon_{p+1} + \cdots + \alpha_n\varepsilon_n$ appartient au noyau de f. Il existe donc des scalaires $\lambda_1,\dots,\lambda_p$ tels que

$$\alpha_{p+1}\varepsilon_{p+1} + \cdots + \alpha_n\varepsilon_n = \lambda_1\varepsilon_1 + \cdots + \lambda_p\varepsilon_p.$$

Comme $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)$ est une base de E, les vecteurs $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$ sont linéairement indépendants et par conséquent pour tout $i=1,\dots,p$, $\lambda_i=0$, et pour tout $i=p+1,\dots,n$, $\alpha_i=0$. Les vecteurs $f(\varepsilon_{p+1}),\dots,f(\varepsilon_n)$ définissent donc bien une base de $\operatorname{Im} f$. Ainsi le sous-espace vectoriel $\operatorname{Im} f$ est de dimension $n-p$, ce qui achève la démonstration.

On remarquera le rôle essentiel joué par le théorème de la base incomplète dans cette démonstration.

2.4. Application linéaire entre deux espaces de même dimension

Rappelons qu'un ***isomorphisme*** est une application linéaire bijective. Un isomorphisme implique que les espaces vectoriels de départ et d'arrivée ont la même dimension.

Proposition 34

Soit $f : E \to F$ un isomorphisme d'espaces vectoriels. Si E (respectivement F) est de dimension finie, alors F (respectivement E) est aussi de dimension finie et on a $\dim E = \dim F$.

Démonstration

Si E est de dimension finie, alors comme f est surjective, $F = \operatorname{Im} f$, donc F est engendré par l'image d'une base de E. On a donc F de dimension finie et $\dim F \leqslant \dim E$. De même $f^{-1} : F \to E$ est un isomorphisme, donc $f^{-1}(F) = E$, ce qui prouve cette fois $\dim E \leqslant \dim F$.
Si c'est F qui est de dimension finie, on fait le même raisonnement avec f^{-1}.

Nous allons démontrer une sorte de réciproque, qui est extrêmement utile.

Théorème 30

Soit $f : E \to F$ une application linéaire avec E et F de dimension finie.
Supposons $\dim E = \dim F$. Alors les assertions suivantes sont équivalentes :
- (i) f est bijective
- (ii) f est injective
- (iii) f est surjective

Autrement dit, dans le cas d'une application linéaire entre deux espaces de ***même*** dimension, pour démontrer qu'elle est bijective, il suffit de démontrer l'une des deux propriétés : injectivité ou surjectivité.

Démonstration

C'est immédiat à partir du théorème du rang. En effet, la propriété f injective équivaut à $\operatorname{Ker} f = \{0\}$, donc d'après le théorème du rang, f est injective si et seulement si $\dim \operatorname{Im} f = \dim E$. D'après l'hypothèse sur l'égalité des dimensions de E et de F, ceci équivaut à $\dim \operatorname{Im} f = \dim F$. Cela équivaut donc à $\operatorname{Im} f = F$, c'est-à-dire f est surjective.

Exemple 105

Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (x-y, x+y)$. Une façon simple de montrer que l'application linéaire f est bijective est de remarquer que l'espace de départ et l'espace d'arrivée ont même dimension. Ensuite on calcule le noyau :

$$(x,y) \in \operatorname{Ker} f \iff f(x,y) = 0 \iff (x-y, x+y) = (0,0) \iff \left\{ \begin{array}{rcl} x+y & = & 0 \\ x-y & = & 0 \end{array} \right. \iff (x,y) = (0,0)$$

Ainsi $\operatorname{Ker} f = \{(0,0)\}$ est réduit au vecteur nul, ce qui prouve que f est injective et donc, par le théorème 30, que f est un isomorphisme.

Mini-exercices

1. Soit (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. Donner l'expression de $f(x,y,z)$ où $f : \mathbb{R}^3 \to \mathbb{R}^3$ est l'application linéaire qui envoie e_1 sur son opposé, qui envoie e_2 sur le vecteur nul et qui envoie e_3 sur la somme des trois vecteurs e_1, e_2, e_3.
2. Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par $f(x,y,z) = (x-2y-3z, 2y+3z)$. Calculer une base du noyau de f, une base de l'image de f et vérifier le théorème du rang.
3. Même question avec $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (-y+z, x+z, x+y)$.
4. Même question avec l'application linéaire $f : \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ qui à X^k associe X^{k-1} pour $1 \leqslant k \leqslant n$ et qui à 1 associe 0.
5. Lorsque c'est possible, calculer la dimension du noyau, le rang et dire si f peut être injective, surjective, bijective :
   - Une application linéaire surjective $f : \mathbb{R}^7 \to \mathbb{R}^4$.
   - Une application linéaire injective $f : \mathbb{R}^5 \to \mathbb{R}^8$.
   - Une application linéaire surjective $f : \mathbb{R}^4 \to \mathbb{R}^4$.
   - Une application linéaire injective $f : \mathbb{R}^6 \to \mathbb{R}^6$.

3. Matrice d'une application linéaire

Nous allons voir qu'il existe un lien étroit entre les matrices et les applications linéaires. À une matrice on associe naturellement une application linéaire. Et réciproquement, étant donné une application linéaire, et des bases pour les espaces vectoriels de départ et d'arrivée, on associe une matrice.

Dans cette section, tous les espaces vectoriels sont de dimension finie.

3.1. Matrice associée à une application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Soient p la dimension de E et $\mathscr{B} = (e_1,\dots,e_p)$ une base de E. Soient n la dimension de F et $\mathscr{B}' = (f_1,\dots,f_n)$ une base de F. Soit enfin $f : E \to F$ une application linéaire.

Les propriétés des applications linéaires entre deux espaces de dimension finie permettent d'affirmer que :

- l'application linéaire f est déterminée de façon unique par l'image d'une base de E, donc par les vecteurs $f(e_1), f(e_2), \dots, f(e_p)$.
- Pour $j \in \{1,\dots,p\}$, $f(e_j)$ est un vecteur de F et s'écrit de manière unique comme combinaison linéaire des vecteurs de la base $\mathscr{B}' = (f_1, f_2, \dots, f_n)$ de F.
  Il existe donc n scalaires uniques $a_{1,j}, a_{2,j}, \dots, a_{n,j}$ (parfois aussi notés $a_{1j}, a_{2j}, \dots, a_{nj}$) tels que
$$f(e_j) = a_{1,j} f_1 + a_{2,j} f_2 + \cdots + a_{n,j} f_n = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}_{\mathscr{B}'} .$$

Ainsi, l'application linéaire f est entièrement déterminée par les coefficients $(a_{i,j})_{(i,j)\in\{1,\dots,n\}\times\{1,\dots,p\}}$. Il est donc naturel d'introduire la définition suivante :

Définition 38

La ***matrice de l'application linéaire*** f par rapport aux bases $\mathscr{B}$ et $\mathscr{B}'$ est la matrice $(a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$ dont la j-ème colonne est constituée par les coordonnées du vecteur $f(e_j)$ dans la base $\mathscr{B}' = (f_1, f_2, \dots, f_n)$:

$$\begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_n \end{array} \begin{array}{c} \begin{array}{ccccc} f(e_1) & \dots & f(e_j) & \dots & f(e_p) \end{array} \\ \begin{pmatrix} a_{11} & & a_{1j} & \dots & a_{1p} \\ a_{21} & & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{nj} & \dots & a_{np} \end{pmatrix} \end{array}$$

En termes plus simples : c'est la matrice dont les vecteurs colonnes sont l'image par f des vecteurs de la base de départ $\mathscr{B}$, exprimée dans la base d'arrivée $\mathscr{B}'$. On note cette matrice $\mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$.

Remarque

- La taille de la matrice $\mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$ dépend uniquement de la dimension de E et de celle de F.
- Par contre, les coefficients de la matrice dépendent du choix de la base $\mathscr{B}$ de E et de la base $\mathscr{B}'$ de F.

Exemple 106

Soit f l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par

$$\begin{array}{rccl} f & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & & (x_1, x_2, x_3) & \longmapsto & (x_1 + x_2 - x_3, x_1 - 2x_2 + 3x_3) \end{array}$$

Il est utile d'identifier vecteurs lignes et vecteurs colonnes ; ainsi f peut être vue comme l'application $f : \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} x_1+x_2-x_3 \\ x_1-2x_2+3x_3 \end{smallmatrix}\right)$.

Soient $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et $\mathcal{B}' = (f_1, f_2)$ la base canonique de $\mathbb{R}^2$. C'est-à-dire :

$$e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} \quad e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \quad e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix} \qquad\qquad f_1 = \begin{pmatrix}1\\0\end{pmatrix} \quad f_2 = \begin{pmatrix}0\\1\end{pmatrix}$$

1. Quelle est la matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$?
   - On a $f(e_1) = f(1,0,0) = (1,1) = f_1 + f_2$. La première colonne de la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ est donc $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$.
   - De même $f(e_2) = f(0,1,0) = (1,-2) = f_1 - 2f_2$. La deuxième colonne de la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ est donc $\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)$.
   - Enfin $f(e_3) = f(0,0,1) = (-1,3) = -f_1 + 3f_2$. La troisième colonne de la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ est donc $\left(\begin{smallmatrix}-1\\3\end{smallmatrix}\right)$.

   Ainsi :
   $$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{pmatrix}1 & 1 & -1\\ 1 & -2 & 3\end{pmatrix}$$
2. On va maintenant changer la base de l'espace de départ et conserver celle de l'espace d'arrivée. Soient les vecteurs de $\mathbb{R}^3$
   $$\varepsilon_1 = \begin{pmatrix}1\\1\\0\end{pmatrix} \quad \varepsilon_2 = \begin{pmatrix}1\\0\\1\end{pmatrix} \quad \varepsilon_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}$$
   On montre facilement que $\mathcal{B}_0 = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une base de $\mathbb{R}^3$. Quelle est la matrice de f dans les bases $\mathcal{B}_0$ et $\mathcal{B}'$?
   $f(\varepsilon_1) = f(1,1,0) = (2,-1) = 2f_1 - f_2$, $f(\varepsilon_2) = f(1,0,1) = (0,4) = 4f_2$, $f(\varepsilon_3) = f(0,1,1) = (0,1) = f_2$, donc
   $$\mathrm{Mat}_{\mathcal{B}_0,\mathcal{B}'}(f) = \begin{pmatrix}2 & 0 & 0\\ -1 & 4 & 1\end{pmatrix}.$$
   Cet exemple illustre bien le fait que la matrice dépend du choix des bases.

3.2. Opérations sur les applications linéaires et les matrices

Proposition 35

Soient $f, g : E \to F$ deux applications linéaires et soient $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de F. Alors :
- $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f+g) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) + \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(g)$
- $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\lambda f) = \lambda\, \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$

Autrement dit, si on note :

$$A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) \qquad B = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(g) \qquad C = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f+g) \qquad D = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\lambda f)$$

Alors :

$$C = A + B \qquad\qquad D = \lambda A$$

Autrement dit : la matrice associée à la somme de deux applications linéaires est la somme des matrices (à condition de considérer la même base sur l'espace de départ pour les deux applications et la même base sur l'espace d'arrivée). Idem avec le produit par un scalaire.

Ce qui est le plus important va être la composition des applications linéaires.

Proposition 36

Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires et soient $\mathscr{B}$ une base de E, $\mathscr{B}'$ une base de F et $\mathscr{B}''$ une base de G. Alors :

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

Autrement dit, si on note :

$$A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) \qquad B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}''}(g) \qquad C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f)$$

Alors

$$C = B \times A$$

Autrement dit, à condition de bien choisir les bases, la matrice associée à la composition de deux applications linéaires est le produit des matrices associées à chacune d'elles, dans le même ordre. En fait, le produit de matrices, qui semble compliqué au premier abord, est défini afin de correspondre à la composition des applications linéaires.

Démonstration

Posons $p = \dim(E)$ et $\mathscr{B} = (e_1, \dots, e_p)$ une base de E ; $n = \dim F$ et $\mathscr{B}' = (f_1, \dots, f_n)$ une base de F ; $q = \dim G$ et $\mathscr{B}'' = (g_1, \dots, g_q)$ une base de G. Écrivons $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) = (a_{ij}) \in \mathscr{M}_{n,p}$ la matrice de f, $B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}''}(g) = (b_{ij}) \in \mathscr{M}_{q,n}$ la matrice de g, $C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f) = (c_{ij}) \in \mathscr{M}_{q,p}$ la matrice de $g \circ f$.
On a

$$\begin{aligned}
(g \circ f)(e_1) &= g\big(f(e_1)\big) \\
&= g(a_{11} f_1 + \cdots + a_{n1} f_n) \\
&= a_{11} g(f_1) + \cdots + a_{n1} g(f_n) \\
&= a_{11}\Big(b_{11} g_1 + \cdots + b_{q1} g_q\Big) + \cdots + a_{n1}\Big(b_{1n} g_1 + \cdots + b_{qn} g_q\Big)
\end{aligned}$$

Ainsi, la première colonne de $C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f)$ est

$$\begin{pmatrix} a_{11} b_{11} + \cdots + a_{n1} b_{1n} \\ a_{11} b_{21} + \cdots + a_{n1} b_{2n} \\ \vdots \\ a_{11} b_{q1} + \cdots + a_{n1} b_{qn} \end{pmatrix}.$$

Mais ceci est aussi la première colonne de la matrice BA. En faisant la même chose avec les autres colonnes, on remarque que $C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f)$ et BA sont deux matrices ayant leurs colonnes égales. On a donc bien l'égalité cherchée.

Exemple 107

On considère deux applications linéaires : $f : \mathbb{R}^2 \to \mathbb{R}^3$ et $g : \mathbb{R}^3 \to \mathbb{R}^2$. On pose $E = \mathbb{R}^2$, $F = \mathbb{R}^3$, $G = \mathbb{R}^2$ avec $f : E \to F$, $g : F \to G$. On se donne des bases : $\mathscr{B} = (e_1, e_2)$ une base de E, $\mathscr{B}' = (f_1, f_2, f_3)$ une base de F, et $\mathscr{B}'' = (g_1, g_2)$ une base de G.
On suppose connues les matrices de f et g :

$$A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \in \mathscr{M}_{3,2} \qquad B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}''}(g) = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 1 & 2 \end{pmatrix} \in \mathscr{M}_{2,3}$$

Calculons la matrice associée à $g \circ f : E \to G$, $C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f)$, de deux façons différentes.

1. **Première méthode.** Revenons à la définition de la matrice de l'application linéaire $g \circ f$. Il s'agit d'exprimer l'image des vecteurs de la base de départ $\mathscr{B}$ dans la base d'arrivée $\mathscr{B}''$. C'est-à-dire qu'il faut exprimer $g \circ f(e_j)$ dans la base (g_1, g_2).
   - Calcul des $f(e_j)$. On sait par définition de la matrice A que $f(e_1)$ correspond au premier vecteur colonne : plus précisément, $f(e_1) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}_{\mathscr{B}'} = 1f_1 + 1f_2 + 0f_3 = f_1 + f_2$. De même, $f(e_2) = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}_{\mathscr{B}'} = 0f_1 + 1f_2 + 2f_3 = f_2 + 2f_3$.
   - Calcul des $g(f_j)$. Par définition, $g(f_j)$ correspond à la j-ème colonne de la matrice B :
   $$g(f_1) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}_{\mathscr{B}''} = 2g_1 + 3g_2 \qquad g(f_2) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}_{\mathscr{B}''} = -g_1 + g_2 \qquad g(f_3) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}_{\mathscr{B}''} = 2g_2$$
   - Calcul des $g \circ f(e_j)$. Pour cela on combine les deux séries de calculs précédents :
   $$g \circ f(e_1) = g(f_1 + f_2) = g(f_1) + g(f_2) = (2g_1 + 3g_2) + (-g_1 + g_2) = g_1 + 4g_2$$
   $$g \circ f(e_2) = g(f_2 + 2f_3) = g(f_2) + 2g(f_3) = (-g_1 + g_2) + 2(2g_2) = -g_1 + 5g_2$$
   - Calcul de la matrice $C = \mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f)$: cette matrice est composée des vecteurs $g \circ f(e_j)$ exprimés dans la base $\mathscr{B}''$. Comme
   $$g \circ f(e_1) = g_1 + 4g_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}_{\mathscr{B}''} \qquad g \circ f(e_2) = -g_1 + 5g_2 = \begin{pmatrix} -1 \\ 5 \end{pmatrix}_{\mathscr{B}''} \qquad \text{alors} \qquad C = \begin{pmatrix} 1 & -1 \\ 4 & 5 \end{pmatrix}$$

   On trouve bien une matrice de taille 2×2 (car l'espace de départ et d'arrivée de $g \circ f$ est $\mathbb{R}^2$).
2. **Deuxième méthode.** Utilisons le produit de matrices : on sait que $C = BA$. Donc
$$\mathrm{Mat}_{\mathscr{B},\mathscr{B}''}(g \circ f) = C = B \times A = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 1 & 2 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 4 & 5 \end{pmatrix}$$

Cet exemple met bien en évidence le gain, en termes de quantité de calculs, réalisé en passant par l'intermédiaire des matrices.

3.3. Matrice d'un endomorphisme

Dans cette section, on étudie la cas où l'espace de départ et l'espace d'arrivée sont identiques : $f : E \to E$ est un endomorphisme. Si $\dim E = n$, alors chaque matrice associée à f est une matrice carrée de taille $n \times n$.

Deux situations :

- Si on choisit la même base $\mathscr{B}$ au départ et à l'arrivée, alors on note simplement $\mathrm{Mat}_{\mathscr{B}}(f)$ la matrice associée à f.
- Mais on peut aussi choisir deux bases distinctes pour le même espace vectoriel E ; on note alors comme précédemment $\mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$.

Exemple 108

- Cas de l'identité : $\mathrm{id} : E \to E$ est définie par $\mathrm{id}(x) = x$. Alors quelle que soit la base $\mathscr{B}$ de E, la matrice associée est la matrice identité : $\mathrm{Mat}_{\mathscr{B}}(f) = I_n$. (Attention ! Ce n'est plus vrai si la base d'arrivée est différente de la base de départ.)
- Cas d'une homothétie $h_\lambda : E \to E$, $h_\lambda(x) = \lambda \cdot x$ (où $\lambda \in \mathbb{K}$ est le rapport de l'homothétie) : $\mathrm{Mat}_{\mathscr{B}}(h_\lambda) = \lambda I_n$.
- Cas d'une symétrie centrale $s : E \to E$, $s(x) = -x$: $\mathrm{Mat}_{\mathscr{B}}(s) = -I_n$.
- Cas de $r_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la rotation d'angle θ, centrée à l'origine, dans l'espace vectoriel $\mathbb{R}^2$ muni de la base canonique $\mathscr{B}$. Alors $r_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$. On a

$$\mathrm{Mat}_{\mathscr{B}}(r_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Dans le cas particulier de la puissance d'un endomorphisme de E, nous obtenons :

Corollaire 8

Soient E un espace vectoriel de dimension finie et $\mathscr{B}$ une base de E. Soit $f : E \to E$ une application linéaire. Alors, quel que soit $p \in \mathbb{N}$:

$$\mathrm{Mat}_{\mathscr{B}}(f^p) = \big(\mathrm{Mat}_{\mathscr{B}}(f)\big)^p$$

Autrement dit, si A est la matrice associée à f, alors la matrice associée à $f^p = \underbrace{f \circ f \circ \cdots \circ f}_{p \text{ occurrences}}$ est $A^p = \underbrace{A \times A \times \cdots \times A}_{p \text{ facteurs}}$. La démonstration est une récurrence sur p en utilisant la proposition 36.

Exemple 109

Soit r_θ la matrice de la rotation d'angle θ dans $\mathbb{R}^2$. La matrice de r_θ^p est :

$$\mathrm{Mat}_{\mathscr{B}}(r_\theta^p) = \big(\mathrm{Mat}_{\mathscr{B}}(r_\theta)\big)^p = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^p$$

Un calcul par récurrence montre ensuite que

$$\mathrm{Mat}_{\mathscr{B}}(r_\theta^p) = \begin{pmatrix} \cos(p\theta) & -\sin(p\theta) \\ \sin(p\theta) & \cos(p\theta) \end{pmatrix},$$

ce qui est bien la matrice de la rotation d'angle $p\theta$: composer p fois la rotation d'angle θ revient à effectuer une rotation d'angle $p\theta$.

3.4. Matrice d'un isomorphisme

Passons maintenant aux isomorphismes. Rappelons qu'un isomorphisme $f : E \to F$ est une application linéaire bijective. Nous avons vu que cela entraîne $\dim E = \dim F$.

Théorème 31. Caractérisation de la matrice d'un isomorphisme

Soient E et F deux $\mathbb{K}$-espaces vectoriels de même dimension finie. Soit $f : E \to F$ une application linéaire. Soient $\mathscr{B}$ une base de E, $\mathscr{B}'$ une base de F et $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$.

1. f est bijective si et seulement si la matrice A est inversible. Autrement dit, f est un isomorphisme si et seulement si sa matrice associée $\mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$ est inversible.
2. De plus, si $f : E \to F$ est bijective, alors la matrice de l'application linéaire $f^{-1} : F \to E$ est la matrice A^{-1}. Autrement dit, $\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f^{-1}) = \left(\mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)\right)^{-1}$.

Voici le cas particulier très important d'un endomorphisme $f : E \to E$ où E est muni de la même base $\mathscr{B}$ au départ et à l'arrivée et $A = \mathrm{Mat}_{\mathscr{B}}(f)$.

Corollaire 9

- f est bijective si et seulement si A est inversible.
- Si f est bijective, alors la matrice associée à f^{-1} dans la base $\mathscr{B}$ est A^{-1}.

Exemple 110

Soient $r : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\frac{\pi}{6}$ (centrée à l'origine) et s la réflexion par rapport à l'axe $(y = x)$. Quelle est la matrice associée à $(s \circ r)^{-1}$ dans la base canonique $\mathscr{B}$?

- Pour $\theta = \frac{\pi}{6}$, on trouve la matrice $A = \mathrm{Mat}_{\mathscr{B}}(r) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$.
- La matrice associée à la réflexion est $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- La matrice de $s \circ r$ est $B \times A = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$.
- La matrice de $(s \circ r)^{-1}$ est $(BA)^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. On aurait aussi pu calculer ainsi : $(BA)^{-1} = A^{-1}B^{-1} = \cdots$
- On note que $(BA)^{-1} = BA$ ce qui, en termes d'applications linéaires, signifie que $(s \circ r)^{-1} = s \circ r$. Autrement dit, $s \circ r$ est son propre inverse.

Démonstration . Preuve du théorème 31

On note $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$.

- Si f est bijective, notons $B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f^{-1})$. Alors par la proposition 36 on sait que

$$BA = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f^{-1}) \times \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f^{-1} \circ f) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}}(\mathrm{id}_E) = I.$$

  De même $AB = I$. Ainsi $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$ est inversible et son inverse est $B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f^{-1})$.
- Réciproquement, si $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$ est une matrice inversible, notons $B = A^{-1}$. Soit $g : F \to E$ l'application linéaire telle que $B = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(g)$. Alors, toujours par la proposition 36 :

$$\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(g \circ f) = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(g) \times \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) = BA = I$$

  Donc la matrice de $g \circ f$ est l'identité, ce qui implique $g \circ f = \mathrm{id}_E$. De même $f \circ g = \mathrm{id}_F$. Ainsi f est bijective (et sa bijection réciproque est g).

Mini-exercices

1. Calculer la matrice associée aux applications linéaires $f_i : \mathbb{R}^2 \to \mathbb{R}^2$:
   (a) f_1 la symétrie par rapport à l'axe (Oy),
   (b) f_2 la symétrie par rapport à l'axe $(y = x)$,
   (c) f_3 le projection orthogonale sur l'axe (Oy),
   (d) f_4 la rotation d'angle $\frac{\pi}{4}$.
   Calculer quelques matrices associées à $f_i \circ f_j$ et, lorsque c'est possible, à f_i^{-1}.
2. Même travail pour $f_i : \mathbb{R}^3 \to \mathbb{R}^3$:
   (a) f_1 l'homothétie de rapport λ,
   (b) f_2 la réflexion orthogonale par rapport au plan (Oxz),
   (c) f_3 la rotation d'axe (Oz) d'angle $-\frac{\pi}{2}$,
   (d) f_4 la projection orthogonale sur le plan (Oyz).

4. Changement de bases

4.1. Application linéaire, matrice, vecteur

Soit E un espace vectoriel de dimension finie et soit $\mathscr{B} = (e_1, e_2, \dots, e_p)$ une base de E. Pour chaque $x \in E$, il existe un p-uplet unique d'éléments de $\mathbb{K}$ $(x_1, x_2, \dots, x_p)$ tel que

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_p e_p.$$

La matrice des coordonnées de x est un vecteur colonne, noté $\mathrm{Mat}_{\mathscr{B}}(v)$ ou encore $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}_{\mathscr{B}}$.

Dans $\mathbb{R}^p$, si $\mathscr{B}$ est la base canonique, alors on note simplement $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ en omettant de mentionner la base.

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f : E \to F$ une application linéaire. Le but de ce paragraphe est de traduire l'égalité vectorielle $y = f(x)$ par une égalité matricielle.
Soient $\mathscr{B}$ une base de E et $\mathscr{B}'$ une base de F.

Proposition 37

- Soit $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$.
- Pour $x \in E$, notons $X = \mathrm{Mat}_{\mathscr{B}}(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}_{\mathscr{B}}$.
- Pour $y \in F$, notons $Y = \mathrm{Mat}_{\mathscr{B}'}(y) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{\mathscr{B}'}$.

Alors, si $y = f(x)$, on a

$$\boxed{Y = AX}$$

Autrement dit :

$$\boxed{\mathrm{Mat}_{\mathscr{B}'}\left(f(x)\right) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f) \times \mathrm{Mat}_{\mathscr{B}}(x)}$$

Démonstration

- On pose $\mathscr{B} = (e_1, \dots, e_p)$, $\mathscr{B}' = (f_1, f_2, \dots, f_n)$, $A = (a_{i,j}) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(f)$ et $X = \mathrm{Mat}_{\mathscr{B}}(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$.
- On a
$$f(x) = f\left(\sum_{j=1}^{p} x_j e_j\right) = \sum_{j=1}^{p} x_j f(e_j) = \sum_{j=1}^{p} x_j \left(\sum_{i=1}^{n} a_{i,j} f_i\right).$$
En utilisant la commutativité de $\mathbb{K}$, on a
$$f(x) = \left(\sum_{j=1}^{p} a_{1,j} x_j\right) f_1 + \cdots + \left(\sum_{j=1}^{p} a_{n,j} x_j\right) f_n.$$
- La matrice colonne des coordonnées de $y = f(x)$ dans la base $(f_1, f_2, \dots, f_n)$ est $\begin{pmatrix} \sum_{j=1}^{p} a_{1,j} x_j \\ \sum_{j=1}^{p} a_{2,j} x_j \\ \vdots \\ \sum_{j=1}^{p} a_{n,j} x_j \end{pmatrix}$.
- Ainsi la matrice $Y = \mathrm{Mat}_{\mathscr{B}'}\left(f(x)\right) = \begin{pmatrix} \sum_{j=1}^{p} a_{1,j} x_j \\ \sum_{j=1}^{p} a_{2,j} x_j \\ \vdots \\ \sum_{j=1}^{p} a_{n,j} x_j \end{pmatrix}$ n'est autre que $A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$.

Exemple 111

Soient E un $\mathbb{K}$-espace vectoriel de dimension 3 et $\mathscr{B} = (e_1, e_2, e_3)$ une base de E. Soit f l'endomorphisme de E dont la matrice dans la base $\mathscr{B}$ est égale à

$$A = \mathrm{Mat}_{\mathscr{B}}(f) = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

On se propose de déterminer le noyau de f et l'image de f.
Les éléments x de E sont des combinaisons linéaires de e_1, e_2 et e_3 : $x = x_1 e_1 + x_2 e_2 + x_3 e_3$.
On a

$$\begin{aligned} x \in \operatorname{Ker} f \quad & \iff \quad f(x) = 0_E \iff \mathrm{Mat}_{\mathscr{B}}\left(f(x)\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ & \iff \quad AX = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ & \iff \quad \begin{cases} x_1 + 2x_2 + x_3 = 0 \\ 2x_1 + 3x_2 + x_3 = 0 \\ x_1 + x_2 = 0 \end{cases} \end{aligned}$$

On résout ce système par la méthode du pivot de Gauss. On trouve

$$\operatorname{Ker} f = \left\{x_1 e_1 + x_2 e_2 + x_3 e_3 \in E \mid x_1 + 2x_2 + x_3 = 0 \text{ et } x_2 + x_3 = 0\right\} = \left\{(t, -t, t) \mid t \in \mathbb{K}\right\} = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}_{\mathscr{B}}\right)$$

Le noyau est donc de dimension 1. Par le théorème du rang, l'image $\operatorname{Im} f$ est de dimension 2. Les deux premiers vecteurs de la matrice A étant linéairement indépendants, ils engendrent

Im f : $\operatorname{Im} f = \operatorname{Vect}\left(\begin{pmatrix}1\\2\\1\end{pmatrix}_{\mathscr{B}}, \begin{pmatrix}2\\3\\1\end{pmatrix}_{\mathscr{B}}\right)$.

4.2. Matrice de passage d'une base à une autre

Soit E un espace vectoriel de dimension finie n. On sait que toutes les bases de E ont n éléments.

Définition 39

Soit $\mathscr{B}$ une base de E. Soit $\mathscr{B}'$ une autre base de E.
On appelle ***matrice de passage*** de la base $\mathscr{B}$ vers la base $\mathscr{B}'$, et on note $P_{\mathscr{B},\mathscr{B}'}$, la matrice carrée de taille $n \times n$ dont la j-ème colonne est formée des coordonnées du j-ème vecteur de la base $\mathscr{B}'$, par rapport à la base $\mathscr{B}$.

On résume en :

La matrice de passage $P_{\mathscr{B},\mathscr{B}'}$ contient - en colonnes - les coordonnées des vecteurs de la nouvelle base $\mathscr{B}'$ exprimés dans l'ancienne base $\mathscr{B}$.

C'est pourquoi on note parfois aussi $P_{\mathscr{B},\mathscr{B}'}$ par $\operatorname{Mat}_{\mathscr{B}}(\mathscr{B}')$.

Exemple 112

Soit l'espace vectoriel réel $\mathbb{R}^2$. On considère

$$e_1 = \begin{pmatrix}1\\0\end{pmatrix} \quad e_2 = \begin{pmatrix}1\\1\end{pmatrix} \qquad \varepsilon_1 = \begin{pmatrix}1\\2\end{pmatrix} \quad \varepsilon_2 = \begin{pmatrix}5\\4\end{pmatrix}.$$

On considère la base $\mathscr{B} = (e_1, e_2)$ et la base $\mathscr{B}' = (\varepsilon_1, \varepsilon_2)$.
Quelle est la matrice de passage de la base $\mathscr{B}$ vers la base $\mathscr{B}'$?
Il faut exprimer ε_1 et ε_2 en fonction de (e_1, e_2). On calcule que :

$$\varepsilon_1 = -e_1 + 2e_2 = \begin{pmatrix}-1\\2\end{pmatrix}_{\mathscr{B}} \qquad \varepsilon_2 = e_1 + 4e_2 = \begin{pmatrix}1\\4\end{pmatrix}_{\mathscr{B}}$$

La matrice de passage est donc :

$$P_{\mathscr{B},\mathscr{B}'} = \begin{pmatrix}-1 & 1\\2 & 4\end{pmatrix}$$

On va interpréter une matrice de passage comme la matrice associée à l'application identique de E par rapport à des bases bien choisies.

Proposition 38

La matrice de passage $P_{\mathscr{B},\mathscr{B}'}$ de la base $\mathscr{B}$ vers la base $\mathscr{B}'$ est la matrice associée à l'identité $\mathrm{id}_E : (E, \mathscr{B}') \to (E, \mathscr{B})$ où E est l'espace de départ muni de la base $\mathscr{B}'$, et E est aussi l'espace d'arrivée, mais muni de la base $\mathscr{B}$:

$$\boxed{P_{\mathscr{B},\mathscr{B}'} = \operatorname{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E)}$$

Faites bien attention à l'inversion de l'ordre des bases !

Cette interprétation est un outil fondamental pour ce qui suit. Elle permet d'obtenir les résultats de façon très élégante et avec un minimum de calculs.

Démonstration

On pose $\mathscr{B} = (e_1, e_2, \dots, e_n)$ et $\mathscr{B}' = (e'_1, e'_2, \dots, e'_n)$. On considère

$$\begin{array}{rccc} \mathrm{id}_E & : & (E, \mathscr{B}') & \longrightarrow & (E, \mathscr{B}) \\ & & x & \longmapsto & \mathrm{id}_E(x) = x \end{array}$$

On a $\mathrm{id}_E(e'_j) = e'_j = \sum_{i=1}^n a_{i,j} e_i$ et $\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E)$ est la matrice dont la j-ème colonne est formée des coordonnées de e'_j par rapport à $\mathscr{B}$, soit $\begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$. Cette colonne est la j-ème colonne de $\mathrm{P}_{\mathscr{B},\mathscr{B}'}$.

Proposition 39

1. La matrice de passage d'une base $\mathscr{B}$ vers une base $\mathscr{B}'$ est inversible et son inverse est égale à la matrice de passage de la base $\mathscr{B}'$ vers la base $\mathscr{B}$: $\boxed{\mathrm{P}_{\mathscr{B}',\mathscr{B}} = \left(\mathrm{P}_{\mathscr{B},\mathscr{B}'}\right)^{-1}}$
2. Si $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ sont trois bases, alors $\boxed{\mathrm{P}_{\mathscr{B},\mathscr{B}''} = \mathrm{P}_{\mathscr{B},\mathscr{B}'} \times \mathrm{P}_{\mathscr{B}',\mathscr{B}''}}$

Démonstration

1. On a $\mathrm{P}_{\mathscr{B},\mathscr{B}'} = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}\left(\mathrm{id}_E\right)$. Donc, d'après le théorème 31 caractérisant la matrice d'un isomorphisme, $\mathrm{P}^{-1}_{\mathscr{B},\mathscr{B}'} = \left(\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}\left(\mathrm{id}_E\right)\right)^{-1} = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}\left(\mathrm{id}_E^{-1}\right)$. Or $\mathrm{id}_E^{-1} = \mathrm{id}_E$, donc $\mathrm{P}^{-1}_{\mathscr{B},\mathscr{B}'} = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}\left(\mathrm{id}_E\right) = \mathrm{P}_{\mathscr{B}',\mathscr{B}}$.
2. $\mathrm{id}_E : (E, \mathscr{B}'') \to (E, \mathscr{B})$ se factorise de la façon suivante :
$$(E, \mathscr{B}'') \xrightarrow{\mathrm{id}_E} (E, \mathscr{B}') \xrightarrow{\mathrm{id}_E} (E, \mathscr{B}).$$
Autrement dit, on écrit $\mathrm{id}_E = \mathrm{id}_E \circ \mathrm{id}_E$. Cette factorisation permet d'écrire l'égalité suivante : $\mathrm{Mat}_{\mathscr{B}'',\mathscr{B}}\left(\mathrm{id}_E\right) = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}\left(\mathrm{id}_E\right) \times \mathrm{Mat}_{\mathscr{B}'',\mathscr{B}'}\left(\mathrm{id}_E\right)$, soit $\mathrm{P}_{\mathscr{B},\mathscr{B}''} = \mathrm{P}_{\mathscr{B},\mathscr{B}'} \times \mathrm{P}_{\mathscr{B}',\mathscr{B}''}$.

Exemple 113

Soit $E = \mathbb{R}^3$ muni de sa base canonique $\mathscr{B}$. Définissons

$$\mathscr{B}_1 = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \right) \quad \text{et} \quad \mathscr{B}_2 = \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right).$$

Quelle est la matrice de passage de $\mathscr{B}_1$ vers $\mathscr{B}_2$?
On a d'abord

$$\mathrm{P}_{\mathscr{B},\mathscr{B}_1} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{et} \quad \mathrm{P}_{\mathscr{B},\mathscr{B}_2} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

La proposition 39 implique que $\mathrm{P}_{\mathscr{B},\mathscr{B}_2} = \mathrm{P}_{\mathscr{B},\mathscr{B}_1} \times \mathrm{P}_{\mathscr{B}_1,\mathscr{B}_2}$. Donc $\mathrm{P}_{\mathscr{B}_1,\mathscr{B}_2} = \mathrm{P}^{-1}_{\mathscr{B},\mathscr{B}_1} \times \mathrm{P}_{\mathscr{B},\mathscr{B}_2}$. En appliquant la méthode de Gauss pour calculer $\mathrm{P}^{-1}_{\mathscr{B},\mathscr{B}_1}$, on trouve alors :

$$\mathrm{P}_{\mathscr{B}_1,\mathscr{B}_2} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix}^{-1} \times \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -3 \\ 2 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nous allons maintenant étudier l'effet d'un changement de bases sur les coordonnées d'un vecteur.

- Soient $\mathscr{B} = (e_1, e_2, \dots, e_n)$ et $\mathscr{B}' = (e'_1, e'_2, \dots, e'_n)$ deux bases d'un même $\mathbb{K}$-espace vectoriel E.
- Soit $P_{\mathscr{B},\mathscr{B}'}$ la matrice de passage de la base $\mathscr{B}$ vers la base $\mathscr{B}'$.
- Pour $x \in E$, il se décompose en $x = \sum_{i=1}^n x_i e_i$ dans la base $\mathscr{B}$ et on note $X = \mathrm{Mat}_{\mathscr{B}}(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathscr{B}}$.
- Ce même $x \in E$ se décompose en $x = \sum_{i=1}^n x'_i e'_i$ dans la base $\mathscr{B}'$ et on note $X' = \mathrm{Mat}_{\mathscr{B}'}(x) = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}_{\mathscr{B}'}$.

Proposition 40

$$\boxed{X = P_{\mathscr{B},\mathscr{B}'} \times X'}$$

Notez bien l'ordre !

Démonstration

$P_{\mathscr{B},\mathscr{B}'}$ est la matrice de $\mathrm{id}_E : (E, \mathscr{B}') \to (E, \mathscr{B})$. On utilise que $x = \mathrm{id}_E(x)$ et la proposition 37. On a :

$$X = \mathrm{Mat}_{\mathscr{B}}(x) = \mathrm{Mat}_{\mathscr{B}}\big(\mathrm{id}_E(x)\big) = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E) \times \mathrm{Mat}_{\mathscr{B}'}(x) = P_{\mathscr{B},\mathscr{B}'} \times X'$$

4.3. Formule de changement de base

- Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie.
- Soit $f : E \to F$ une application linéaire.
- Soient $\mathscr{B}_E$, $\mathscr{B}'_E$ deux bases de E.
- Soient $\mathscr{B}_F$, $\mathscr{B}'_F$ deux bases de F.
- Soit $P = P_{\mathscr{B}_E,\mathscr{B}'_E}$ la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$.
- Soit $Q = P_{\mathscr{B}_F,\mathscr{B}'_F}$ la matrice de passage de $\mathscr{B}_F$ à $\mathscr{B}'_F$.
- Soit $A = \mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}(f)$ la matrice de l'application linéaire f de la base $\mathscr{B}_E$ vers la base $\mathscr{B}_F$.
- Soit $B = \mathrm{Mat}_{\mathscr{B}'_E,\mathscr{B}'_F}(f)$ la matrice de l'application linéaire f de la base $\mathscr{B}'_E$ vers la base $\mathscr{B}'_F$.

Théorème 32. Formule de changement de base

$$\boxed{B = Q^{-1} A P}$$

Démonstration

L'application $f : (E, \mathscr{B}'_E) \to (F, \mathscr{B}'_F)$ se factorise de la façon suivante :

$$(E, \mathscr{B}'_E) \xrightarrow{\mathrm{id}_E} (E, \mathscr{B}_E) \xrightarrow{f} (F, \mathscr{B}_F) \xrightarrow{\mathrm{id}_F} (F, \mathscr{B}'_F),$$

c'est-à-dire que $f = \mathrm{id}_F \circ f \circ \mathrm{id}_E$.

On a donc l'égalité de matrices suivante :

$$\begin{aligned} B &= \mathrm{Mat}_{\mathscr{B}'_E,\mathscr{B}'_F}(f) \\ &= \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}'_F}(\mathrm{id}_F)\times \mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}(f)\times \mathrm{Mat}_{\mathscr{B}'_E,\mathscr{B}_E}(\mathrm{id}_E) \\ &= \mathrm{P}_{\mathscr{B}'_F,\mathscr{B}_F}\times \mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}(f)\times \mathrm{P}_{\mathscr{B}_E,\mathscr{B}'_E} \\ &= Q^{-1}AP \end{aligned}$$

Dans le cas particulier d'un endomorphisme, nous obtenons une formule plus simple :

- Soit $f : E \to E$ une application linéaire.
- Soient $\mathscr{B}$, $\mathscr{B}'$ deux bases de E.
- Soit $P = \mathrm{P}_{\mathscr{B},\mathscr{B}'}$ la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.
- Soit $A = \mathrm{Mat}_{\mathscr{B}}(f)$ la matrice de l'application linéaire f dans la base $\mathscr{B}$.
- Soit $B = \mathrm{Mat}_{\mathscr{B}'}(f)$ la matrice de l'application linéaire f dans la base $\mathscr{B}'$.

Le théorème 32 devient alors :

Corollaire 10

$$\boxed{B = P^{-1}AP}$$

Exemple 114

Reprenons les deux bases de $\mathbb{R}^3$ de l'exemple 113 :

$$\mathscr{B}_1 = \left(\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}0\\-1\\0\end{pmatrix},\begin{pmatrix}3\\2\\-1\end{pmatrix}\right) \qquad \text{et} \qquad \mathscr{B}_2 = \left(\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\-1\end{pmatrix}\right).$$

Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire dont la matrice dans la base $\mathscr{B}_1$ est :

$$A = \mathrm{Mat}_{\mathscr{B}_1}(f) = \begin{pmatrix}1 & 0 & -6\\ -2 & 2 & -7\\ 0 & 0 & 3\end{pmatrix}$$

Que vaut la matrice de f dans la base $\mathscr{B}_2$, $B = \mathrm{Mat}_{\mathscr{B}_2}(f)$?

1. Nous avions calculé que la matrice de passage de $\mathscr{B}_1$ vers $\mathscr{B}_2$ était
$$P = \mathrm{P}_{\mathscr{B}_1,\mathscr{B}_2} = \begin{pmatrix}1 & 0 & -3\\ 2 & -1 & -1\\ 0 & 0 & 1\end{pmatrix}.$$
2. On calcule aussi $P^{-1} = \begin{pmatrix}1 & 0 & 3\\ 2 & -1 & 5\\ 0 & 0 & 1\end{pmatrix}$.
3. On applique la formule du changement de base du corollaire 10 :

$$B = P^{-1}AP = \begin{pmatrix}1 & 0 & 3\\ 2 & -1 & 5\\ 0 & 0 & 1\end{pmatrix}\times\begin{pmatrix}1 & 0 & -6\\ -2 & 2 & -7\\ 0 & 0 & 3\end{pmatrix}\times\begin{pmatrix}1 & 0 & -3\\ 2 & -1 & -1\\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{pmatrix}$$

C'est souvent l'intérêt des changements de base, se ramener à une matrice plus simple. Par exemple ici, il est facile de calculer les puissances B^k, pour en déduire les A^k.

4.4. Matrices semblables

Les matrices considérées dans ce paragraphe sont des matrices carrées, éléments de $\mathcal{M}_n(\mathbb{K})$.

Définition 40

Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$. On dit que la matrice B est ***semblable*** à la matrice A s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

C'est un bon exercice de montrer que la relation « être semblable » est une relation d'équivalence dans l'ensemble $\mathcal{M}_n(\mathbb{K})$:

Proposition 41

- La relation est ***réflexive*** : une matrice A est semblable à elle-même.
- La relation est ***symétrique*** : si A est semblable à B, alors B est semblable à A.
- La relation est ***transitive*** : si A est semblable à B, et B est semblable à C, alors A est semblable à C.

Vocabulaire :
Compte tenu de ces propriétés, on peut dire indifféremment que la matrice A est semblable à la matrice B ou que les matrices A et B sont semblables.
Le corollaire 10 se reformule ainsi :

Corollaire 11

Deux matrices semblables représentent le même endomorphisme, mais exprimé dans des bases différentes.

Mini-exercices

Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (2x+y, 3x-2y)$, Soit $v = \begin{pmatrix} 3 \\ -4 \end{pmatrix} \in \mathbb{R}^2$ avec ses coordonnées dans la base canonique $\mathcal{B}_0$ de $\mathbb{R}^2$. Soit $\mathcal{B}_1 = \left(\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right)$ une autre base de $\mathbb{R}^2$.

1. Calculer la matrice de f dans la base canonique.
2. Calculer les coordonnées de $f(v)$ dans la base canonique.
3. Calculer la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}_1$.
4. En déduire les coordonnées de v dans la base $\mathcal{B}_1$, et de $f(v)$ dans la base $\mathcal{B}_1$.
5. Calculer la matrice de f dans la base $\mathcal{B}_1$.

Même exercice dans $\mathbb{R}^3$ avec $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x,y,z) = (x-2y, y-2z, z-2x)$, $v = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ et $\mathcal{B}_1 = \left(\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}\right)$.

Auteurs

- D'après un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties d'un cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- réécrit et complété par Arnaud Bodin. Relu par Vianney Combet.

7 Vector products

This chapter is extracted from the book by Kenneth Kuttler *Elementary linear algebra*. This book can be downloaded from author's web page : link to the full text of the book.

Vector Products

3.1 The Dot Product

There are two ways of multiplying vectors which are of great importance in applications. The first of these is called the **dot product**, also called the **scalar product** and sometimes the **inner product**.

Definition 3.1.1 *Let* $\mathbf{a}, \mathbf{b}$ *be two vectors in* $\mathbb{R}^n$ *define* $\mathbf{a} \cdot \mathbf{b}$ *as*

$$\mathbf{a} \cdot \mathbf{b} \equiv \sum_{k=1}^{n} a_k b_k.$$

The dot product $\mathbf{a} \cdot \mathbf{b}$ *is sometimes denoted as* $(\mathbf{a}, \mathbf{b})$ *of* $\langle \mathbf{a}, \mathbf{b} \rangle$ *where a comma replaces* $\cdot$.

With this definition, there are several important properties satisfied by the dot product. In the statement of these properties, α and β will denote scalars and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ will denote vectors.

Proposition 3.1.2 *The dot product satisfies the following properties.*

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \tag{3.1}$$

$$\mathbf{a} \cdot \mathbf{a} \geq 0 \text{ and equals zero if and only if } \mathbf{a} = \mathbf{0} \tag{3.2}$$

$$(\alpha \mathbf{a} + \beta \mathbf{b}) \cdot \mathbf{c} = \alpha (\mathbf{a} \cdot \mathbf{c}) + \beta (\mathbf{b} \cdot \mathbf{c}) \tag{3.3}$$

$$\mathbf{c} \cdot (\alpha \mathbf{a} + \beta \mathbf{b}) = \alpha (\mathbf{c} \cdot \mathbf{a}) + \beta (\mathbf{c} \cdot \mathbf{b}) \tag{3.4}$$

$$|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} \tag{3.5}$$

You should verify these properties. Also be sure you understand that 3.4 follows from the first three and is therefore redundant. It is listed here for the sake of convenience.

Example 3.1.3 *Find* $(1, 2, 0, -1) \cdot (0, 1, 2, 3)$.

This equals $0 + 2 + 0 + -3 = -1$.

Example 3.1.4 *Find the magnitude of* $\mathbf{a} = (2, 1, 4, 2)$. *That is, find* $|\mathbf{a}|$.

This is $\sqrt{(2, 1, 4, 2) \cdot (2, 1, 4, 2)} = 5$.
The dot product satisfies a fundamental inequality known as the **Cauchy Schwarz inequality.**

Theorem 3.1.5 *The dot product satisfies the inequality*

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}| |\mathbf{b}|. \tag{3.6}$$

Furthermore equality is obtained if and only if one of $\mathbf{a}$ *or* $\mathbf{b}$ *is a scalar multiple of the other.*

Proof: First note that if $\mathbf{b} = \mathbf{0}$ both sides of 3.6 equal zero and so the inequality holds in this case. Therefore, it will be assumed in what follows that $\mathbf{b} \neq \mathbf{0}$.

Define a function of $t \in \mathbb{R}$

$$f(t) = (\mathbf{a} + t\mathbf{b}) \cdot (\mathbf{a} + t\mathbf{b}).$$

Then by 3.2, $f(t) \geq 0$ for all $t \in \mathbb{R}$. Also from 3.3,3.4,3.1, and 3.5

$$\begin{aligned} f(t) &= \mathbf{a} \cdot (\mathbf{a} + t\mathbf{b}) + t\mathbf{b} \cdot (\mathbf{a} + t\mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + t(\mathbf{a} \cdot \mathbf{b}) + t\mathbf{b} \cdot \mathbf{a} + t^2 \mathbf{b} \cdot \mathbf{b} \\ &= |\mathbf{a}|^2 + 2t(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2 t^2. \end{aligned}$$

Now this means the graph, $y = f(t)$ is a polynomial which opens up and either its vertex touches the t axis or else the entire graph is above the x axis. In the first case, there exists some t where $f(t) = 0$ and this requires $\mathbf{a} + t\mathbf{b} = \mathbf{0}$ so one vector is a multiple of the other. Then clearly equality holds in 3.6. In the case where $\mathbf{b}$ is not a multiple of $\mathbf{a}$, it follows $f(t) > 0$ for all t which says $f(t)$ has no real zeros and so from the quadratic formula,

$$(2(\mathbf{a} \cdot \mathbf{b}))^2 - 4|\mathbf{a}|^2 |\mathbf{b}|^2 < 0$$

which is equivalent to $|(\mathbf{a} \cdot \mathbf{b})| < |\mathbf{a}| |\mathbf{b}|$. ■

You should note that the entire argument was based only on the properties of the dot product listed in 3.1 - 3.5. This means that whenever something satisfies these properties, the Cauchy Schwarz inequality holds. There are many other instances of these properties besides vectors in $\mathbb{R}^n$.

The Cauchy Schwarz inequality allows a proof of the **triangle inequality** for distances in $\mathbb{R}^n$ in much the same way as the triangle inequality for the absolute value.

Theorem 3.1.6 *(Triangle inequality) For* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}| \tag{3.7}$$

and equality holds if and only if one of the vectors is a nonnegative scalar multiple of the other. Also

$$||\mathbf{a}| - |\mathbf{b}|| \leq |\mathbf{a} - \mathbf{b}| \tag{3.8}$$

Proof: By properties of the dot product and the Cauchy Schwarz inequality,

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= (\mathbf{a} \cdot \mathbf{a}) + (\mathbf{a} \cdot \mathbf{b}) + (\mathbf{b} \cdot \mathbf{a}) + (\mathbf{b} \cdot \mathbf{b}) \\ &= |\mathbf{a}|^2 + 2(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2 \\ &\leq |\mathbf{a}|^2 + 2|\mathbf{a} \cdot \mathbf{b}| + |\mathbf{b}|^2 \\ &\leq |\mathbf{a}|^2 + 2|\mathbf{a}| |\mathbf{b}| + |\mathbf{b}|^2 \\ &= (|\mathbf{a}| + |\mathbf{b}|)^2. \end{aligned}$$

Taking square roots of both sides you obtain 3.7.

It remains to consider when equality occurs. If either vector equals zero, then that vector equals zero times the other vector and the claim about when equality occurs is verified. Therefore, it can be assumed both vectors are nonzero. To get equality in the second inequality above, Theorem 3.1.5 implies one of the vectors must be a multiple of the other. Say $\mathbf{b} = \alpha\mathbf{a}$. If $\alpha < 0$ then equality cannot occur in the first inequality because in this case

$$(\mathbf{a} \cdot \mathbf{b}) = \alpha |\mathbf{a}|^2 < 0 < |\alpha| |\mathbf{a}|^2 = |\mathbf{a} \cdot \mathbf{b}|$$

Therefore, $\alpha \geq 0$.

To get the other form of the triangle inequality,

$$\mathbf{a} = \mathbf{a} - \mathbf{b} + \mathbf{b}$$

so

$$\begin{aligned}|\mathbf{a}| &= |\mathbf{a}-\mathbf{b}+\mathbf{b}| \\ &\leq |\mathbf{a}-\mathbf{b}| + |\mathbf{b}| .\end{aligned}$$

Therefore,

$$|\mathbf{a}| - |\mathbf{b}| \leq |\mathbf{a}-\mathbf{b}| \tag{3.9}$$

Similarly,

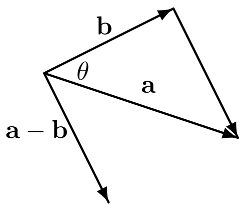
$$|\mathbf{b}| - |\mathbf{a}| \leq |\mathbf{b}-\mathbf{a}| = |\mathbf{a}-\mathbf{b}| . \tag{3.10}$$

It follows from 3.9 and 3.10 that 3.8 holds. This is because $||\mathbf{a}| - |\mathbf{b}||$ equals the left side of either 3.9 or 3.10 and either way, $||\mathbf{a}| - |\mathbf{b}|| \leq |\mathbf{a}-\mathbf{b}|$. ■

3.2 The Geometric Significance Of The Dot Product

3.2.1 The Angle Between Two Vectors

Given two vectors, **a** and **b**, the included angle is the angle between these two vectors which is less than or equal to 180 degrees. The dot product can be used to determine the included angle between two vectors. To see how to do this, consider the following picture.



By the law of cosines,

$$|\mathbf{a}-\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\,|\mathbf{a}|\,|\mathbf{b}|\cos\theta.$$

Also from the properties of the dot product,

$$\begin{aligned}|\mathbf{a}-\mathbf{b}|^2 &= (\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b}) \\ &= |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\mathbf{a}\cdot\mathbf{b}\end{aligned}$$

and so comparing the above two formulas,

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\cos\theta. \tag{3.11}$$

In words, the dot product of two vectors equals the product of the magnitude of the two vectors multiplied by the cosine of the included angle. Note this gives a geometric description of the dot product which does not depend explicitly on the coordinates of the vectors.

Example 3.2.1 *Find the angle between the vectors* $2\mathbf{i}+\mathbf{j}-\mathbf{k}$ *and* $3\mathbf{i}+4\mathbf{j}+\mathbf{k}$.

The dot product of these two vectors equals $6+4-1=9$ and the norms are $\sqrt{4+1+1}=\sqrt{6}$ and $\sqrt{9+16+1}=\sqrt{26}$. Therefore, from 3.11 the cosine of the included angle equals

$$\cos\theta = \frac{9}{\sqrt{26}\sqrt{6}} = .720\,58$$

Now the cosine is known, the angle can be determines by solving the equation, $\cos\theta = .720\,58$. This will involve using a calculator or a table of trigonometric functions. The answer is $\theta = .766\,16$ radians or in terms of degrees, $\theta = .766\,16 \times \frac{360}{2\pi} = 43.898^\circ$. Recall how this last computation is done. Set up a proportion, $\frac{x}{.76616} = \frac{360}{2\pi}$ because 360° corresponds to 2π radians. However, in calculus, you should get used to thinking in terms of radians and not degrees. This is because all the important calculus formulas are defined in terms of radians.

Example 3.2.2 *Let* $\mathbf{u}, \mathbf{v}$ *be two vectors whose magnitudes are equal to 3 and 4 respectively and such that if they are placed in standard position with their tails at the origin, the angle between* $\mathbf{u}$ *and the positive* x *axis equals 30°* *and the angle between* $\mathbf{v}$ *and the positive* x *axis is* $-30°$. *Find* $\mathbf{u} \cdot \mathbf{v}$.

From the geometric description of the dot product in 3.11

$$\mathbf{u} \cdot \mathbf{v} = 3 \times 4 \times \cos(60°) = 3 \times 4 \times 1/2 = 6.$$

Observation 3.2.3 *Two vectors are said to be* ***perpendicular*** *if the included angle is* $\pi/2$ *radians (90°). You can tell if two nonzero vectors are perpendicular by simply taking their dot product. If the answer is zero, this means they are perpendicular because* $\cos\theta = 0$.

Example 3.2.4 *Determine whether the two vectors,* $2\mathbf{i} + \mathbf{j} - \mathbf{k}$ *and* $1\mathbf{i} + 3\mathbf{j} + 5\mathbf{k}$ *are perpendicular.*

When you take this dot product you get $2 + 3 - 5 = 0$ and so these two are indeed perpendicular.

Definition 3.2.5 *When two lines intersect, the angle between the two lines is the smaller of the two angles determined.*

Example 3.2.6 *Find the angle between the two lines,* $(1, 2, 0) + t(1, 2, 3)$ *and* $(0, 4, -3) + t(-1, 2, -3)$.

These two lines intersect, when $t = 0$ in the first and $t = -1$ in the second. It is only a matter of finding the angle between the direction vectors. One angle determined is given by

$$\cos\theta = \frac{-6}{14} = \frac{-3}{7}. \tag{3.12}$$

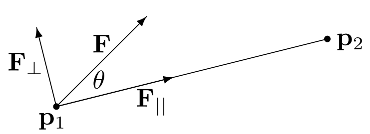
We don't want this angle because it is obtuse. The angle desired is the acute angle given by

$$\cos\theta = \frac{3}{7}.$$

It is obtained by replacing one of the direction vectors with -1 times it.

3.2.2 Work And Projections

Our first application will be to the concept of work. The physical concept of work does not in any way correspond to the notion of work employed in ordinary conversation. For example, if you were to slide a 150 pound weight off a table which is three feet high and shuffle along the floor for 50 yards, sweating profusely and exerting all your strength to keep the weight from falling on your feet, keeping the height always three feet and then deposit this weight on another three foot high table, the physical concept of work would indicate that the force exerted by your arms did no work during this project even though the muscles in your hands and arms would likely be very tired. The reason for such an unusual definition is that even though your arms exerted considerable force on the weight, enough to keep it from falling, the direction of motion was at right angles to the force they exerted. The only part of a force which does work in the sense of physics is the component of the force in the direction of motion (This is made more precise below.). The work is defined to be the magnitude of the component of this force times the distance over which it acts in the case where this component of force points in the direction of motion and (-1) times the magnitude of this component times the distance in case the force tends to impede the motion. Thus the work done by a force on an object as the object moves from one point to another is a measure of the extent to which the force contributes to the motion. This is illustrated in the following picture in the case where the given force contributes to the motion.



In this picture the force, $\mathbf{F}$ is applied to an object which moves on the straight line from $\mathbf{p}_1$ to $\mathbf{p}_2$. There are two vectors shown, $\mathbf{F}_{||}$ and $\mathbf{F}_{\perp}$ and the picture is intended to indicate that when you add these two vectors you get $\mathbf{F}$ while $\mathbf{F}_{||}$ acts in the direction of motion and $\mathbf{F}_{\perp}$ acts perpendicular to the direction of motion. Only $\mathbf{F}_{||}$ contributes to the work done by $\mathbf{F}$ on the object as it moves from $\mathbf{p}_1$ to $\mathbf{p}_2$. $\mathbf{F}_{||}$ is called the **component of the force** in the direction of motion. From trigonometry, you see the magnitude of $\mathbf{F}_{||}$ should equal $|\mathbf{F}|\,|\cos\theta|$. Thus, since $\mathbf{F}_{||}$ points in the direction of the vector from $\mathbf{p}_1$ to $\mathbf{p}_2$, the total work done should equal

$$|\mathbf{F}|\left|\overrightarrow{\mathbf{p}_1\mathbf{p}_2}\right|\cos\theta = |\mathbf{F}|\,|\mathbf{p}_2-\mathbf{p}_1|\cos\theta$$

If the included angle had been obtuse, then the work done by the force, $\mathbf{F}$ on the object would have been negative because in this case, the force tends to impede the motion from $\mathbf{p}_1$ to $\mathbf{p}_2$ but in this case, $\cos\theta$ would also be negative and so it is still the case that the work done would be given by the above formula. Thus from the geometric description of the dot product given above, the work equals

$$|\mathbf{F}|\,|\mathbf{p}_2-\mathbf{p}_1|\cos\theta = \mathbf{F}\cdot(\mathbf{p}_2-\mathbf{p}_1).$$

This explains the following definition.

Definition 3.2.7 *Let $\mathbf{F}$ be a force acting on an object which moves from the point $\mathbf{p}_1$ to the point $\mathbf{p}_2$. Then the **work** done on the object by the given force equals $\mathbf{F}\cdot(\mathbf{p}_2-\mathbf{p}_1)$.*

The concept of writing a given vector $\mathbf{F}$ in terms of two vectors, one which is parallel to a given vector $\mathbf{D}$ and the other which is perpendicular can also be explained with no reliance on trigonometry, completely in terms of the algebraic properties of the dot product. As before, this is mathematically more significant than any approach involving geometry or trigonometry because it extends to more interesting situations. This is done next.

Theorem 3.2.8 *Let $\mathbf{F}$ and $\mathbf{D}$ be nonzero vectors. Then there exist unique vectors $\mathbf{F}_{||}$ and $\mathbf{F}_{\perp}$ such that*

$$\mathbf{F} = \mathbf{F}_{||} + \mathbf{F}_{\perp} \tag{3.13}$$

where $\mathbf{F}_{||}$ is a scalar multiple of $\mathbf{D}$, also referred to as

$$\mathrm{proj}_{\mathbf{D}}(\mathbf{F}),$$

*and $\mathbf{F}_{\perp}\cdot\mathbf{D}=0$. The vector $\mathrm{proj}_{\mathbf{D}}(\mathbf{F})$ is called the **projection** of $\mathbf{F}$ onto $\mathbf{D}$.*

Proof: Suppose 3.13 and $\mathbf{F}_{||}=\alpha\mathbf{D}$. Taking the dot product of both sides with $\mathbf{D}$ and using $\mathbf{F}_{\perp}\cdot\mathbf{D}=0$, this yields

$$\mathbf{F}\cdot\mathbf{D} = \alpha\,|\mathbf{D}|^2$$

which requires $\alpha = \mathbf{F}\cdot\mathbf{D}/|\mathbf{D}|^2$. Thus there can be no more than one vector $\mathbf{F}_{||}$. It follows $\mathbf{F}_{\perp}$ must equal $\mathbf{F}-\mathbf{F}_{||}$. This verifies there can be no more than one choice for both $\mathbf{F}_{||}$ and $\mathbf{F}_{\perp}$.

Now let

$$\mathbf{F}_{||} \equiv \frac{\mathbf{F}\cdot\mathbf{D}}{|\mathbf{D}|^2}\mathbf{D}$$

and let

$$\mathbf{F}_{\perp} = \mathbf{F}-\mathbf{F}_{||} = \mathbf{F}-\frac{\mathbf{F}\cdot\mathbf{D}}{|\mathbf{D}|^2}\mathbf{D}$$

Then $\mathbf{F}_{||}=\alpha\,\mathbf{D}$ where $\alpha=\frac{\mathbf{F}\cdot\mathbf{D}}{|\mathbf{D}|^2}$. It only remains to verify $\mathbf{F}_{\perp}\cdot\mathbf{D}=0$. But

$$\begin{aligned}\mathbf{F}_{\perp}\cdot\mathbf{D} &= \mathbf{F}\cdot\mathbf{D}-\frac{\mathbf{F}\cdot\mathbf{D}}{|\mathbf{D}|^2}\mathbf{D}\cdot\mathbf{D}\\ &= \mathbf{F}\cdot\mathbf{D}-\mathbf{F}\cdot\mathbf{D}=0.\end{aligned}$$

■

Example 3.2.9 *Let* $\mathbf{F} = 2\mathbf{i}+7\mathbf{j} - 3\mathbf{k}$ *Newtons. Find the work done by this force in moving from the point* $(1, 2, 3)$ *to the point* $(-9, -3, 4)$ *along the straight line segment joining these points where distances are measured in meters.*

According to the definition, this work is

$$\begin{aligned}(2\mathbf{i}+7\mathbf{j} - 3\mathbf{k}) \cdot (-10\mathbf{i} - 5\mathbf{j} + \mathbf{k}) &= -20 + (-35) + (-3) \\ &= -58 \text{ Newton meters.}\end{aligned}$$

Note that if the force had been given in pounds and the distance had been given in feet, the units on the work would have been foot pounds. In general, work has units equal to units of a force times units of a length. Instead of writing Newton meter, people write joule because a joule is by definition a Newton meter. That word is pronounced "jewel" and it is the unit of work in the metric system of units. Also be sure you observe that the work done by the force can be negative as in the above example. In fact, work can be either positive, negative, or zero. You just have to do the computations to find out.

Example 3.2.10 *Find* $\text{proj}_{\mathbf{u}}(\mathbf{v})$ *if* $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$ *and* $\mathbf{v} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$.

From the above discussion in Theorem 3.2.8, this is just

$$\begin{aligned}&\frac{1}{4+9+16}(\mathbf{i} - 2\mathbf{j} + \mathbf{k}) \cdot (2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k})(2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}) \\ = &\frac{-8}{29}(2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}) = -\frac{16}{29}\mathbf{i} - \frac{24}{29}\mathbf{j} + \frac{32}{29}\mathbf{k}.\end{aligned}$$

Example 3.2.11 *Suppose* $\mathbf{a}$, *and* $\mathbf{b}$ *are vectors and* $\mathbf{b}_\perp = \mathbf{b} - \text{proj}_{\mathbf{a}}(\mathbf{b})$. *What is the magnitude of* $\mathbf{b}_\perp$ *in terms of the included angle?*

$$\begin{aligned}|\mathbf{b}_\perp|^2 &= (\mathbf{b} - \text{proj}_{\mathbf{a}}(\mathbf{b})) \cdot (\mathbf{b} - \text{proj}_{\mathbf{a}}(\mathbf{b})) = \left(\mathbf{b} - \frac{\mathbf{b} \cdot \mathbf{a}}{|\mathbf{a}|^2}\mathbf{a}\right) \cdot \left(\mathbf{b} - \frac{\mathbf{b} \cdot \mathbf{a}}{|\mathbf{a}|^2}\mathbf{a}\right) \\ &= |\mathbf{b}|^2 - 2\frac{(\mathbf{b} \cdot \mathbf{a})^2}{|\mathbf{a}|^2} + \left(\frac{\mathbf{b} \cdot \mathbf{a}}{|\mathbf{a}|^2}\right)^2 |\mathbf{a}|^2 = |\mathbf{b}|^2 \left(1 - \frac{(\mathbf{b} \cdot \mathbf{a})^2}{|\mathbf{a}|^2 |\mathbf{b}|^2}\right) \\ &= |\mathbf{b}|^2 \left(1 - \cos^2\theta\right) = |\mathbf{b}|^2 \sin^2(\theta)\end{aligned}$$

where θ is the included angle between $\mathbf{a}$ and $\mathbf{b}$ which is less than π radians. Therefore, taking square roots,

$$|\mathbf{b}_\perp| = |\mathbf{b}| \sin\theta.$$

3.2.3 The Inner Product And Distance In $\mathbb{C}^n$

It is necessary to give a generalization of the dot product for vectors in $\mathbb{C}^n$. This is often called the inner product. It reduces to the definition of the dot product in the case the components of the vector are real.

Definition 3.2.12 *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. *Thus* $\mathbf{x} = (x_1, \cdots, x_n)$ *where each* $x_k \in \mathbb{C}$ *and a similar formula holding for* $\mathbf{y}$. *Then the inner product of these two vectors is defined to be*

$$\mathbf{x} \cdot \mathbf{y} \equiv \sum_j x_j \overline{y_j} \equiv x_1 \overline{y_1} + \cdots + x_n \overline{y_n}.$$

The inner product is often denoted as $(\mathbf{x}, \mathbf{y})$ *or* $\langle \mathbf{x}, \mathbf{y} \rangle$.

Notice how you put the conjugate on the entries of the vector $\mathbf{y}$. It makes no difference if the vectors happen to be real vectors but with complex vectors you must do it this way. The reason for this is that when you take the inner product of a vector with itself, you want to get the square of the length of the vector, a positive number. Placing the conjugate on the components of $\mathbf{y}$ in the above definition assures this will take place. Thus

$$\mathbf{x}\cdot\mathbf{x} = \sum_j x_j \overline{x_j} = \sum_j |x_j|^2 \geq 0.$$

If you didn't place a conjugate as in the above definition, things wouldn't work out correctly. For example,

$$(1+i)^2 + 2^2 = 4 + 2i$$

and this is not a positive number.

The following properties of the inner product follow immediately from the definition and you should verify each of them.

Properties of the inner product:

1. $\mathbf{u}\cdot\mathbf{v} = \overline{\mathbf{v}\cdot\mathbf{u}}$.
2. If a, b are numbers and $\mathbf{u}, \mathbf{v}, \mathbf{z}$ are vectors then $(a\mathbf{u} + b\mathbf{v})\cdot\mathbf{z} = a(\mathbf{u}\cdot\mathbf{z}) + b(\mathbf{v}\cdot\mathbf{z})$.
3. $\mathbf{u}\cdot\mathbf{u} \geq 0$ and it equals 0 if and only if $\mathbf{u} = \mathbf{0}$.

Note this implies $(\mathbf{x}\cdot\alpha\mathbf{y}) = \overline{\alpha}(\mathbf{x}\cdot\mathbf{y})$ because

$$(\mathbf{x}\cdot\alpha\mathbf{y}) = \overline{(\alpha\mathbf{y}\cdot\mathbf{x})} = \overline{\alpha(\mathbf{y}\cdot\mathbf{x})} = \overline{\alpha}(\mathbf{x}\cdot\mathbf{y})$$

The norm is defined in the usual way.

Definition 3.2.13 *For* $\mathbf{x} \in \mathbb{C}^n$,

$$|\mathbf{x}| \equiv \left(\sum_{k=1}^n |x_k|^2\right)^{1/2} = (\mathbf{x}\cdot\mathbf{x})^{1/2}$$

Here is a fundamental inequality called the **Cauchy Schwarz inequality** which is stated here in $\mathbb{C}^n$. First here is a simple lemma.

Lemma 3.2.14 *If* $z \in \mathbb{C}$ *there exists* $\theta \in \mathbb{C}$ *such that* $\theta z = |z|$ *and* $|\theta| = 1$.

Proof: Let $\theta = 1$ if $z = 0$ and otherwise, let $\theta = \dfrac{\overline{z}}{|z|}$. Recall that for $z = x + iy, \overline{z} = x - iy$ and $\overline{z}z = |z|^2$.

I will give a proof of this important inequality which depends only on the above list of properties of the inner product. It will be slightly different than the earlier proof.

Theorem 3.2.15 *(Cauchy Schwarz)The following inequality holds for* $\mathbf{x}$ *and* $\mathbf{y} \in \mathbb{C}^n$.

$$|(\mathbf{x}\cdot\mathbf{y})| \leq (\mathbf{x}\cdot\mathbf{x})^{1/2}(\mathbf{y}\cdot\mathbf{y})^{1/2} \tag{3.14}$$

Equality holds in this inequality if and only if one vector is a multiple of the other.

Proof: Let $\theta \in \mathbb{C}$ such that $|\theta| = 1$ and

$$\theta(\mathbf{x}\cdot\mathbf{y}) = |(\mathbf{x}\cdot\mathbf{y})|$$

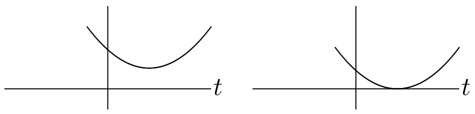
Consider $p(t) \equiv \left(\mathbf{x}+\overline{\theta}t\mathbf{y}, \mathbf{x}+t\overline{\theta}\mathbf{y}\right)$ where $t \in \mathbb{R}$. Then from the above list of properties of the dot product,

$$\begin{aligned}
0 &\leq p(t) = (\mathbf{x}\cdot\mathbf{x}) + t\theta(\mathbf{x}\cdot\mathbf{y}) + t\overline{\theta}(\mathbf{y}\cdot\mathbf{x}) + t^2(\mathbf{y}\cdot\mathbf{y}) \\
&= (\mathbf{x}\cdot\mathbf{x}) + t\theta(\mathbf{x}\cdot\mathbf{y}) + t\overline{\theta(\mathbf{x}\cdot\mathbf{y})} + t^2(\mathbf{y}\cdot\mathbf{y}) \\
&= (\mathbf{x}\cdot\mathbf{x}) + 2t\,\mathrm{Re}(\theta(\mathbf{x}\cdot\mathbf{y})) + t^2(\mathbf{y}\cdot\mathbf{y}) \\
&= (\mathbf{x}\cdot\mathbf{x}) + 2t\,|(\mathbf{x}\cdot\mathbf{y})| + t^2(\mathbf{y}\cdot\mathbf{y}) \qquad (3.15)
\end{aligned}$$

and this must hold for all $t \in \mathbb{R}$. Therefore, if $(\mathbf{y}\cdot\mathbf{y}) = 0$ it must be the case that $|(\mathbf{x}\cdot\mathbf{y})| = 0$ also since otherwise the above inequality would be violated. Therefore, in this case,

$$|(\mathbf{x}\cdot\mathbf{y})| \leq (\mathbf{x}\cdot\mathbf{x})^{1/2}(\mathbf{y}\cdot\mathbf{y})^{1/2}.$$

On the other hand, if $(\mathbf{y}\cdot\mathbf{y}) \neq 0$, then $p(t) \geq 0$ for all t means the graph of $y = p(t)$ is a parabola which opens up and it either has exactly one real zero in the case its vertex touches the t axis or it has no real zeros.



From the quadratic formula this happens exactly when

$$4|(\mathbf{x}\cdot\mathbf{y})|^2 - 4(\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y}) \leq 0$$

which is equivalent to 3.14.

It is clear from a computation that if one vector is a scalar multiple of the other that equality holds in 3.14. Conversely, suppose equality does hold. Then this is equivalent to saying $4|(\mathbf{x}\cdot\mathbf{y})|^2 - 4(\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y}) = 0$ and so from the quadratic formula, there exists one real zero to $p(t) = 0$. Call it t_0. Then

$$p(t_0) \equiv \left(\left(\mathbf{x}+\overline{\theta}t_0\mathbf{y}\right)\cdot\left(\mathbf{x}+t_0\overline{\theta}\mathbf{y}\right)\right) = \left|\mathbf{x}+\overline{\theta}t\mathbf{y}\right|^2 = 0$$

and so $\mathbf{x} = -\overline{\theta}t_0\mathbf{y}$. ■

Note that I only used part of the above properties of the inner product. It was not necessary to use the one which says that if $(\mathbf{x}\cdot\mathbf{x}) = 0$ then $\mathbf{x} = \mathbf{0}$.

By analogy to the case of $\mathbb{R}^n$, length or magnitude of vectors in $\mathbb{C}^n$ can be defined.

Definition 3.2.16 *Let $\mathbf{z} \in \mathbb{C}^n$. Then $|\mathbf{z}| \equiv (\mathbf{z}\cdot\mathbf{z})^{1/2}$.*

The conclusions of the following theorem are also called the **axioms for a norm.**

Theorem 3.2.17 *For length defined in Definition 3.2.16, the following hold.*

$$|\mathbf{z}| \geq 0 \text{ and } |\mathbf{z}| = 0 \text{ if and only if } \mathbf{z} = \mathbf{0} \qquad (3.16)$$

$$\text{If } \alpha \text{ is a scalar, } |\alpha\mathbf{z}| = |\alpha|\,|\mathbf{z}| \qquad (3.17)$$

$$|\mathbf{z}+\mathbf{w}| \leq |\mathbf{z}| + |\mathbf{w}|. \qquad (3.18)$$

Proof: The first two claims are left as exercises. To establish the third, you use the same argument which was used in $\mathbb{R}^n$.

$$\begin{aligned}
|\mathbf{z}+\mathbf{w}|^2 &= (\mathbf{z}+\mathbf{w}, \mathbf{z}+\mathbf{w}) \\
&= \mathbf{z}\cdot\mathbf{z} + \mathbf{w}\cdot\mathbf{w} + \mathbf{w}\cdot\mathbf{z} + \mathbf{z}\cdot\mathbf{w} \\
&= |\mathbf{z}|^2 + |\mathbf{w}|^2 + 2\,\mathrm{Re}\,\mathbf{w}\cdot\mathbf{z} \\
&\leq |\mathbf{z}|^2 + |\mathbf{w}|^2 + 2\,|\mathbf{w}\cdot\mathbf{z}| \\
&\leq |\mathbf{z}|^2 + |\mathbf{w}|^2 + 2\,|\mathbf{w}|\,|\mathbf{z}| = (|\mathbf{z}| + |\mathbf{w}|)^2. \ ■
\end{aligned}$$

Occasionally, I may refer to the inner product in $\mathbb{C}^n$ as the dot product. They are the same thing for $\mathbb{R}^n$. However, it is convenient to draw a distinction when discussing matrix multiplication a little later.

3.3 Exercises

1. Use formula 3.11 to verify the Cauchy Schwarz inequality and to show that equality occurs if and only if one of the vectors is a scalar multiple of the other.

2. For $\mathbf{u}, \mathbf{v}$ vectors in $\mathbb{R}^3$, define the product, $\mathbf{u} * \mathbf{v} \equiv u_1 v_1 + 2u_2 v_2 + 3u_3 v_3$. Show the axioms for a dot product all hold for this funny product. Prove
$$|\mathbf{u} * \mathbf{v}| \leq (\mathbf{u} * \mathbf{u})^{1/2} (\mathbf{v} * \mathbf{v})^{1/2}.$$
Hint: Do not try to do this with methods from trigonometry.

3. Find the angle between the vectors $3\mathbf{i} - \mathbf{j} - \mathbf{k}$ and $\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$.

4. Find the angle between the vectors $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$ and $\mathbf{i} + 2\mathbf{j} - 7\mathbf{k}$.

5. Find $\text{proj}_{\mathbf{u}} (\mathbf{v})$ where $\mathbf{v} = (1, 0, -2)$ and $\mathbf{u} = (1, 2, 3)$.

6. Find $\text{proj}_{\mathbf{u}} (\mathbf{v})$ where $\mathbf{v} = (1, 2, -2)$ and $\mathbf{u} = (1, 0, 3)$.

7. Find $\text{proj}_{\mathbf{u}} (\mathbf{v})$ where $\mathbf{v} = (1, 2, -2, 1)$ and $\mathbf{u} = (1, 2, 3, 0)$.

8. Does it make sense to speak of $\text{proj}_{\mathbf{0}} (\mathbf{v})$?

9. If $\mathbf{F}$ is a force and $\mathbf{D}$ is a vector, show $\text{proj}_{\mathbf{D}} (\mathbf{F}) = (|\mathbf{F}| \cos \theta) \mathbf{u}$ where $\mathbf{u}$ is the unit vector in the direction of $\mathbf{D}$, $\mathbf{u} = \mathbf{D}/|\mathbf{D}|$ and θ is the included angle between the two vectors, $\mathbf{F}$ and $\mathbf{D}$. $|\mathbf{F}| \cos \theta$ is sometimes called the component of the force, $\mathbf{F}$ in the direction, $\mathbf{D}$.

10. Prove the Cauchy Schwarz inequality in $\mathbb{R}^n$ as follows. For $\mathbf{u}, \mathbf{v}$ vectors, consider
$$(\mathbf{u} - \text{proj}_{\mathbf{v}} \mathbf{u}) \cdot (\mathbf{u} - \text{proj}_{\mathbf{v}} \mathbf{u}) \geq 0$$
Now simplify using the axioms of the dot product and then put in the formula for the projection. Of course this expression equals 0 and you get equality in the Cauchy Schwarz inequality if and only if $\mathbf{u} = \text{proj}_{\mathbf{v}} \mathbf{u}$. What is the geometric meaning of $\mathbf{u} = \text{proj}_{\mathbf{v}} \mathbf{u}$?

11. A boy drags a sled for 100 feet along the ground by pulling on a rope which is 20 degrees from the horizontal with a force of 40 pounds. How much work does this force do?

12. A girl drags a sled for 200 feet along the ground by pulling on a rope which is 30 degrees from the horizontal with a force of 20 pounds. How much work does this force do?

13. A large dog drags a sled for 300 feet along the ground by pulling on a rope which is 45 degrees from the horizontal with a force of 20 pounds. How much work does this force do?

14. How much work in Newton meters does it take to slide a crate 20 meters along a loading dock by pulling on it with a 200 Newton force at an angle of $30°$ from the horizontal?

15. An object moves 10 meters in the direction of $\mathbf{j}$. There are two forces acting on this object, $\mathbf{F}_1 = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$, and $\mathbf{F}_2 = -5\mathbf{i} + 2\mathbf{j} - 6\mathbf{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force. Why?

16. An object moves 10 meters in the direction of $\mathbf{j} + \mathbf{i}$. There are two forces acting on this object, $\mathbf{F}_1 = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$, and $\mathbf{F}_2 = 5\mathbf{i} + 2\mathbf{j} - 6\mathbf{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force. Why?

17. An object moves 20 meters in the direction of $\mathbf{k} + \mathbf{j}$. There are two forces acting on this object, $\mathbf{F}_1 = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$, and $\mathbf{F}_2 = \mathbf{i} + 2\mathbf{j} - 6\mathbf{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force.

18. If $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are vectors. Show that $(\mathbf{b}+\mathbf{c})_{\perp} = \mathbf{b}_{\perp} + \mathbf{c}_{\perp}$ where $\mathbf{b}_{\perp} = \mathbf{b} - \operatorname{proj}_{\mathbf{a}}(\mathbf{b})$.

19. Find $(1,2,3,4) \cdot (2,0,1,3)$.

20. Show that $(\mathbf{a} \cdot \mathbf{b}) = \frac{1}{4}\left[|\mathbf{a}+\mathbf{b}|^2 - |\mathbf{a}-\mathbf{b}|^2\right]$.

21. Prove from the axioms of the dot product the parallelogram identity, $|\mathbf{a}+\mathbf{b}|^2 + |\mathbf{a}-\mathbf{b}|^2 = 2|\mathbf{a}|^2 + 2|\mathbf{b}|^2$.

22. Recall that the open ball having center at $\mathbf{a}$ and radius r is given by

$$B(\mathbf{a},r) \equiv \{\mathbf{x} : |\mathbf{x}-\mathbf{a}| < r\}$$

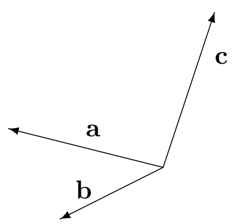
Show that if $\mathbf{y} \in B(\mathbf{a}, r)$, then there exists a positive number δ such that $B(\mathbf{y}, \delta) \subseteq B(\mathbf{a}, r)$. (The symbol $\subseteq$ means that every point in $B(\mathbf{y}, \delta)$ is also in $B(\mathbf{a}, r)$. In words, it states that $B(\mathbf{y}, \delta)$ is contained in $B(\mathbf{a}, r)$. The statement $\mathbf{y} \in B(\mathbf{a}, r)$ says that $\mathbf{y}$ is one of the points of $B(\mathbf{a}, r)$.) When you have done this, you will have shown that an open ball is open. This is a fantastically important observation although its major implications will not be explored very much in this book.

3.4 The Cross Product

The cross product is the other way of multiplying two vectors in $\mathbb{R}^3$. It is very different from the dot product in many ways. First the geometric meaning is discussed and then a description in terms of coordinates is given. Both descriptions of the cross product are important. The geometric description is essential in order to understand the applications to physics and geometry while the coordinate description is the only way to practically compute the cross product.

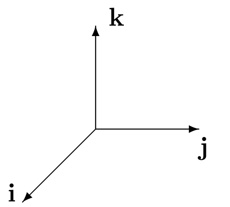
Definition 3.4.1 *Three vectors,* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *form a right handed system if when you extend the fingers of your right hand along the vector* $\mathbf{a}$ *and close them in the direction of* $\mathbf{b}$, *the thumb points roughly in the direction of* $\mathbf{c}$.

For an example of a right handed system of vectors, see the following picture.



In this picture the vector $\mathbf{c}$ points upwards from the plane determined by the other two vectors. You should consider how a right hand system would differ from a left hand system. Try using your left hand and you will see that the vector $\mathbf{c}$ would need to point in the opposite direction as it would for a right hand system.

From now on, the vectors, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ will always form a right handed system. To repeat, if you extend the fingers of your right hand along $\mathbf{i}$ and close them in the direction $\mathbf{j}$, the thumb points in the direction of $\mathbf{k}$.



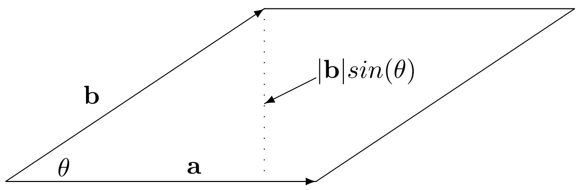
The following is the geometric description of the cross product. It gives both the direction and the magnitude and therefore specifies the vector.

Definition 3.4.2 *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be two vectors in* $\mathbb{R}^3$*. Then* $\mathbf{a}\times\mathbf{b}$ *is defined by the following two rules.*

1. $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\sin\theta$ *where* θ *is the included angle.*

2. $\mathbf{a}\times\mathbf{b}\cdot\mathbf{a} = 0$, $\mathbf{a}\times\mathbf{b}\cdot\mathbf{b} = 0$, *and* $\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}$ *forms a right hand system.*

Note that $|\mathbf{a}\times\mathbf{b}|$ is the area of the parallelogram determined by $\mathbf{a}$ and $\mathbf{b}$.



The cross product satisfies the following properties.

$$\mathbf{a}\times\mathbf{b} = -(\mathbf{b}\times\mathbf{a})\ ,\ \mathbf{a}\times\mathbf{a} = \mathbf{0}, \tag{3.19}$$

For α a scalar,

$$(\alpha\mathbf{a})\times\mathbf{b} = \alpha(\mathbf{a}\times\mathbf{b}) = \mathbf{a}\times(\alpha\mathbf{b}), \tag{3.20}$$

For $\mathbf{a},\mathbf{b}$, and $\mathbf{c}$ vectors, one obtains the distributive laws,

$$\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}, \tag{3.21}$$

$$(\mathbf{b}+\mathbf{c})\times\mathbf{a} = \mathbf{b}\times\mathbf{a}+\mathbf{c}\times\mathbf{a}. \tag{3.22}$$

Formula 3.19 follows immediately from the definition. The vectors $\mathbf{a}\times\mathbf{b}$ and $\mathbf{b}\times\mathbf{a}$ have the same magnitude, $|\mathbf{a}|\,|\mathbf{b}|\sin\theta$, and an application of the right hand rule shows they have opposite direction. Formula 3.20 is also fairly clear. If α is a nonnegative scalar, the direction of $(\alpha\mathbf{a})\times\mathbf{b}$ is the same as the direction of $\mathbf{a}\times\mathbf{b}$,$\alpha(\mathbf{a}\times\mathbf{b})$ and $\mathbf{a}\times(\alpha\mathbf{b})$ while the magnitude is just α times the magnitude of $\mathbf{a}\times\mathbf{b}$ which is the same as the magnitude of $\alpha(\mathbf{a}\times\mathbf{b})$ and $\mathbf{a}\times(\alpha\mathbf{b})$. Using this yields equality in 3.20. In the case where $\alpha<0$, everything works the same way except the vectors are all pointing in the opposite direction and you must multiply by $|\alpha|$ when comparing their magnitudes. The distributive laws are much harder to establish but the second follows from the first quite easily. Thus, assuming the first, and using 3.19,

$$\begin{aligned}(\mathbf{b}+\mathbf{c})\times\mathbf{a} &= -\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = -(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c})\\ &= \mathbf{b}\times\mathbf{a}+\mathbf{c}\times\mathbf{a}.\end{aligned}$$

A proof of the distributive law is given in a later section for those who are interested.

Now from the definition of the cross product,

$$\begin{array}{ll}\mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{j}\times\mathbf{i}=-\mathbf{k}\\ \mathbf{k}\times\mathbf{i}=\mathbf{j} & \mathbf{i}\times\mathbf{k}=-\mathbf{j}\\ \mathbf{j}\times\mathbf{k}=\mathbf{i} & \mathbf{k}\times\mathbf{j}=-\mathbf{i}\end{array}$$

With this information, the following gives the coordinate description of the cross product.

Proposition 3.4.3 *Let* $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ *and* $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ *be two vectors. Then*

$$\begin{aligned} \mathbf{a} \times \mathbf{b} = & (a_2b_3 - a_3b_2)\,\mathbf{i} + (a_3b_1 - a_1b_3)\,\mathbf{j} + \\ & + (a_1b_2 - a_2b_1)\,\mathbf{k}. \end{aligned} \tag{3.23}$$

Proof: From the above table and the properties of the cross product listed,

$$(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) =$$

$$\begin{aligned} & a_1b_2\mathbf{i} \times \mathbf{j} + a_1b_3\mathbf{i} \times \mathbf{k} + a_2b_1\mathbf{j} \times \mathbf{i} + a_2b_3\mathbf{j} \times \mathbf{k} + \\ & + a_3b_1\mathbf{k} \times \mathbf{i} + a_3b_2\mathbf{k} \times \mathbf{j} \\ & = a_1b_2\mathbf{k} - a_1b_3\mathbf{j} - a_2b_1\mathbf{k} + a_2b_3\mathbf{i} + a_3b_1\mathbf{j} - a_3b_2\mathbf{i} \\ & = (a_2b_3 - a_3b_2)\,\mathbf{i} + (a_3b_1 - a_1b_3)\,\mathbf{j} + (a_1b_2 - a_2b_1)\,\mathbf{k} \end{aligned} \tag{3.24}$$

■

It is probably impossible for most people to remember 3.23. Fortunately, there is a somewhat easier way to remember it. Define the determinant of a 2×2 matrix as follows

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \equiv ad - bc$$

Then

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{3.25}$$

where you expand the determinant along the top row. This yields

$$\mathbf{i}\,(-1)^{1+1} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \mathbf{j}\,(-1)^{2+1} \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \mathbf{k}\,(-1)^{3+1} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

$$= \mathbf{i} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \mathbf{j} \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \mathbf{k} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

Note that to get the scalar which multiplies $\mathbf{i}$ you take the determinant of what is left after deleting the first row and the first column and multiply by $(-1)^{1+1}$ because $\mathbf{i}$ is in the first row and the first column. Then you do the same thing for the $\mathbf{j}$ and $\mathbf{k}$. In the case of the $\mathbf{j}$ there is a minus sign because $\mathbf{j}$ is in the first row and the second column and so$(-1)^{1+2} = -1$ while the $\mathbf{k}$ is multiplied by $(-1)^{3+1} = 1$. The above equals

$$(a_2b_3 - a_3b_2)\,\mathbf{i} - (a_1b_3 - a_3b_1)\,\mathbf{j} + (a_1b_2 - a_2b_1)\,\mathbf{k} \tag{3.26}$$

which is the same as 3.24. There will be much more presented on determinants later. For now, consider this an introduction if you have not seen this topic.

Example 3.4.4 *Find* $(\mathbf{i} - \mathbf{j} + 2\mathbf{k}) \times (3\mathbf{i} - 2\mathbf{j} + \mathbf{k})$.

Use 3.25 to compute this.

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -1 & 2 \\ 3 & -2 & 1 \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ -2 & 1 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & -1 \\ 3 & -2 \end{vmatrix} \mathbf{k} = 3\mathbf{i} + 5\mathbf{j} + \mathbf{k}.$$

Example 3.4.5 *Find the area of the parallelogram determined by the vectors,*

$$(\mathbf{i} - \mathbf{j} + 2\mathbf{k}),\ (3\mathbf{i} - 2\mathbf{j} + \mathbf{k}).$$

These are the same two vectors in Example 3.4.4.

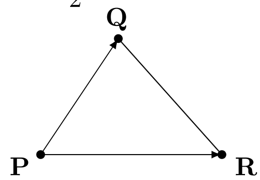
From Example 3.4.4 and the geometric description of the cross product, the area is just the norm of the vector obtained in Example 3.4.4. Thus the area is $\sqrt{9+25+1}=\sqrt{35}$.

Example 3.4.6 *Find the area of the triangle determined by* $(1,2,3),(0,2,5),(5,1,2)$.

This triangle is obtained by connecting the three points with lines. Picking $(1,2,3)$ as a starting point, there are two displacement vectors, $(-1,0,2)$ and $(4,-1,-1)$ such that the given vector added to these displacement vectors gives the other two vectors. The area of the triangle is half the area of the parallelogram determined by $(-1,0,2)$ and $(4,-1,-1)$. Thus $(-1,0,2)\times(4,-1,-1)=(2,7,1)$ and so the area of the triangle is $\frac{1}{2}\sqrt{4+49+1}=\frac{3}{2}\sqrt{6}$.

Observation 3.4.7 *In general, if you have three points (vectors) in* $\mathbb{R}^3$, $\mathbf{P},\mathbf{Q},\mathbf{R}$ *the area of the triangle is given by*

$$\frac{1}{2}\left|(\mathbf{Q}-\mathbf{P})\times(\mathbf{R}-\mathbf{P})\right|.$$

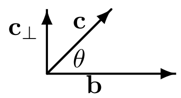


3.4.1 The Distributive Law For The Cross Product

This section gives a proof for 3.21, a fairly difficult topic. It is included here for the interested student. If you are satisfied with taking the distributive law on faith, it is not necessary to read this section. The proof given here is quite clever and follows the one given in [3]. Another approach, based on volumes of parallelepipeds is found in [15] and is discussed a little later.

Lemma 3.4.8 *Let* $\mathbf{b}$ *and* $\mathbf{c}$ *be two vectors. Then* $\mathbf{b}\times\mathbf{c}=\mathbf{b}\times\mathbf{c}_\perp$ *where* $\mathbf{c}_{||}+\mathbf{c}_\perp=\mathbf{c}$ *and* $\mathbf{c}_\perp\cdot\mathbf{b}=0$.

Proof: Consider the following picture.



Now $\mathbf{c}_\perp=\mathbf{c}-\mathbf{c}\cdot\frac{\mathbf{b}}{|\mathbf{b}|}\frac{\mathbf{b}}{|\mathbf{b}|}$ and so $\mathbf{c}_\perp$ is in the plane determined by $\mathbf{c}$ and $\mathbf{b}$. Therefore, from the geometric definition of the cross product, $\mathbf{b}\times\mathbf{c}$ and $\mathbf{b}\times\mathbf{c}_\perp$ have the same direction. Now, referring to the picture,

$$|\mathbf{b}\times\mathbf{c}_\perp|=|\mathbf{b}|\,|\mathbf{c}_\perp|=|\mathbf{b}|\,|\mathbf{c}|\sin\theta=|\mathbf{b}\times\mathbf{c}|.$$

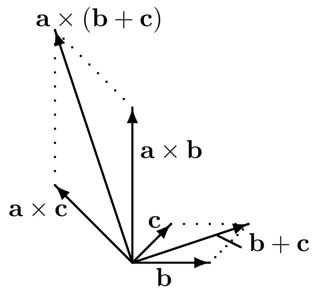
Therefore, $\mathbf{b}\times\mathbf{c}$ and $\mathbf{b}\times\mathbf{c}_\perp$ also have the same magnitude and so they are the same vector. ■

With this, the proof of the distributive law is in the following theorem.

Theorem 3.4.9 *Let* $\mathbf{a},\mathbf{b},$ *and* $\mathbf{c}$ *be vectors in* $\mathbb{R}^3$. *Then*

$$\mathbf{a}\times(\mathbf{b}+\mathbf{c})=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c} \tag{3.27}$$

Proof: Suppose first that $\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\mathbf{c}=0$. Now imagine $\mathbf{a}$ is a vector coming out of the page and let $\mathbf{b},\mathbf{c}$ and $\mathbf{b}+\mathbf{c}$ be as shown in the following picture.



Then $\mathbf{a}\times\mathbf{b}, \mathbf{a}\times(\mathbf{b}+\mathbf{c})$, and $\mathbf{a}\times\mathbf{c}$ are each vectors in the same plane, perpendicular to $\mathbf{a}$ as shown. Thus $\mathbf{a}\times\mathbf{c}\cdot\mathbf{c}=0, \mathbf{a}\times(\mathbf{b}+\mathbf{c})\cdot(\mathbf{b}+\mathbf{c})=0$, and $\mathbf{a}\times\mathbf{b}\cdot\mathbf{b}=0$. This implies that to get $\mathbf{a}\times\mathbf{b}$ you move counterclockwise through an angle of $\pi/2$ radians from the vector $\mathbf{b}$. Similar relationships exist between the vectors $\mathbf{a}\times(\mathbf{b}+\mathbf{c})$ and $\mathbf{b}+\mathbf{c}$ and the vectors $\mathbf{a}\times\mathbf{c}$ and $\mathbf{c}$. Thus the angle between $\mathbf{a}\times\mathbf{b}$ and $\mathbf{a}\times(\mathbf{b}+\mathbf{c})$ is the same as the angle between $\mathbf{b}+\mathbf{c}$ and $\mathbf{b}$ and the angle between $\mathbf{a}\times\mathbf{c}$ and $\mathbf{a}\times(\mathbf{b}+\mathbf{c})$ is the same as the angle between $\mathbf{c}$ and $\mathbf{b}+\mathbf{c}$. In addition to this, since $\mathbf{a}$ is perpendicular to these vectors,

$$|\mathbf{a}\times\mathbf{b}|=|\mathbf{a}|\,|\mathbf{b}|\,, |\mathbf{a}\times(\mathbf{b}+\mathbf{c})|=|\mathbf{a}|\,|\mathbf{b}+\mathbf{c}|\,, \text{ and}$$

$$|\mathbf{a}\times\mathbf{c}|=|\mathbf{a}|\,|\mathbf{c}|\,.$$

Therefore,

$$\frac{|\mathbf{a}\times(\mathbf{b}+\mathbf{c})|}{|\mathbf{b}+\mathbf{c}|}=\frac{|\mathbf{a}\times\mathbf{c}|}{|\mathbf{c}|}=\frac{|\mathbf{a}\times\mathbf{b}|}{|\mathbf{b}|}=|\mathbf{a}|$$

and so

$$\frac{|\mathbf{a}\times(\mathbf{b}+\mathbf{c})|}{|\mathbf{a}\times\mathbf{c}|}=\frac{|\mathbf{b}+\mathbf{c}|}{|\mathbf{c}|},\frac{|\mathbf{a}\times(\mathbf{b}+\mathbf{c})|}{|\mathbf{a}\times\mathbf{b}|}=\frac{|\mathbf{b}+\mathbf{c}|}{|\mathbf{b}|}$$

showing the triangles making up the parallelogram on the right and the four sided figure on the left in the above picture are similar. It follows the four sided figure on the left is in fact a parallelogram and this implies the diagonal is the vector sum of the vectors on the sides, yielding 3.27.

Now suppose it is not necessarily the case that $\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\mathbf{c}=0$. Then write $\mathbf{b}=\mathbf{b}_{||}+\mathbf{b}_{\perp}$ where $\mathbf{b}_{\perp}\cdot\mathbf{a}=0$. Similarly $\mathbf{c}=\mathbf{c}_{||}+\mathbf{c}_{\perp}$. By the above lemma and what was just shown,

$$\begin{aligned}\mathbf{a}\times(\mathbf{b}+\mathbf{c}) &= \mathbf{a}\times(\mathbf{b}+\mathbf{c})_{\perp}=\mathbf{a}\times(\mathbf{b}_{\perp}+\mathbf{c}_{\perp})\\ &= \mathbf{a}\times\mathbf{b}_{\perp}+\mathbf{a}\times\mathbf{c}_{\perp}=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}.\ \blacksquare\end{aligned}$$

The result of Problem 18 of the exercises 3.3 is used to go from the first to the second line.

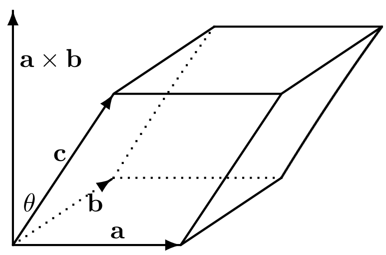
3.4.2 The Box Product

Definition 3.4.10 *A parallelepiped determined by the three vectors,* $\mathbf{a},\mathbf{b}$, *and* $\mathbf{c}$ *consists of*

$$\{r\mathbf{a}+s\mathbf{b}+t\mathbf{c}:r,s,t\in[0,1]\}\,.$$

That is, if you pick three numbers, r,s, *and* t *each in* $[0,1]$ *and form* $r\mathbf{a}+s\mathbf{b}+t\mathbf{c}$, *then the collection of all such points is what is meant by the parallelepiped determined by these three vectors.*

The following is a picture of such a thing.



You notice the area of the base of the parallelepiped, the parallelogram determined by the vectors, $\mathbf{a}$ and $\mathbf{b}$ has area equal to $|\mathbf{a}\times\mathbf{b}|$ while the altitude of the parallelepiped is $|\mathbf{c}|\cos\theta$ where θ is the angle shown in the picture between $\mathbf{c}$ and $\mathbf{a}\times\mathbf{b}$. Therefore, the volume of this parallelepiped is the area of the base times the altitude which is just

$$|\mathbf{a}\times\mathbf{b}|\,|\mathbf{c}|\cos\theta=\mathbf{a}\times\mathbf{b}\cdot\mathbf{c}.$$

This expression is known as the box product and is sometimes written as $[\mathbf{a},\mathbf{b},\mathbf{c}]$. You should consider what happens if you interchange the $\mathbf{b}$ with the $\mathbf{c}$ or the $\mathbf{a}$ with the $\mathbf{c}$. You can see geometrically from drawing pictures that this merely introduces a minus sign. In any case the box product of three vectors always equals either the volume of the parallelepiped determined by the three vectors or else minus this volume.

Example 3.4.11 *Find the volume of the parallelepiped determined by the vectors,* $\mathbf{i}+2\mathbf{j}-5\mathbf{k},\mathbf{i}+3\mathbf{j}-6\mathbf{k}$,$3\mathbf{i}+2\mathbf{j}+3\mathbf{k}$.

According to the above discussion, pick any two of these, take the cross product and then take the dot product of this with the third of these vectors. The result will be either the desired volume or minus the desired volume.

$$(\mathbf{i}+2\mathbf{j}-5\mathbf{k})\times(\mathbf{i}+3\mathbf{j}-6\mathbf{k})=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & 2 & -5\\ 1 & 3 & -6\end{vmatrix}=3\mathbf{i}+\mathbf{j}+\mathbf{k}$$

Now take the dot product of this vector with the third which yields

$$(3\mathbf{i}+\mathbf{j}+\mathbf{k})\cdot(3\mathbf{i}+2\mathbf{j}+3\mathbf{k})=9+2+3=14.$$

This shows the volume of this parallelepiped is 14 cubic units.

There is a fundamental observation which comes directly from the geometric definitions of the cross product and the dot product.

Lemma 3.4.12 *Let* $\mathbf{a},\mathbf{b}$, *and* $\mathbf{c}$ *be vectors. Then* $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$.

Proof: This follows from observing that either $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ and $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ both give the volume of the parallelepiped or they both give -1 times the volume. ■

Notation 3.4.13 *The box product* $\mathbf{a}\times\mathbf{b}\cdot\mathbf{c}=\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$ *is denoted more compactly as* $[\mathbf{a},\mathbf{b},\mathbf{c}]$.

3.4.3 A Proof Of The Distributive Law

Here is another proof of the distributive law for the cross product. Let $\mathbf{x}$ be a vector. From the above observation,

$$\begin{aligned}\mathbf{x}\cdot\mathbf{a}\times(\mathbf{b}+\mathbf{c}) &= (\mathbf{x}\times\mathbf{a})\cdot(\mathbf{b}+\mathbf{c})\\ &= (\mathbf{x}\times\mathbf{a})\cdot\mathbf{b}+(\mathbf{x}\times\mathbf{a})\cdot\mathbf{c}\\ &= \mathbf{x}\cdot\mathbf{a}\times\mathbf{b}+\mathbf{x}\cdot\mathbf{a}\times\mathbf{c}\\ &= \mathbf{x}\cdot(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}).\end{aligned}$$

Therefore,

$$\mathbf{x}\cdot[\mathbf{a}\times(\mathbf{b}+\mathbf{c})-(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c})]=0$$

for all $\mathbf{x}$. In particular, this holds for $\mathbf{x}=\mathbf{a}\times(\mathbf{b}+\mathbf{c})-(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c})$ showing that $\mathbf{a}\times(\mathbf{b}+\mathbf{c})=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$ and this proves the distributive law for the cross product another way.

Observation 3.4.14 *Suppose you have three vectors,* $\mathbf{u}=(a,b,c)$, $\mathbf{v}=(d,e,f)$, *and* $\mathbf{w}=(g,h,i)$. *Then* $\mathbf{u}\cdot\mathbf{v}\times\mathbf{w}$ *is given by the following.*

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v}\times\mathbf{w} &= (a,b,c)\cdot\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ d & e & f\\ g & h & i\end{vmatrix}\\ &= a\begin{vmatrix} e & f\\ h & i\end{vmatrix}-b\begin{vmatrix} d & f\\ g & i\end{vmatrix}+c\begin{vmatrix} d & e\\ g & h\end{vmatrix}\\ &\equiv \det\begin{pmatrix} a & b & c\\ d & e & f\\ g & h & i\end{pmatrix}.\end{aligned}$$

The message is that to take the box product, you can simply take the determinant of the matrix which results by letting the rows be the rectangular components of the given vectors in the order in which they occur in the box product. More will be presented on determinants later.

3.5 The Vector Identity Machine

In practice, you often have to deal with combinations of several cross products mixed in with dot products. It is extremely useful to have a technique which will allow you to discover vector identities and simplify expressions involving cross and dot products in three dimensions. This involves two special symbols, δ_{ij} and ε_{ijk} which are very useful in dealing with vector identities. To begin with, here is the definition of these symbols.

Definition 3.5.1 *The symbol δ_{ij}, called the Kronecker delta symbol is defined as follows.*

$$\delta_{ij} \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} .$$

With the Kronecker symbol i and j can equal any integer in $\{1, 2, \cdots, n\}$ for any $n \in \mathbb{N}$.

Definition 3.5.2 *For i, j, and k integers in the set, $\{1, 2, 3\}$, ε_{ijk} is defined as follows.*

$$\varepsilon_{ijk} \equiv \begin{cases} 1 \text{ if } (i, j, k) = (1, 2, 3), (2, 3, 1), \text{ or } (3, 1, 2) \\ -1 \text{ if } (i, j, k) = (2, 1, 3), (1, 3, 2), \text{ or } (3, 2, 1) \\ 0 \text{ if there are any repeated integers} \end{cases} .$$

The subscripts ijk and ij in the above are called indices. A single one is called an index. This symbol ε_{ijk} is also called the permutation symbol.

The way to think of ε_{ijk} is that $\varepsilon_{123} = 1$ and if you switch any two of the numbers in the list i, j, k, it changes the sign. Thus $\varepsilon_{ijk} = -\varepsilon_{jik}$ and $\varepsilon_{ijk} = -\varepsilon_{kji}$ etc. You should check that this rule reduces to the above definition. For example, it immediately implies that if there is a repeated index, the answer is zero. This follows because $\varepsilon_{iij} = -\varepsilon_{iij}$ and so $\varepsilon_{iij} = 0$.

It is useful to use the Einstein summation convention when dealing with these symbols. Simply stated, the convention is that you sum over the repeated index. Thus $a_i b_i$ means $\sum_i a_i b_i$. Also, $\delta_{ij} x_j$ means $\sum_j \delta_{ij} x_j = x_i$. Thus $\delta_{ij} x_j = x_i$, $\delta_{ii} = 3$, $\delta_{ij} x_{jkl} = x_{ikl}$. When you use this convention, there is one very important thing to never forget. It is this: **Never have an index be repeated more than once**. Thus $a_i b_i$ is all right but $a_{ii} b_i$ is not. The reason for this is that you end up getting confused about what is meant. If you want to write $\sum_i a_i b_i c_i$ it is best to simply use the summation notation. There is a very important reduction identity connecting these two symbols.

Lemma 3.5.3 *The following holds.*

$$\varepsilon_{ijk} \varepsilon_{irs} = (\delta_{jr} \delta_{ks} - \delta_{kr} \delta_{js}) .$$

Proof: If $\{j, k\} \neq \{r, s\}$ then every term in the sum on the left must have either ε_{ijk} or ε_{irs} contains a repeated index. Therefore, the left side equals zero. The right side also equals zero in this case. To see this, note that if the two sets are not equal, then there is one of the indices in one of the sets which is not in the other set. For example, it could be that j is not equal to either r or s. Then the right side equals zero.

Therefore, it can be assumed $\{j, k\} = \{r, s\}$. If $i = r$ and $j = s$ for $s \neq r$, then there is exactly one term in the sum on the left and it equals 1. The right also reduces to 1 in this case. If $i = s$ and $j = r$, there is exactly one term in the sum on the left which is nonzero and it must equal -1. The right side also reduces to -1 in this case. If there is a repeated index in $\{j, k\}$, then every term in the sum on the left equals zero. The right also reduces to zero in this case because then $j = k = r = s$ and so the right side becomes $(1)(1) - (-1)(-1) = 0$. ■

Proposition 3.5.4 *Let $\mathbf{u}, \mathbf{v}$ be vectors in $\mathbb{R}^n$ where the Cartesian coordinates of $\mathbf{u}$ are $(u_1, \cdots, u_n)$ and the Cartesian coordinates of $\mathbf{v}$ are $(v_1, \cdots, v_n)$. Then $\mathbf{u} \cdot \mathbf{v} = u_i v_i$. If $\mathbf{u}, \mathbf{v}$ are vectors in $\mathbb{R}^3$, then*

$$(\mathbf{u} \times \mathbf{v})_i = \varepsilon_{ijk} u_j v_k .$$

Also, $\delta_{ik} a_k = a_i$.

Proof: The first claim is obvious from the definition of the dot product. The second is verified by simply checking that it works. For example,

$$\mathbf{u} \times \mathbf{v} \equiv \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

and so

$$(\mathbf{u} \times \mathbf{v})_1 = (u_2 v_3 - u_3 v_2).$$

From the above formula in the proposition,

$$\varepsilon_{1jk} u_j v_k \equiv u_2 v_3 - u_3 v_2,$$

the same thing. The cases for $(\mathbf{u} \times \mathbf{v})_2$ and $(\mathbf{u} \times \mathbf{v})_3$ are verified similarly. The last claim follows directly from the definition. ■

With this notation, you can easily discover vector identities and simplify expressions which involve the cross product.

Example 3.5.5 *Discover a formula which simplifies* $(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{z} \times \mathbf{w})$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$.

From the above description of the cross product and dot product, along with the reduction identity,

$$(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{z} \times \mathbf{w}) =$$

$$\begin{aligned} \varepsilon_{ijk} u_j v_k \varepsilon_{irs} z_r w_s &= (\delta_{jr} \delta_{ks} - \delta_{js} \delta_{kr}) u_j v_k z_r w_s \\ &= u_j v_k z_j w_k - u_j v_k z_k w_j \\ &= (\mathbf{u} \cdot \mathbf{z})(\mathbf{v} \cdot \mathbf{w}) - (\mathbf{u} \cdot \mathbf{w})(\mathbf{v} \cdot \mathbf{z}) \end{aligned}$$

Example 3.5.6 *Simplify* $\mathbf{u} \times (\mathbf{u} \times \mathbf{v})$.

The i^{th} component is

$$\begin{aligned} \varepsilon_{ijk} u_j (\mathbf{u} \times \mathbf{v})_k &= \varepsilon_{ijk} u_j \varepsilon_{krs} u_r v_s = \varepsilon_{kij} \varepsilon_{krs} u_j u_r v_s \\ &= (\delta_{ir} \delta_{js} - \delta_{jr} \delta_{is}) u_j u_r v_s \\ &= u_j u_i v_j - u_j u_j v_i \\ &= (\mathbf{u} \cdot \mathbf{v}) u_i - |\mathbf{u}|^2 v_i \end{aligned}$$

Hence

$$\mathbf{u} \times (\mathbf{u} \times \mathbf{v}) = (\mathbf{u} \cdot \mathbf{v}) \mathbf{u} - |\mathbf{u}|^2 \mathbf{v}$$

because the i^{th} components of the two sides are equal for any i.

3.6 Exercises

1. Show that if $\mathbf{a} \times \mathbf{u} = \mathbf{0}$ for all unit vectors, $\mathbf{u}$, then $\mathbf{a} = \mathbf{0}$.
2. Find the area of the triangle determined by the three points, $(1, 2, 3)$, $(4, 2, 0)$ and $(-3, 2, 1)$.
3. Find the area of the triangle determined by the three points, $(1, 0, 3)$, $(4, 1, 0)$ and $(-3, 1, 1)$.
4. Find the area of the triangle determined by the three points, $(1, 2, 3)$, $(2, 3, 4)$ and $(3, 4, 5)$. Did something interesting happen here? What does it mean geometrically?
5. Find the area of the parallelogram determined by the vectors, $(1, 2, 3)$, $(3, -2, 1)$.
6. Find the area of the parallelogram determined by the vectors, $(1, 0, 3)$, $(4, -2, 1)$.

7. Find the volume of the parallelepiped determined by the vectors, $\mathbf{i}-7\mathbf{j}-5\mathbf{k}, \mathbf{i}-2\mathbf{j}-6\mathbf{k}, 3\mathbf{i}+2\mathbf{j}+3\mathbf{k}$.

8. Suppose $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ are three vectors whose components are all integers. Can you conclude the volume of the parallelepiped determined from these three vectors will always be an integer?

9. What does it mean geometrically if the box product of three vectors gives zero?

10. Using Problem 9, find an equation of a plane containing the two position vectors, $\mathbf{a}$ and $\mathbf{b}$ and the point $\mathbf{0}$. **Hint:** If (x, y, z) is a point on this plane the volume of the parallelepiped determined by (x, y, z) and the vectors $\mathbf{a}, \mathbf{b}$ equals 0.

11. Using the notion of the box product yielding either plus or minus the volume of the parallelepiped determined by the given three vectors, show that

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

In other words, the dot and the cross can be switched as long as the order of the vectors remains the same. **Hint:** There are two ways to do this, by the coordinate description of the dot and cross product and by geometric reasoning. It is better if you use geometric reasoning.

12. Is $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$? What is the meaning of $\mathbf{a} \times \mathbf{b} \times \mathbf{c}$? Explain. **Hint:** Try $(\mathbf{i} \times \mathbf{j}) \times \mathbf{j}$.

13. Discover a vector identity for $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$ and one for $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$.

14. Discover a vector identity for $(\mathbf{u} \times \mathbf{v}) \times (\mathbf{z} \times \mathbf{w})$.

15. Simplify $(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{v} \times \mathbf{w}) \times (\mathbf{w} \times \mathbf{z})$.

16. Simplify $|\mathbf{u} \times \mathbf{v}|^2 + (\mathbf{u} \cdot \mathbf{v})^2 - |\mathbf{u}|^2 |\mathbf{v}|^2$.

17. For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ functions of t, $\mathbf{u}'(t)$ is defined as the limit of the difference quotient as in calculus, $(\lim_{h \to 0} \mathbf{w}(h))_i \equiv \lim_{h \to 0} w_i(h)$. Show the following

$$(\mathbf{u} \times \mathbf{v})' = \mathbf{u}' \times \mathbf{v} + \mathbf{u} \times \mathbf{v}', \ (\mathbf{u} \cdot \mathbf{v})' = \mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}'$$

18. If $\mathbf{u}$ is a function of t, and the magnitude $|\mathbf{u}(t)|$ is a constant, show from the above problem that the velocity $\mathbf{u}'$ is perpendicular to $\mathbf{u}$.

19. When you have a rotating rigid body with angular velocity vector $\boldsymbol{\Omega}$, then the velocity vector $\mathbf{v} \equiv \mathbf{u}'$ is given by

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{u}$$

where $\mathbf{u}$ is a position vector. The acceleration is the derivative of the velocity. Show that if $\boldsymbol{\Omega}$ is a constant vector, then the acceleration vector $\mathbf{a} = \mathbf{v}'$ is given by the formula

$$\mathbf{a} = \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{u}).$$

Now simplify the expression. It turns out this is centripetal acceleration.

20. Verify directly that the coordinate description of the cross product, $\mathbf{a} \times \mathbf{b}$ has the property that it is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. Then show by direct computation that this coordinate description satisfies

$$\begin{aligned} |\mathbf{a} \times \mathbf{b}|^2 &= |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 \\ &= |\mathbf{a}|^2 |\mathbf{b}|^2 \left(1 - \cos^2(\theta)\right) \end{aligned}$$

where θ is the angle included between the two vectors. Explain why $|\mathbf{a} \times \mathbf{b}|$ has the correct magnitude. All that is missing is the material about the right hand rule. Verify directly from the coordinate description of the cross product that the right thing happens with regards to the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$. Next verify that the distributive law holds for the coordinate description of the cross product. This gives another way to approach the cross product. First define it in terms of coordinates and then get the geometric properties from this. However, this approach does not yield the right hand rule property very easily.