

Math S2 PeiP

Chapitre 5

Applications linéaires et calcul matriciel

Michel Rumin

1 Matrices et applications linéaires

1.1 Applications linéaires

Soient n et p deux entiers non nuls. Nous allons définir les applications $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ qui nous intéressent.

Définition vectorielle. Il y a deux présentations utiles et complémentaires de la notion d'application linéaire. La première est abstraite et géométrique car elle ne fait pas appel explicitement aux composantes des vecteurs.

Définition 1.1. Une application $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est dite **linéaire** si pour tout $\vec{u}, \vec{v} \in \mathbb{R}^p$, et tout $\lambda \in \mathbb{R}$, on a :

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

On peut aussi regrouper les deux propriétés en disant que :

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

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

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

Plus précisément, l'image par f de la droite passant D par \vec{u} et de direction \vec{v} est la droite $f(D)$ passant par $f(\vec{u})$ et de direction $f(\vec{v})$. Notez que si $f(\vec{v}) = \vec{0}$ alors l'image $f(D)$ se réduit au point $f(\vec{u})$.

Application linéaire et image d'une base. Voici une propriété cruciale des applications linéaires, qui permet de les fabriquer et de les reconnaître.

Théorème 1.2. Soient $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ une **base** de \mathbb{R}^p , et $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ des vecteurs **quelconques** de \mathbb{R}^n . Alors il existe une **unique** application linéaire $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ telle que

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

Si $\vec{u} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \dots + \lambda_p \vec{u}_p$ est un vecteur quelconque de \mathbb{R}^p , de coordonnées $(\lambda_1, \lambda_2, \dots, \lambda_p)$ dans la base B , on a :

$$f(\vec{u}) = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_p \vec{v}_p. \quad (1)$$

En d'autres termes, une application linéaire est déterminée par son action sur une base.

Démonstration. La formule vient directement de la linéarité. Si une telle f existe, on doit avoir :

$$\begin{aligned} f(\vec{u}) &= f(\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \dots + \lambda_p \vec{u}_p) \\ &= \lambda_1 f(\vec{u}_1) + \lambda_2 f(\vec{u}_2) + \dots + \lambda_p f(\vec{u}_p) \\ &= \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_p \vec{v}_p. \end{aligned}$$

Le fait que cette formule définit bien une application linéaire provient de la « linéarité » des coordonnées des vecteurs par somme et homothétie. \square

Ainsi, une application linéaire est une combinaison linéaire de vecteurs fixes **déterminés** par l'image d'une base de l'ensemble de départ.

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

est la diagonale du parallélogramme de côtés $xf(\vec{e}_1)$ et $yf(\vec{e}_2)$.

Expression en coordonnées. L'équation (1) permet d'avoir une deuxième définition, plus concrète, d'une application linéaire $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$. Si on travaille avec les bases canoniques B_p et B_n de \mathbb{R}^p et \mathbb{R}^n , alors les coordonnées de

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

sont ses composantes x_j . On peut aussi décomposer chaque vecteur $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ de l'espace d'arrivée \mathbb{R}^n :

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

La formule (1) se traduit alors de la façon suivante :

Définition 1.3. Une application linéaire f de \mathbb{R}^p dans \mathbb{R}^n est une expression qui transforme un vecteur $\vec{u} = (x_1, x_2, \dots, x_p)$ de \mathbb{R}^p en un vecteur $\vec{v} = f(\vec{u}) = (y_1, y_2, \dots, y_n)$ de \mathbb{R}^n , où chaque composante y_i est donnée par une combinaison linéaire des coordonnées x_j de x . C'est-à-dire qu'il existe des constantes a_{ij} telles que

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

Comme dans le cas des systèmes linéaires, le passage de la combinaison linéaire (1) à l'écriture (2) se fait en mettant **en colonnes** les coordonnées des vecteurs \vec{v}_j image de la base.

Par exemple, l'application $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ telle que $f(x, y, z) = (12x - y, 2x + y + z)$ est une telle expression, par contre $f(x, y, z) = (x^2 + y^2, z)$ ou $f(x, y, z) = (x + 1, y)$ n'en sont pas.

1.2 Matrices et applications linéaires

On voit qu'une application linéaire de \mathbb{R}^p dans \mathbb{R}^n est uniquement déterminée par un tableau A à n lignes et p colonnes de coefficients

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

où a_{ij} est l'élément de A situé sur la i -ème ligne et la j -ème colonne de A .¹

Définition 1.4. A s'appelle la **matrice** de l'application linéaire f dans les bases canoniques, et on écrit $A = \text{Mat}(f)$.

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

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

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

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

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

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

1. C'est la même convention d'indice que pour les systèmes linéaires !

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

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

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

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

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

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

$$\begin{aligned} f : \vec{u} \mapsto \vec{v} = f(\vec{u}) &= x_1 f(\vec{u}_1) + x_2 f(\vec{u}_2) + \cdots + x_p f(\vec{u}_p) \\ &\quad \updownarrow \text{passage en colonne} \\ A : X \mapsto Y = AX &= x_1 C_1 + x_2 C_2 + \cdots + x_p C_p, \end{aligned}$$

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

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

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

Pour calculer $f(1, -1, 2)$ on écrit

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

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

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

Par exemple le produit suivant

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

Il ne faut pas compléter les « composantes fantômes » par des 0 s'il en manque. Ce calcul n'a pas de sens² car A est associée à une application de \mathbb{R}^3 dans \mathbb{R}^2 mais $(1, -1)$ est un

2. d'ailleurs, Scilab refusera de le faire!

vecteur de \mathbb{R}^2 .

1.3 Exemples d'applications linéaires du plan : les similitudes

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

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

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

par la formule

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto Y = AX = \begin{pmatrix} ax + cy \\ bx + dy \end{pmatrix}.$$

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

Définition 1.5. Soit $\lambda = a + ib$ un nombre complexe. L'application

$$f_\lambda : \mathbb{C} \rightarrow \mathbb{C}$$

$$z = x + iy \mapsto \lambda z$$

s'appelle une *similitude directe* du plan.

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

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

Autrement dit, la similitude f_λ est l'application linéaire associée à la matrice

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

Voici leur interprétation géométrique.

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

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

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

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

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

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

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

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

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

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

En effet, on a

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

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

On a ici,

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

La matrice de R_θ est donc

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

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

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

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

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

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

est associée à l'application linéaire $p : (x, y) \mapsto (x, 0)$ qui est la **projection orthogonale**

sur l'axe des x .

Nous allons voir comment définir les projections en général.

1.4 Les projections

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

Définition générale. Si on a

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

alors tout vecteur \vec{u} de \mathbb{R}^n se décompose alors de manière unique en

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

Définition 1.6. La projection $p_{E/F}$ de \mathbb{R}^n sur E parallèlement à F (ou le long de F) est l'application définie par

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

On vérifie facilement (exercice) que $p_{E/F}$ est une application linéaire.

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

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

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

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

$$p_{D_1/D_2}(\vec{v}) = \frac{\ell(\vec{v})}{\ell(\vec{u}_1)} \vec{u}_1. \quad (4)$$

En effet, la formule proposée est clairement une application linéaire. De plus, on a

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

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

Projection orthogonale. Si de plus D_2 est la droite orthogonale à D_1 , alors on peut prendre $\ell(\vec{v}) = \langle \vec{u}_1, \vec{v} \rangle$, car alors $\ell(\vec{v}) = \langle \vec{u}_1, \vec{v} \rangle = 0$ est bien l'équation définissant les vecteurs de D_2 , orthogonaux à \vec{v}_1 . Dans ce cas la formule (4) s'écrit

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

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

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

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

Si D est engendré par \vec{u}_1 et P a pour équation $\ell(x, y, z) = ax + by + cz = 0$, alors la formule (4)

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

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

La projection $p_{P/D}$ sur P le long de D s'obtient par

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

En effet, d'après la définition 1.6, on a de manière générale

$$p_{E/F} + p_{F/E} = \text{Id} \quad \text{si} \quad \mathbb{R}^n = E \oplus F.$$

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

En algèbre linéaire, il est souvent utile de travailler dans des bases adaptées au problème considéré, et qui ne sont pas nécessairement les bases canoniques de \mathbb{R}^p et \mathbb{R}^n . Par exemple dans le cas d'une projection sur E par rapport à F , on préfère travailler dans une base du type $B = B_E \cup B_F$ car l'action de $p_{E/F}$ est simple sur B_E et B_F .

On revient sur la relation entre applications linéaires et matrices dans des bases quelconques.

1.5 Matrices et applications linéaires dans des bases générales

Soient $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ et $B' = (\vec{u}'_1, \vec{u}'_2, \dots, \vec{u}'_n)$ des bases **quelconques** de \mathbb{R}^p et \mathbb{R}^n , et soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application linéaire.

D'après le théorème 1.2, on a toujours pour

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

la formule

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

On peut comme dans la définition 1.3 écrire cette relation en coordonnées dans la base B' , ce qui donne en colonne une expression similaire

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

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_p \end{pmatrix} \longmapsto Y = \begin{pmatrix} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ \vdots \\ y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p \\ \vdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p \end{pmatrix} \quad (5)$$

c'est-à-dire

$$Y = AX$$

avec

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

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

Définition 1.7. On appelle cette matrice A la matrice de l'application linéaire f dans les bases B et B' , et on la note

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

Les formules de calcul sont les mêmes que dans les bases canoniques, mais on manipule ici les coordonnées de \vec{u} dans la base B de départ, et on obtient des coordonnées de $f(\vec{u})$ dans la base B' d'arrivée. Lorsque l'on utilise les bases canoniques, on allège la notation en écrivant $A = \text{Mat}(f)$.

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

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

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

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

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

puisque

$$f(\vec{e}_1) = (1, 2) = \vec{u}_1 = (1, 0)_{B'_2} \Rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ première colonne de } A',$$

$$f(\vec{e}_2) = (3, 4) = \vec{u}_2 = (0, 1)_{B'_2} \Rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ deuxième colonne de } A',$$

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

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

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

On a $f(x, y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right) = \frac{x+y}{2}(1, 1)$. En particulier, on a

$$f(1, 1) = (1, 1) \text{ tandis que } f(-1, 1) = (0, 0).$$

Dans la base $B' = (\vec{u}_1 = (1, 1), \vec{u}_2 = (-1, 1))$ de \mathbb{R}^2 , on a donc

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

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

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

L'application f est en fait la projection (orthogonale) sur la droite $y = x$, le long de la droite $y = -x$.

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

En fait, **toutes** les projections sur des droites dans \mathbb{R}^2 ont une matrice de cette forme, dans une base convenable! En effet, celles-ci sont définies par deux droites distinctes : D_1 , la droite sur laquelle on projette, et D_2 la direction de projection (pas toujours orthogonale, penser à l'ombre du soleil). Si on prend une base $B' = (\vec{u}_1, \vec{u}_2)$ de \mathbb{R}^2 avec $\vec{u}_1 \in D_1$ et $\vec{u}_2 \in D_2$, alors la formule de projection est d'après la définition 1.6

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

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

2 Opérations sur les matrices et les applications linéaires

2.1 Structure d'espace vectoriel

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

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

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

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

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

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

Ces deux opérations sont compatibles dans toute base de travail de \mathbb{R}^p et \mathbb{R}^n .

Proposition 2.1. Si $A = \text{Mat}(f)$ et $B = \text{Mat}(g)$, alors $A + B = \text{Mat}(f + g)$.

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

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

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

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

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

Ces deux opérations sont compatibles.

Proposition 2.2. Si $A = \text{Mat}(f)$ alors $\lambda A = \text{Mat}(\lambda f)$.

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

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

sont des espaces vectoriels de dimension $n \times p$. Les $n \times p$ coefficients a_{ij} d'une matrice $A = (a_{ij}) \in M_{n,p}(\mathbb{R})$ donnent des coordonnées.

2.2 Produit de matrices et composition d'applications linéaires

Composition d'applications linéaires. Quand on a deux applications

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

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

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

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

Proposition 2.3. *Si f et g sont des applications linéaires composables, alors $f \circ g$ est aussi une application **linéaire**.*

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

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

□

Produit de matrices. L'opération précédente de composition d'applications linéaires est très simple et naturelle, et définie indépendamment de tout choix de base. Comme on sait que l'on peut associer des matrices à des applications linéaires, cela signifie que l'on va pouvoir **multiplier** certaines matrices entre elles.

Soient $f \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^n)$ et $g \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$. Soient B une base de \mathbb{R}^p , B' une base de \mathbb{R}^q et B'' une base de \mathbb{R}^n . Alors on peut considérer les matrices

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

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

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

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

où B_j est la j -ème colonne de B .

Pour calculer AB , il suffit donc de multiplier une à une les colonnes de B par la matrice A . On présente les calculs de la façon suivante :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & & \vdots \\ \mathbf{a}_{i1} & \rightarrow & \mathbf{a}_{ik} & \rightarrow & \mathbf{a}_{iq} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nq} \end{pmatrix} \begin{pmatrix} \cdots & & \cdots \\ \vdots & & \vdots \\ \downarrow & & \\ \mathbf{c}_{ij} & & \end{pmatrix} = C = AB$$

avec

$$c_{ij} = \sum_{k=1}^q a_{ik} b_{kj}. \quad (6)$$

Exemple. Soient $A = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Calculer le produit AB . On a

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$A = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 3 & 3 \\ 4 & 5 & 6 \end{pmatrix} = AB$$

On a défini le produit de matrice pour qu'il soit compatible avec l'opération de composition naturelle des applications linéaires. C'est sa propriété principale.

Théorème 2.4. Soient $f \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^n)$ et $g \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$. Alors on a

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

Propriétés du produit et de la composition.

Lorsqu'elle est définie, la composition est :

1. associative : $(f \circ g) \circ h = f \circ (g \circ h)$,
2. distributive par rapport à l'addition : $(f+g) \circ h = f \circ h + g \circ h$ et $f \circ (g+h) = f \circ g + f \circ h$.

On a donc aussi au niveau matriciel :

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

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

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

Définition 2.5. - Une application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ s'appelle un **endomorphisme**. On note $\text{End}(\mathbb{R}^n) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ cet espace.

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

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

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

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

a pour matrice (dans toute base de \mathbb{R}^n) la **matrice identité** de taille n ,

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

On a

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

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

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

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

$$\begin{aligned} AB &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = A \\ A^2 &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \text{ avec } A \neq 0! \end{aligned}$$

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

3 Noyau et image des application linéaires

3.1 Définitions et premières propriétés

Soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application linéaire. On peut lui associer deux sous-espaces vectoriels importants.

Définition 3.1. - L'**image** de f est l'ensemble des vecteurs de \mathbb{R}^n qui sont de la forme $f(\vec{u})$ avec $\vec{u} \in \mathbb{R}^p$:

$$\text{Im}(f) = \{\vec{v} = f(\vec{u}), \vec{u} \in \mathbb{R}^p\} = f(\mathbb{R}^p).$$

- Le **noyau** de f est l'ensemble des vecteurs $\vec{u} \in \mathbb{R}^p$ tels que $f(\vec{u}) = \vec{0}$:

$$\ker f = \{\vec{u} \in \mathbb{R}^p, f(\vec{u}) = \vec{0}\}.$$

Ce sont des sous-espaces vectoriels qui déterminent si f est surjective, injective.

Proposition 3.2. a) $\ker f$ est un sev de \mathbb{R}^p et $\text{Im } f$ est un sev de \mathbb{R}^n ,

b) f est surjective $\Leftrightarrow \text{Im } f = \mathbb{R}^n$,

c) f est injective $\Leftrightarrow \ker f = \{\vec{0}\}$.

Démonstration. a) • Soient \vec{u} et $\vec{v} \in \ker f$, alors par linéarité de f , on a

$$f(\vec{u} + \lambda\vec{v}) = f(\vec{u}) + \lambda f(\vec{v}) = \vec{0} + \lambda\vec{0} = \vec{0},$$

d'où $\vec{u} + \lambda\vec{v} \in \ker f$ et $\ker f$ est un sous-espace vectoriel de \mathbb{R}^p .

• Soient \vec{v}_1 et $\vec{v}_2 \in \text{Im } f$, alors il existe \vec{u}_1 et $\vec{u}_2 \in \mathbb{R}^p$ tels que $f(\vec{u}_1) = \vec{v}_1$ et $f(\vec{u}_2) = \vec{v}_2$.
On a alors

$$\vec{v}_1 + \lambda\vec{v}_2 = f(\vec{u}_1) + \lambda f(\vec{u}_2) = f(\vec{u}_1 + \lambda\vec{u}_2) \in \text{Im } f,$$

et $\text{Im } f$ est un sev de \mathbb{R}^n .

b) Par définition, f est surjective ssi tout vecteur \vec{v} de l'espace d'arrivée \mathbb{R}^n possède au moins un antécédent par f , c'est-à-dire s'écrit $\vec{v} = f(\vec{u})$. Cela signifie que $\text{Im } f = \mathbb{R}^n$.

c) On rappelle que f est injective si tout élément $\vec{v} \in \mathbb{R}^n$ possède **au plus un** antécédent. Cela signifie que

$$f(\vec{u}_1) = f(\vec{u}_2) \iff \vec{u}_1 = \vec{u}_2.$$

Si f est injective et $\vec{u} \in \ker f$ alors

$$f(\vec{u}) = \vec{0} = f(\vec{0}) \Rightarrow \vec{u} = \vec{0} \Rightarrow \ker f = \{\vec{0}\}.$$

Inversement, si $\ker f = \{\vec{0}\}$ et $f(\vec{u}_1) = f(\vec{u}_2)$ alors par linéarité de f , on a

$$f(\vec{u}_1) - f(\vec{u}_2) = f(\vec{u}_1 - \vec{u}_2) = \vec{0} \iff \vec{u}_1 - \vec{u}_2 \in \ker f \iff \vec{u}_1 - \vec{u}_2 = \vec{0},$$

et f est injective. □

La propriété c) est très efficace pour étudier l'injectivité d'une **application linéaire**. En effet $\vec{u} \in \ker f \Leftrightarrow f(\vec{u}) = \vec{0}$ est un système linéaire homogène à p inconnues (les coordonnées de $\vec{u} \in \mathbb{R}^p$), alors que la définition générale de l'injectivité dépend de l'équation $f(\vec{u}_1) = f(\vec{u}_2)$, qui est un système à $2p$ inconnues (deux fois plus)! C'est la linéarité de f qui permet cette simplification.

La propriété : f injective $\Leftrightarrow \ker f = \{\vec{0}\}$ s'interprète géométriquement. Deux vecteurs \vec{u}_1 et \vec{u}_2 ont même image par f si et seulement si $\vec{u}_2 - \vec{u}_1 \in \ker f$, c'est-à-dire ssi \vec{u}_2 appartient à l'espace affine $E_{\vec{u}_1}$ passant par \vec{u}_1 et de direction vectorielle $\ker f$. On a f injective si et seulement si $E_{\vec{u}_1}$ est réduit à un point, c'est-à-dire ssi $\ker f = \{\vec{0}\}$. Lorsque f n'est pas injective, elle « écrase » les directions dans $\ker f$, comme le fait une projection.

En résumé, l'étude de l'équation $f(\vec{u}) = \vec{v}$ s'appuie sur le noyau et l'image de f :

Proposition 3.3. Soient $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ et $\vec{v} \in \mathbb{R}^n$ donnés. Alors :

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

Le noyau d'une application f est défini par une équation homogène $f(\vec{u}) = \vec{0}$, qui conduit en coordonnées au système $AX = 0$ avec $A = \text{Mat}(f)$. Ainsi le noyau est présenté comme l'espace des solutions d'un système linéaire homogène.

Ce n'est pas le cas de l'image, qui est définie comme l'ensemble des seconds membres $\vec{v} \in \mathbb{R}^n$ tels que l'équation $f(\vec{u}) = \vec{v}$ possède au moins une solution. On peut donner une présentation beaucoup plus pratique lorsque f est linéaire.

Image de f et image d'une base.

Proposition 3.4. Soit $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ et $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ une base quelconque de \mathbb{R}^p . Alors l'image de f est engendrée par l'image de B :

$$\text{Im } f = \text{Vect}(f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_p)).$$

Démonstration. Tout \vec{u} de \mathbb{R}^p s'écrit $\vec{u} = x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_p\vec{u}_p$. On a alors par linéarité de f ,

$$f(\vec{u}) = x_1f(\vec{u}_1) + x_2f(\vec{u}_2) + \dots + x_pf(\vec{u}_p).$$

Ainsi, les vecteurs \vec{v} de $\text{Im } f$ sont les combinaisons linéaires de $f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_p)$. \square

Noyau et image des matrices. Soit $A \in M_{n,p}(\mathbb{R})$ une matrice. On lui associe une application linéaire

$$\begin{aligned} f : \mathbb{R}^p &\rightarrow \mathbb{R}^n \\ X &\mapsto Y = AX. \end{aligned}$$

Définition 3.5. - Le noyau de A est le noyau de f c'est-à-dire

$$\ker A = \{X \in \mathbb{R}^p, AX = 0\}.$$

C'est l'ensemble des solutions d'un système homogène à p équations et n inconnues.

- L'image de A est l'image de f , c'est-à-dire les vecteurs Y de \mathbb{R}^n de la forme $Y = AX$.

D'après la proposition 3.4, $\text{Im } A$ est le sev de \mathbb{R}^n engendré par les p vecteurs colonnes de A , puisque ces colonnes sont l'image de la base canonique de \mathbb{R}^p .

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

• On a $X = (x, y, z) \in \ker f = \ker A \Leftrightarrow$

$$\begin{aligned} AX &= \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} \boxed{x} + 4y + 7z = 0 \\ 2x + 5y + 8z = 0 \\ 3x + 6y + 9z = 0 \end{cases} \\ \Leftrightarrow \begin{cases} \boxed{x} + 4y + 7z = 0 \\ \boxed{-3y} - 6z = 0 \\ -6y - 12z = 0 \end{cases} &\Leftrightarrow \begin{cases} \boxed{x} + 4y + 7z = 0 \\ \boxed{y} + 2z = 0 \end{cases} \Leftrightarrow y = -2z, x = z, z \in \mathbb{R}. \end{aligned}$$

On a donc que $\ker A$ est la droite vectorielle engendrée par le vecteur $X = (1, -2, 1)$. On note que $X = (1, -2, 1) \in \ker A$ se traduit par une relation de dépendance linéaire entre les colonnes de A . On a

$$C_1 - 2C_2 + C_3 = 0. \tag{7}$$

C'est une propriété générale utile.

Proposition 3.6. Soit $A \in M_{n,p}(\mathbb{R})$. On a

$$X = (x_1, x_2, \dots, x_p) \in \ker A \Leftrightarrow AX = x_1 C_1 + x_2 C_2 + \dots + x_p C_p = 0,$$

avec les C_j vecteurs colonnes de A .

En effet, on a $C_j = Ae_j$ avec e_j base canonique.

• D'après la proposition 3.4, l'image de A est engendrée par ses colonnes : C_1, C_2, C_3 . D'après (7) elles sont liées, et en fait $\text{Im } A$ est un plan vectoriel de base C_1 et C_2 par exemple.

3.2 Le théorème du rang et ses applications

Le rang d'une application.

Définition 3.7. Soit f une application linéaire. On appelle rang de f la dimension de $\text{Im } f$, et on le note $\text{rang}(f)$.

D'après la proposition 3.4, le rang de $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ est le rang de l'image d'une base $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ de \mathbb{R}^p :

$$\text{rang}(f) = \text{rang}(f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_p)) = \text{rang}(C_1, C_2, \dots, C_p),$$

pour les colonnes C_j de $A = \text{Mat}_{B,B'}(f)$.

Par extension, le rang d'une matrice A est le rang de ces vecteurs colonnes, c'est-à-dire le rang de l'application $X \mapsto AX$ associée. On peut donc calculer le rang d'une matrice A en l'échelonnant à l'aide de la méthode du pivot. Le rang de A est finalement le nombre de pivots d'une matrice échelonnée équivalente.

Exemple. Pour λ paramètre réel, on calcule le rang de

$$A_\lambda = \begin{pmatrix} 1 & 0 & 1 & 2 \\ \lambda & \lambda - 1 & 3\lambda & 3\lambda \\ 1 & 0 & \lambda + 1 & 2 \\ 1 & 0 & \lambda + 1 & \lambda + 2 \end{pmatrix}.$$

On a

$$\text{rang}(A_\lambda) = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & \lambda - 1 & 2\lambda & \lambda \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & \lambda & \lambda \end{pmatrix} = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & \lambda - 1 & 2\lambda & \lambda \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

• Cette matrice triangulaire est échelonnée de rang 4 si tous les nombres sur la diagonale sont non nuls (sont des pivots). On a donc $\text{rang} A_\lambda = 4$ si $\lambda \neq 0$ et 1. Dans ce cas $\text{Im } A_\lambda = \mathbb{R}^4$ et A_λ est surjective.

• Si $\lambda = 0$, alors

$$\text{rang}(A_0) = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & \boxed{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 2$$

et une base de $\text{Im } A_0$ est constituée de deux première colonnes de A_0 : $C_1 = (1, 0, 1, 1)$ et $C_2 = (0, -1, 0, 0)$.

• Si $\lambda = 1$, alors

$$\begin{aligned} \text{rang}(A_1) &= \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & 0 & \boxed{2} & 1 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \text{rang} \begin{pmatrix} \boxed{1} & 0 & 1 & 2 \\ 0 & 0 & \boxed{2} & 1 \\ 0 & 0 & 0 & \boxed{-\frac{1}{2}} \\ 0 & 0 & 0 & 0 \end{pmatrix} = 3, \end{aligned}$$

et une base de $\text{Im } A_1$ est donnée par les colonnes C_1, C_3 et C_4 de A_1 .

Le théorème du rang. Il y a une relation importante et générale entre les dimensions du noyau et de l'image d'une application linéaire.

Théorème 3.8 (Théorème du rang). Soit $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ linéaire, alors on a :

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

Démonstration. Soit $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ une base de \mathbb{R}^p . Le noyau de f est donné par l'équation

$$(S) : \quad f(\vec{u}) = x_1 f(\vec{u}_1) + x_2 f(\vec{u}_2) + \dots + x_p f(\vec{u}_p) = \vec{0}$$

qui est un système linéaire homogène à p inconnues et n équations. On a donc :

$$\dim \ker f = \text{nombre d'inconnues non principales de } (S).$$

Par ailleurs, on a

$$\begin{aligned} \text{rang}(f) &= \text{rang}(f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_p)) \\ &= \text{nombre d'inconnues principales de } (S). \end{aligned}$$

On a bien

$$p = \text{nombre d'inconnues de } (S) = \text{rang}(f) + \dim \ker f$$

□

Attention, si $p = n$, le théorème du rang peut donner l'impression que noyau et image sont en somme directe dans \mathbb{R}^n , car il satisfait la bonne condition sur les dimensions. Mais cela n'est pas vrai en général car il arrive que

$$\ker f \cap \text{Im } f \neq \{\vec{0}\}.$$

Par exemple pour $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ associée à $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, on a $\ker f = \text{Im}(f) = \text{axe des } x$.

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

Proposition 3.9. Soit $f \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$. Alors :

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

Démonstration.

- a) On a toujours $\text{rang}(f) = p - \dim \ker f \leq p$, avec égalité ssi $\ker f = \{\vec{0}\}$.
- b) On a toujours $\text{Im } f \subset \mathbb{R}^n$ et donc, si f est injective, $p = \text{rang}(f) = \dim \text{Im } f \leq n$.
- c) Si f est surjective alors $\text{rang}(f) = \dim \text{Im } f = n \leq p$ d'après a).
- d) Synthèse de b) et c).
- e) Si $p = n$ et $\ker f = \{\vec{0}\}$, alors $\text{rang}(f) = p = n$ d'après a), et donc $\text{Im } f = \mathbb{R}^n$ et f est aussi surjective. Si $p = n$ et $\text{rang}(f) = n$, alors $\dim \ker f = p - \text{rang}(f) = 0$ et f est injective.

□

4 Isomorphismes et matrices inversibles

4.1 Les isomorphismes

Une application linéaire bijective s'appelle un **isomorphisme**.

On sait d'après la proposition précédente que si $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est un isomorphisme, alors $p = n$.

Si f est un isomorphisme, on note $f^{-1}(\vec{v})$ l'unique solution de l'équation $f(\vec{u}) = \vec{v}$. Cette application $f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ s'appelle **l'application réciproque** de f . On a par définition

$$f \circ f^{-1} = \text{Id} = f^{-1} \circ f.$$

Proposition 4.1. *Si f est un isomorphisme, alors son application réciproque f^{-1} est aussi une application linéaire.*

Démonstration. Soient $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$. On a

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

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

4.2 Matrices inversibles

Définition 4.2. Une matrice carré $A \in M_n(\mathbb{R})$ est dite inversible si l'application linéaire associée $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $X \mapsto AX$ est un isomorphisme, c'est-à-dire si pour tout $Y \in \mathbb{R}^n$, l'équation

$$AX = Y$$

possède une unique solution.

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

Remarque 4.3. Si $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$ est un isomorphisme, on sait que $n = p$ nécessairement, et donc les matrices associées (dans n'importe quelles bases) sont nécessairement des matrices carrées. C'est pourquoi les matrices inversibles sont nécessairement carrées.

Proposition 4.4. *Si $A \in M_n(\mathbb{R})$ est inversible, alors on a*

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

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

2. On a donc pour tout $Y \in \mathbb{R}^n$, $AA^{-1}Y = AX = Y$, d'où $AA^{-1} = I_n$. De même, pour tout $X \in \mathbb{R}^n$, $A^{-1}AX = A^{-1}Y = X$, d'où $A^{-1}A = I_n$. \square

Matrices inversibles et bases quelconques. Nous allons avoir besoin d'utiliser des matrices associées aux isomorphismes dans des bases quelconques.

Proposition 4.5. Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application linéaire, $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $B' = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ deux bases quelconques de \mathbb{R}^n .

Alors, f est un isomorphisme si et seulement si la matrice $A = \text{Mat}_{B,B'}(f)$ est inversible, auquel cas

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

Démonstration. Soit $\vec{v} \in \mathbb{R}^n$, l'équation $f(\vec{u}) = \vec{v}$ s'écrit en notation matricielle

$$\text{Mat}_{B,B'}(f)X = AX = Y$$

où Y désigne les coordonnées de \vec{v} dans B' , et X celles de \vec{u} dans B .

f est un isomorphisme ssi pour tout $\vec{v} \in \mathbb{R}^n$, l'équation $f(\vec{u}) = \vec{v}$ possède une unique solution, c'est-à-dire ssi $AX = Y$ possède une unique solution $\Leftrightarrow A$ est inversible. On a de plus dans ce cas,

$$\vec{u} = f^{-1}(\vec{v}) \Leftrightarrow X = \text{Mat}_{B',B}(f^{-1})Y = A^{-1}Y$$

d'où $\text{Mat}_{B',B}(f^{-1}) = A^{-1}$. □

Critères d'inversibilité. D'après la proposition 3.9, on a un critère de rang pour déterminer si une matrice est inversible.

Proposition 4.6. Une matrice carrée $A \in M_n(\mathbb{R})$ est inversible si et seulement si $\text{rang}(A) = n$ ou $\ker A = \{\vec{0}\}$.

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

4.3 Exemples

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

1. A est-elle inversible?
2. Si oui, calculer son inverse.

- Pour 1, il suffit de calculer le rang de A . On a

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

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

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

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

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

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

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

Le cas des matrices 2×2 . Il n'y a pas de formule simple de l'inverse d'une matrice carrée en dimension $n \geq 4$.³ Ce n'est pas le cas pour les matrices 2×2 .

Théorème 4.7. Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ est inversible si et seulement si son **déterminant** $\det A = ad - bc \neq 0$, auquel cas on a :

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

3. Heureusement, les programmes comme Scilab accomplissent très efficacement ces calculs !

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

$$\begin{aligned} \begin{cases} ax_1 + bx_2 = y_1 \\ cx_1 + dx_2 = y_2 \end{cases} &\Rightarrow \begin{cases} (ad - bc)x_1 = dy_1 - by_2 & (dL_1 - bL_2) \\ (ad - bc)x_2 = -cy_1 + ay_2 & (aL_2 - cL_1) \end{cases} \\ &\Rightarrow \begin{cases} x_1 = \frac{1}{\det A}(dy_1 - by_2) \\ x_2 = \frac{1}{\det A}(-cy_1 + ay_2) \end{cases} \quad \text{si } \det A \neq 0. \end{aligned}$$

Dans ce cas, A est une matrice carré injective et donc inversible, et A^{-1} est la matrice souhaitée.

- Si $\det A = 0$, on vérifie facilement que $\begin{pmatrix} d \\ -c \end{pmatrix}$ et $\begin{pmatrix} -b \\ a \end{pmatrix} \in \ker A$, et donc A n'est pas inversible. \square

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

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

5 Changement de base

Nous allons étudier l'effet d'un changement de base :

- sur les coordonnées d'un vecteur donné,
- sur les matrices associées aux endomorphismes.

5.1 Matrices de passage

Soient $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $B' = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ deux bases de \mathbb{R}^n . Chaque vecteur \vec{v}_j de la « nouvelle base » B' à des coordonnées dans « l'ancienne base » B :

$$\vec{v}_j = p_{1j}\vec{u}_1 + p_{2j}\vec{u}_2 + \dots + p_{nj}\vec{u}_n$$

Définition 5.1. On appelle **matrice de passage de B à B'** , la matrice

$$P_{B,B'} = (p_{ij}) \in M_n(\mathbb{R})$$

dont la j -ème colonne est constituée des coordonnées de \vec{v}_j dans B .

Proposition 5.2. 1. Toute matrice de passage est inversible.

2. L'inverse de $P_{B,B'}$ est $P_{B,B'}^{-1} = P_{B',B}$: la matrice de passage de B' à B .

Démonstration. En fait, par définition, $P_{B,B'}$ est la matrice de l'application identité

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

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

$$P_{B,B'} = \text{Mat}_{B',B}(\text{Id}). \quad (8)$$

En effet, on sait que par définition, la j -ème colonne de cette matrice est constituée des coordonnées de $\text{Id}(\vec{v}_j) = \vec{v}_j$ dans la base B .

Comme l'identité est un isomorphisme, la matrice $P_{B,B'}$ est inversible d'après la proposition 4.5, et de plus

$$P_{B,B'}^{-1} = \text{Mat}_{B,B'}(\text{Id}^{-1}) = P_{B',B}.$$

□

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

De manière équivalente mais plus géométrique, on a :

Proposition 5.3. Un isomorphisme transforme une base en une base.

5.2 Changement de base pour les vecteurs

Soient $B = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $B' = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ deux bases de \mathbb{R}^n . Un vecteur \vec{v} donné à des coordonnées X_B dans la base B et $X_{B'}$ dans la base B' . La matrice de passage de B à B' permet de lier ces deux systèmes de coordonnées.

Théorème 5.4. On a

$$X_B = P_{B,B'} X_{B'}.$$

Démonstration. D'après (8), on a $P_{B,B'} = \text{Mat}_{B',B}(\text{Id})$. Si \vec{v} est un vecteur, l'identité $\text{Id}(\vec{v}) = \vec{v}$ devient en notation matricielle

$$X_B = \text{Mat}_{B',B}(\text{Id})X_{B'} = P_{B,B'}X_{B'}.$$

□

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

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

$$X_{B'} = P_{B',B}X_B = P_{B,B'}^{-1}X_B.$$

Il faut donc inverser la matrice $P_{B,B'}$.

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

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

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

$$\begin{aligned} X_B = \begin{pmatrix} x \\ y \end{pmatrix} &= P_{B,B'}X_{B'} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \\ \Leftrightarrow \begin{cases} x = x' \cos \theta - y' \sin \theta \\ y = x' \sin \theta + y' \cos \theta \end{cases} \end{aligned}$$

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

Pour avoir $X_{B'}$ à l'aide de X_B , il faut inverser $P_{B,B'}$, soit en utilisant la relation générale du théorème 4.7, soit en remarquant que $P_{B,B'}$ est en fait la matrice de la rotation d'angle θ , et que donc son inverse est la matrice de rotation d'angle $-\theta$:

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

On a donc

$$X_{B'} = P_{B,B'}^{-1} X_B \Leftrightarrow \begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases}$$

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

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

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

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

5.3 Changement de base pour les endomorphismes

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

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

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

Théorème 5.5. *Avec les notations précédentes, on a*

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

Démonstration. On donne deux preuves.

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

$$\begin{aligned} {}_{B'}f_{B'} &= ({}_{B'}\text{Id}_B)({}_Bf_B)({}_B\text{Id}_{B'}) \\ \Leftrightarrow \text{Mat}_{B',B'}(f) &= \text{Mat}_{B',B'}(\text{Id})\text{Mat}_{B,B}(f)\text{Mat}_{B,B}(\text{Id}) \quad \text{d'après (8)} \\ \Leftrightarrow A' &= P^{-1}AP. \end{aligned}$$

- La seconde preuve utilise la formule de changement de base pour les vecteurs. Soit \vec{u} quelconque et $\vec{v} = f(\vec{u})$. On note X_B et $X_{B'}$ les coordonnées de \vec{u} , et Y_B et $Y_{B'}$ celles de \vec{v} . On a, d'après le théorème 5.4,

$$\begin{aligned} AX_B &= Y_B \text{ avec } Y_B = PY_{B'} \text{ et } X_B = PX_{B'} \\ \Rightarrow APX_{B'} &= PY_{B'} \\ \Rightarrow (P^{-1}AP)X_{B'} &= Y_{B'} = A'X_{B'}. \end{aligned}$$

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

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

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

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

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

On a

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

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

Ici, il est plus efficace de remarquer que

$$\begin{aligned} A \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \text{ et } A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ \Leftrightarrow f(\vec{u}_1) &= \vec{u}_1, \quad f(\vec{u}_2) = \vec{u}_2 \text{ et } f(\vec{u}_3) = -\vec{u}_3. \end{aligned}$$

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

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

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

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

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

ce qui est clair sur A' , mais aussi nécessairement

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

ce qui ne saute pas aux yeux !

5.4 La trace d'une matrice

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

Définition 5.6. Soit

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

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

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

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

Théorème 5.7. Si A et $A' \in M_n(\mathbb{R})$ satisfont $A' = P^{-1}AP$, alors

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

Cela provient du lemme suivant.

Lemme 5.8. Soient $A, B \in M_n(\mathbb{R})$, alors on a $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Démonstration du lemme. D'après (6), on a

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

d'où

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

□

Démonstration du théorème. On a

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

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

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

□

email : michel.rumin@math.u-psud.fr

page web : <http://www.math.u-psud.fr/~rumin/enseignement/enseignement.html>