Représentation matricielle des applications linéaires

Dans tout ce chapitre, $\mathbb{K}$ est l'un des corps $\mathbb{R}$ ou $\mathbb{C}$ et les lettres $n, p, q \ldots$ désignent des entiers naturels non nuls. Tous les résultats présentés demeurent cela dit vrais sur un corps $\mathbb{K}$ quelconque.

1 Matrice d'une application linéaire dans des bases

Définition (Matrice d'une application linéaire dans des bases finies) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives p et n, $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E, $\mathscr{C} = (f_1, \ldots, f_n)$ une base de F et $u \in \mathscr{L}(E, F)$. On appelle *matrice de u dans $\mathscr{B}$ et $\mathscr{C}$* et on note $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ la matrice de la famille $u(\mathscr{B}) = \big(u(e_1), \ldots, u(e_p)\big)$ dans la base $\mathscr{C}$.

Si $E = F$ et $\mathscr{B} = \mathscr{C}$, la matrice $\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(u)$ est simplement notée $\mathrm{Mat}_{\mathscr{B}}(u)$.

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u) = \mathrm{Mat}_{\mathscr{C}}\big(u(\mathscr{B})\big) = \begin{array}{cc} \begin{matrix} u(e_1) & & u(e_j) & & u(e_p) \end{matrix} & \\ \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{np} \end{pmatrix} & \begin{matrix} f_1 \\ \\ f_i \\ \\ f_n \end{matrix} \end{array}$$

Coordonnées de $u(e_j)$ dans $\mathscr{C}$ écrites en colonne

On connaît tout d'une application linéaire quand on connaît la valeur qu'elle prend sur une base, donc quand on connaît sa matrice dans deux bases données. Un exercice peut ainsi commencer sans la moindre ambiguïté de la manière suivante : « On note f l'endomorphisme de $\mathbb{R}_2[X]$ de matrice $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \\ 0 & 4 & 5 \end{pmatrix}$ dans la base canonique. » Il faut alors comprendre que : $f(1) = 3X + 1$, $\quad f(X) = 4X^2 + X \quad$ et $\quad f\big(X^2\big) = 5X^2 + 4X + 2$.

Exemple Pour tout $\mathbb{K}$-espace vectoriel E de dimension finie n et pour toute base $\mathscr{B}$ de E : $\quad \mathrm{Mat}_{\mathscr{B}}\big(\mathrm{Id}_E\big) = I_n$.

Exemple On note T l'endomorphisme $P \longmapsto X^2P'' + P(1)$ de $\mathbb{R}_3[X]$ et $\mathscr{B}_3$ la base canonique de $\mathbb{R}_3[X]$.

Alors $\mathrm{Mat}_{\mathscr{B}_3}(T) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}$ car : $\quad T(1) = 1, \quad T(X) = 1, \quad T\big(X^2\big) = 2X^2 + 1 \quad$ et $\quad T\big(X^3\big) = 6X^3 + 1$.

Théorème (Matrice dans les bases canoniques de l'application linéaire canoniquement associée à une matrice) Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Si on note $\widehat{A}$ l'application linéaire canoniquement associée à A et $\mathscr{B}_p$ et $\mathscr{B}_n$ les bases canoniques respectives de $\mathbb{K}^p$ et $\mathbb{K}^n$: $\quad A = \mathrm{Mat}_{\mathscr{B}_p,\mathscr{B}_n}\big(\widehat{A}\big)$.

Démonstration Réfléchissez, il suffit d'appliquer scrupuleusement la définition. Ce petit résultat doit couler dans vos veines.

Exemple On note φ l'application linéaire canoniquement associée à la matrice $\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\mathscr{B}_2'$ la famille $\big((0,1),(1,0)\big)$ et $\mathscr{B}_3'$ la famille $\big((1,1,1),(1,1,0),(1,0,0)\big)$. Ces familles $\mathscr{B}_2'$ et $\mathscr{B}_3'$ sont alors respectivement des bases de $\mathbb{R}^2$ et $\mathbb{R}^3$, et : $\mathrm{Mat}_{\mathscr{B}_2',\mathscr{B}_3'}(\varphi) = \begin{pmatrix} 1 & -1 \\ 0 & 2 \\ -1 & 0 \end{pmatrix}$. $\quad$ En résumé, si on change les bases, on change la matrice !

Démonstration

- La famille $\mathscr{B}_2'$ est une base de $\mathbb{R}^2$ car sa matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans la base canonique est inversible — d'inverse elle-même. Même idée pour $\mathscr{B}_3'$, sa matrice $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ dans la base canonique est inversible car triangulaire à coefficients diagonaux non nuls après échange de ses première et troisième colonne.

- Ensuite : $\varphi(0,1)=\begin{pmatrix}1&0\\1&1\\-1&1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=(0,1,1)=(1,1,1)-(1,0,0)$.

  De même : $\varphi(1,0)=(1,1,-1)=-(1,1,1)+2(1,1,0)$. Les coordonnées de $\varphi(0,1)$ dans $\mathscr{B}_3'$ sont donc $(1,0,-1)$ et celles de $\varphi(1,0)$ sont $(-1,2,0)$. C'est le résultat voulu.

Théorème (Rang d'une application linéaire, rang d'une matrice associée) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F et $u\in\mathscr{L}(E,F)$. Alors : $\mathrm{rg}(u)=\mathrm{rg}\big(\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\big)$.

Tout rang d'application linéaire peut donc être calculé comme le rang d'une matrice grâce à l'**ALGORITHME DU PIVOT**.

Démonstration D'après le théorème analogue pour les familles de vecteurs :

$$\mathrm{rg}\big(\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\big)=\mathrm{rg}\Big(\mathrm{Mat}_{\mathscr{C}}\big(u(\mathscr{B})\big)\Big)=\mathrm{rg}\big(u(\mathscr{B})\big)=\dim\mathrm{Vect}\big(u(\mathscr{B})\big)=\dim\mathrm{Im}\,u=\mathrm{rg}(u).$$

Théorème (Calcul matriciel de l'image d'un vecteur par une application linéaire) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F, $u\in\mathscr{L}(E,F)$ et $x\in E$. Alors :

$$\mathrm{Mat}_{\mathscr{C}}\big(u(x)\big)=\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\times\mathrm{Mat}_{\mathscr{B}}(x).$$

En d'autres termes, l'**ÉVALUATION** par une application linéaire se traduit matriciellement en termes de **PRODUIT**.

Démonstration Introduisons les vecteurs de $\mathscr{B}$ et $\mathscr{C}$: $\mathscr{B}=(e_1,\ldots,e_p)$ et $\mathscr{C}=(f_1,\ldots,f_n)$, ainsi que les coordonnées de x dans $\mathscr{B}$: $X=\mathrm{Mat}_{\mathscr{B}}(x)$ et la matrice de u dans les bases $\mathscr{B}$ et $\mathscr{C}$: $U=\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

$$u(x)=u\left(\sum_{j=1}^p x_je_j\right)=\sum_{j=1}^p x_ju(e_j)=\sum_{j=1}^p x_j\sum_{i=1}^n u_{ij}f_i=\sum_{i=1}^n\left(\sum_{j=1}^p u_{ij}x_j\right)f_i,$$

donc les coordonnées de $u(x)$ dans $\mathscr{C}$ sont $\left(\sum_{j=1}^p u_{1j}x_j,\ldots,\sum_{j=1}^p u_{nj}x_j\right)$, i.e. le produit $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\times\mathrm{Mat}_{\mathscr{B}}(x)$.

Exemple On note f l'endomorphisme de $\mathbb{R}_2[X]$ de matrice $\begin{pmatrix}3&3&6\\0&1&2\\0&2&4\end{pmatrix}$ dans la base canonique.

Alors : $\mathrm{Im}\,f=\mathrm{Vect}\big(1,2X^2+X\big)$ et $\mathrm{Ker}\,f=\mathrm{Vect}\big(X^2-2X\big)$.

Démonstration Pour commencer :

$$\mathrm{Im}\,f=\mathrm{Vect}\big(f(1),f(X),f(X^2)\big)=\mathrm{Vect}\big(3,2X^2+X+3,4X^2+2X+6\big)=\mathrm{Vect}\big(1,2X^2+X+3\big)=\mathrm{Vect}\big(1,2X^2+X\big).$$

Ensuite, pour tout $P=aX^2+bX+c\in\mathbb{R}_2[X]$: $P\in\mathrm{Ker}\,f\iff f(P)=0\iff\begin{pmatrix}3&3&6\\0&1&2\\0&2&4\end{pmatrix}\begin{pmatrix}c\\b\\a\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$

$$\iff\left\{\begin{array}{rcrcrcl}3c&+&3b&+&6a&=&0\\&&b&+&2a&=&0\\&&2b&+&4a&=&0\end{array}\right.\overset{L_1\leftarrow L_1-3L_2}{\iff}c=0\ \text{ et }\ b=-2a\iff P=aX^2-2aX.$$

Conclusion : $\mathrm{Ker}\,f=\mathrm{Vect}\big(X^2-2X\big)$.

Attention ! Deux remarques sur l'exemple précédent.

- Les coordonnées de aX^2+bX+c dans la base canonique de $\mathbb{R}_2[X]$ sont (c,b,a) **ET NON PAS** (a,b,c).
- On raisonne matriciellement sur un squelette numérique, mais il ne faut pas oublier à la fin de l'exemple précédent de réincarner le résultat dans le monde vectoriel $\mathbb{R}_3[X]$. La réponse $\mathrm{Ker}\,R=\mathrm{Vect}\big((0,-2,1)\big)$ n'est pas correcte.

Théorème (Un dictionnaire entre les points de vue vectoriel et matriciel sur les applications linéaires)

(i) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies respectives p et n, $\mathscr{B}$ une base de E et $\mathscr{C}$ une base de F. L'application $u\longmapsto\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ est un isomorphisme de $\mathscr{L}(E,F)$ sur $\mathscr{M}_{n,p}(\mathbb{K})$.

(ii) Soient E,F,G trois $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles de bases respectives $\mathscr{B},\mathscr{C},\mathscr{D}$ et $u\in\mathscr{L}(E,F)$ et $v\in\mathscr{L}(F,G)$. Alors : $\mathrm{Mat}_{\mathscr{B},\mathscr{D}}(v\circ u)=\mathrm{Mat}_{\mathscr{C},\mathscr{D}}(v)\times\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

En particulier, l'application $u\longmapsto\mathrm{Mat}_{\mathscr{B}}(u)$ est un isomorphisme d'anneaux de $\mathscr{L}(E)$ sur $\mathscr{M}_n(\mathbb{K})$ si on pose $n=\dim E$.

(iii) Soient E et F deux $\mathbb{K}$-espaces vectoriels de **MÊMES DIMENSIONS** finies non nulles, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F et $u\in\mathscr{L}(E,F)$. Alors u est un isomorphisme de E sur F si et seulement si $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ est inversible.

Dans ce cas : $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}\big(u^{-1}\big)=\big(\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\big)^{-1}$.

En résumé, l'assertion (i) exprime deux choses :

— une propriété de linéarité : $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(\lambda u+\mu v)=\lambda\,\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)+\mu\,\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(v)$ avec des notations évidentes,

— une propriété de bijectivité déjà mentionnée informellement — on connaît entièrement f quand on connaît $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

Elle relie aussi en passant deux résultats bien connus : $\dim \mathscr{M}_{n,p}=np$ et $\dim \mathscr{L}(E,F)=\dim E\times\dim F$.

L'assertion (ii) montre que le **PRODUIT** est aux matrices ce que la **COMPOSITION** est aux applications linéaires. Nous connaissions déjà ce résultat dans le cas particulier des applications linéaires canoniquement associées à des matrices.

Démonstration Je vous laisse démontrer seuls l'assertion (i).

(ii) Introduisons les vecteurs de $\mathscr{B}$: $\mathscr{B}=(e_1,\dots,e_n)$. Pour tout $j\in[\![1,n]\!]$:

$$\mathrm{Mat}_{\mathscr{B}}(e_j)=\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}\begin{array}{l}\\ \\ \leftarrow \text{position } j\\ \\ \\ \end{array}$$

$$\mathrm{Mat}_{\mathscr{B},\mathscr{D}}(v\circ u)\times\mathrm{Mat}_{\mathscr{B}}(e_j)=\mathrm{Mat}_{\mathscr{D}}\big(v\circ u(e_j)\big)=\mathrm{Mat}_{\mathscr{C},\mathscr{D}}(v)\times\mathrm{Mat}_{\mathscr{C}}\big(u(e_j)\big)=\mathrm{Mat}_{\mathscr{C},\mathscr{D}}(v)\times\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\times\mathrm{Mat}_{\mathscr{B}}(e_j),$$

mais n'oublions pas que $\mathrm{Mat}_{\mathscr{B}}(e_j)$ est le $j^{\text{ème}}$ vecteur de la base canonique de $\mathbb{K}^n$. Nous venons donc de montrer que $\mathrm{Mat}_{\mathscr{B},\mathscr{D}}(v\circ u)$ et $\mathrm{Mat}_{\mathscr{C},\mathscr{D}}(v)\times\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ ont les mêmes $j^{\text{èmes}}$ colonnes, et ce pour tout $j\in[\![1,n]\!]$. Comme voulu : $\mathrm{Mat}_{\mathscr{B},\mathscr{D}}(v\circ u)=\mathrm{Mat}_{\mathscr{C},\mathscr{D}}(v)\times\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

(iii) Si u est bijective et si on pose $n=\dim F$: $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)\times\mathrm{Mat}_{\mathscr{C},\mathscr{B}}\big(u^{-1}\big)=\mathrm{Mat}_{\mathscr{C}}\big(\mathrm{Id}_F\big)=I_n$, donc $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ est inversible d'inverse $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}\big(u^{-1}\big)$.

Réciproquement, si $A=\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ est inversible, notons v l'unique application linéaire de F dans E pour laquelle $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(v)=A^{-1}$. Aussitôt : $\mathrm{Mat}_{\mathscr{B}}(v\circ u)=\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(v)\times\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)=A^{-1}A=I_n$ et de même $\mathrm{Mat}_{\mathscr{C}}(u\circ v)=I_n$, donc $v\circ u=\mathrm{Id}_E$ et $u\circ v=\mathrm{Id}_F$, autrement dit u est bijective de E sur F. ■

Exemple L'endomorphisme ω de $\mathbb{R}_3[X]$ dont la matrice dans la base canonique de $\mathbb{R}_3[X]$ est $\Omega=\begin{pmatrix}0&0&1&-1\\-1&1&1&-1\\0&2&1&-2\\-1&2&1&-2\end{pmatrix}$ est la symétrie par rapport à $\mathrm{Vect}\big(X^3+X^2+X,X^2+1\big)$ parallèlement à $\mathrm{Vect}\big(X^3+X+1,X^3+X^2\big)$.

Démonstration

- Par définition, ω est linéaire. Montrer que c'est une symétrie revient donc à montrer que $\omega^2=\mathrm{Id}_{\mathbb{R}_3[X]}$, ou encore matriciellement que $\Omega^2=I_4$ — ce qui est très facile à vérifier.
- {Etudions le sous-espace vectoriel $\mathrm{Ker}\big(\omega-\mathrm{Id}_{\mathbb{R}_3[X]}\big)$ de $\mathbb{R}_3[X]$ par rapport auquel ω est une symétrie. Pour tout $P=aX^3+bX^2+cX+d\in\mathbb{R}_3[X]$: $P\in\mathrm{Ker}\big(\omega-\mathrm{Id}_{\mathbb{R}_3[X]}\big)\iff\omega(P)=P$

$$\iff\quad\Omega\begin{pmatrix}d\\c\\b\\a\end{pmatrix}=\begin{pmatrix}d\\c\\b\\a\end{pmatrix}\quad\iff\quad\begin{cases}b-a=d\\-d+c+b-a=c\\2c+b-2a=b\\-d+2c+b-2a=a\end{cases}$$

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

$$\iff\quad\exists\,\lambda,\mu\in\mathbb{R},\quad\begin{cases}a=\lambda\\b=\lambda+\mu\\c=\lambda\\d=\mu.\end{cases}\qquad\text{Ainsi }\mathrm{Ker}\big(\omega-\mathrm{Id}_{\mathbb{R}_3[X]}\big)=\mathrm{Vect}\big(X^3+X^2+X,X^2+1\big).$$

- On montre de la même manière que $\mathrm{Ker}\big(\omega+\mathrm{Id}_{\mathbb{R}_3[X]}\big)=\mathrm{Vect}\big(X^3+X+1,X^3+X^2\big)$.

Définition-théorème (Condition nécessaire et suffisante d'inversibilité d'une matrice de Vandermonde) Soient $x_1,\dots,x_n\in\mathbb{K}$.

On appelle *matrice de Vandermonde de* $x_1,\dots,x_n$ la matrice carrée : $\big(x_i^{j-1}\big)_{1\leqslant i,j\leqslant n}=\begin{pmatrix}1&x_1&x_1^2&\cdots&x_1^{n-1}\\1&x_2&x_2^2&\cdots&x_2^{n-1}\\\vdots&\vdots&\vdots&&\vdots\\1&x_n&x_n^2&\cdots&x_n^{n-1}\end{pmatrix}$.

Cette matrice est inversible si et seulement si les scalaires $x_1,\dots,x_n$ sont distincts.

Démonstration

- Si deux des scalaires $x_1,\dots,x_n$ sont égaux, leur matrice de Vandermonde possède deux lignes égales, donc n'est pas inversible.

- Réciproquement, supposons $x_1,\dots,x_n$ distincts et notons φ l'application linéaire $P \longmapsto \big(P(x_1),\dots,P(x_n)\big)$ de $\mathbb{K}_{n-1}[X]$ dans $\mathbb{K}^n$. Cette application est injective car pour tout $P \in \operatorname{Ker}\varphi$: $P(x_1)=\ldots=P(x_n)=0$, donc le polynôme P possède n racines distinctes alors qu'il est de degré au plus $n-1$ — ainsi $P=0$. Comme $\dim \mathbb{K}_{n-1}[X] = n = \dim \mathbb{K}^n$, cette injectivité fait de φ un isomorphisme de $\mathbb{K}_{n-1}[X]$ sur $\mathbb{K}^n$.

  En particulier, la matrice de φ dans la base canonique de $\mathbb{K}_{n-1}[X]$ au départ et la base canonique de $\mathbb{K}^n$ à l'arrivée est inversible d'après le théorème précédent, or cette matrice est exactement la matrice de Vandermonde de $x_1,\dots,x_n$. ∎

Quelques mots à présent sur l'interprétation géométrique des blocs qu'on lit sur la une matrice d'application linéaire. Soient E un $\mathbb{K}$-espace vectoriel, F et G deux sous-espaces vectoriels supplémentaires de E de dimensions respectives p et q et $\mathscr{B}=(e_1,\dots,e_{p+q})$ une base de E adaptée à la décomposition $E=F\oplus G$. Pour rappel, cela signifie que $(e_1,\dots,e_p)$ est une base de F et $(e_{p+1},\dots,e_{p+q})$ une base de G.

Soit $u\in\mathscr{L}(E)$. La matrice de u dans $\mathscr{B}$ s'écrit $\operatorname{Mat}_{\mathscr{B}}(u)=\begin{pmatrix} A & C \\ B & D\end{pmatrix}$ pour certaines $A\in\mathscr{M}_p(\mathbb{K})$, $B\in\mathscr{M}_{q,p}(\mathbb{K})$, $C\in\mathscr{M}_{p,q}(\mathbb{K})$ et $D\in\mathscr{M}_q(\mathbb{K})$. Nous allons tâcher de comprendre sur deux situations importantes de quelle manières les blocs A, B, C et D peuvent être interprétés géométriquement.

- **À quelle condition a-t-on $B=0$?** Tout simplement :
  $$B=0 \iff \forall j\in[\![1,p]\!],\quad u(e_j)\in\operatorname{Vect}(e_1,\dots,e_p) \iff \forall x\in F,\quad u(x)\in F \iff F \text{ est stable par } u.$$
  Dans ces conditions, $u|_F$ est un endomorphisme de F et $A=\operatorname{Mat}_{(e_1,\dots,e_p)}\big(u|_F\big)$.
- **À quelle condition a-t-on $B=C=0$?** Comme au point précédent, $B=C=0$ si et seulement si F et G sont tous les deux stables par u. Dans ces conditions, $u|_F$ et $u|_G$ sont des endomorphismes de F et G respectivement et $A=\operatorname{Mat}_{(e_1,\dots,e_p)}\big(u|_F\big)$ et $D=\operatorname{Mat}_{(e_{p+1},\dots,e_{p+q})}\big(u|_G\big)$.

Exemple Soient E un $\mathbb{K}$-espace vectoriel, F et G deux sous-espaces vectoriels supplémentaires de E de dimensions respectives q et r et $\mathscr{B}$ une base de E adaptée à la décomposition $E=F\oplus G$. Notons p la projection sur F parallèlement à G et s la symétrie par rapport à F parallèlement à G. Alors $\operatorname{Mat}_{\mathscr{B}}(p)=\begin{pmatrix} I_q & \\ & 0_r\end{pmatrix}$ et $\operatorname{Mat}_{\mathscr{B}}(s)=\begin{pmatrix} I_q & \\ & -I_r\end{pmatrix}$.

> **Démonstration** Introduisons les vecteurs de $\mathscr{B}$: $\mathscr{B}=(e_1,\dots,e_{q+r})$. Par définition des projections et des symétries : $p(e_i)=s(e_i)=e_i$ pour tout $i\in[\![1,q]\!]$ et : $p(e_j)=0_E$ et $s(e_j)=-e_j$ pour tout $j\in[\![q+1,q+r]\!]$. Le résultat en découle.

L'exemple qui suit est emblématique de nombreux exercices. En dimension finie, on peut calculer la matrice d'un endomorphisme dans n'importe quelle base, mais n'y a-t-il pas des bases dans lesquelles le résultat est plus simple et plus joli que dans d'autres ? L'étude de cette question est une branche de l'algèbre linéaire que vous étudierez davantage en deuxième année, appelée *réduction*. *Réduire* un endomorphisme, c'est trouver une base dans laquelle sa matrice est facile à interpréter — par exemple diagonale, triangulaire, pleine de zéros… —et *réduire* une matrice carrée, c'est réduire l'endomorphisme qui lui est canoniquement associé.

Exemple Soient E un $\mathbb{K}$-espace vectoriel de dimension 2 et $f\in\mathscr{L}(E)$. Si f est nilpotent d'indice 2, i.e. si $f^2=0_{\mathscr{L}(E)}$ mais $f\neq 0_{\mathscr{L}(E)}$, alors f a pour matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}$ dans une certaine base de E.

> **Démonstration** On cherche une base (e_1,e_2) de E pour laquelle $f(e_1)=0_E$ et $f(e_2)=e_1$. Si une telle base existe, on peut aussi l'écrire $\big(f(e_2),e_2\big)$ où e_2 est un vecteur de E pour lequel $f(e_2)=e_1\neq 0_E$, i.e. n'appartenant pas à $\operatorname{Ker} f$. Un tel choix de vecteur e_2 est-il cependant possible ? Oui, car f étant nilpotent d'indice 2 : $f\neq 0_{\mathscr{L}(E)}$, donc $\operatorname{Ker} f\neq E$. Donnons-nous donc un vecteur e_2 de $E\setminus\operatorname{Ker} f$ et posons $e_1=f(e_2)$. Dans ces conditions, si jamais (e_1,e_2) est une base de E : $\operatorname{Mat}_{(e_1,e_2)}(f)=\begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}$ car $f(e_1)=f^2(e_2)=0_E$ et $f(e_2)=e_1$ par construction. Pour une raison de dimension, il ne nous reste donc plus qu'à montrer la liberté de la famille (e_1,e_2). Soient $\lambda,\mu\in\mathbb{K}$. On suppose que $\lambda e_1+\mu e_2=0_E$. Composons par f : $\mu e_1=0_E$, or $e_1\neq 0_E$, donc $\mu=0$. En retour $\lambda e_1=0_E$, donc de même $\lambda=0$.

2 Changements de bases, équivalence et similitude

2.1 Changements de bases

Définition-théorème (Matrice de passage d'une base à une autre) Soient E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ trois bases de E.

On appelle *matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$* la matrice $\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}') = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)$, souvent notée $P_{\mathscr{B}}^{\mathscr{B}'}$.

Deux choses à savoir : (i) $P_{\mathscr{B}}^{\mathscr{B}'}$ est inversible d'inverse $P_{\mathscr{B}'}^{\mathscr{B}}$. (ii) $P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}''} = P_{\mathscr{B}}^{\mathscr{B}''}$.

Démonstration

(i) $P_{\mathscr{B}}^{\mathscr{B}'}$ est la matrice d'un isomorphisme : $\left(P_{\mathscr{B}}^{\mathscr{B}'}\right)^{-1} = \left(\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)\right)^{-1} = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}\left(\mathrm{Id}_E^{-1}\right) \overset{\mathrm{Id}_E^{-1}=\mathrm{Id}_E}{=} P_{\mathscr{B}'}^{\mathscr{B}}$.

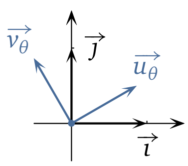
(ii) $P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}''} = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E) \times \mathrm{Mat}_{\mathscr{B}'',\mathscr{B}'}(\mathrm{Id}_E) = \mathrm{Mat}_{\mathscr{B}'',\mathscr{B}}(\mathrm{Id}_E \circ \mathrm{Id}_E) = \mathrm{Mat}_{\mathscr{B}'',\mathscr{B}}(\mathrm{Id}_E) = P_{\mathscr{B}}^{\mathscr{B}''}$.

Théorème (Changement de base pour un vecteur) Soient E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. On pose $P = P_{\mathscr{B}}^{\mathscr{B}'}$. Pour tout $x \in E$ de coordonnées X dans $\mathscr{B}$ et X' dans $\mathscr{B}'$: $X = PX'$.

Démonstration L'égalité $x = \mathrm{Id}_E(x)$ s'écrit matriciellement dans les bases adaptées $X = PX'$.

Exemple Soit $\theta \in \mathbb{R}$ fixé. Notons $\mathscr{B} = \left(\vec{\imath}, \vec{\jmath}\right)$ la base canonique de $\mathbb{R}^2$ et posons : $\begin{cases} \overrightarrow{u_\theta} = \cos\theta\, \vec{\imath} + \sin\theta\, \vec{\jmath} \\ \overrightarrow{v_\theta} = -\sin\theta\, \vec{\imath} + \cos\theta\, \vec{\jmath}. \end{cases}$ On définit ainsi une base $\mathscr{B}_\theta = \left(\overrightarrow{u_\theta}, \overrightarrow{v_\theta}\right)$ de $\mathbb{R}^2$.

En outre, soit $\vec{u} = (x,y) \in \mathbb{R}^2$. Les coordonnées de $\vec{u}$ dans $\mathscr{B}$ sont bien sûr $X = \begin{pmatrix} x \\ y \end{pmatrix}$. Si nous notons $X_\theta = \begin{pmatrix} x_\theta \\ y_\theta \end{pmatrix}$ les coordonnées de $\vec{u}$ dans la base $\mathscr{B}_\theta$: $\begin{cases} x = x_\theta \cos\theta - y_\theta \sin\theta \\ y = x_\theta \sin\theta + y_\theta \cos\theta \end{cases}$ et $\begin{cases} x_\theta = x\cos\theta + y\sin\theta \\ y_\theta = -x\sin\theta + y\cos\theta. \end{cases}$



Démonstration D'abord : $\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, et cette matrice est inversible car son déterminant vaut $\cos^2\theta + \sin^2\theta = 1 \neq 0$. Comme voulu, $\mathscr{B}_\theta$ est une base de $\mathbb{R}^2$.

Ensuite, sachant que $P_{\mathscr{B}}^{\mathscr{B}_\theta} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$: $X = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} X'$. Pour l'autre formule, simplement calculer l'inverse de $P_{\mathscr{B}}^{\mathscr{B}_\theta}$: $P_{\mathscr{B}_\theta}^{\mathscr{B}} = \left(P_{\mathscr{B}}^{\mathscr{B}_\theta}\right)^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

Théorème (Changement de bases pour une application linéaire) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, $\mathscr{C}$ et $\mathscr{C}'$ deux bases de F et $u \in \mathscr{L}(E,F)$.
On pose : $P = P_{\mathscr{B}}^{\mathscr{B}'}$, $Q = P_{\mathscr{C}}^{\mathscr{C}'}$, $A = \mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ et $A' = \mathrm{Mat}_{\mathscr{B}',\mathscr{C}'}(u)$. Alors : $A' = Q^{-1}AP$.

Il est important de se donner des mots pour décrire chacune des données de cet énoncé. Tout simplement, E est l'espace de départ de u, F son espace d'arrivée, $\mathscr{B}$ et $\mathscr{C}$ sont les « anciennes » bases, i.e. les bases avant changement de bases, et $\mathscr{B}'$ et $\mathscr{C}'$ les « nouvelles » bases, i.e. les bases après changement. Départ/arrivée/ancien/nouveau !

$$\begin{tikzcd} E, \mathscr{B}' \arrow[r, "\mathrm{Id}_E, P"] \arrow[d, "u, A'"'] & E, \mathscr{B} \arrow[d, "u, A"] \\ F, \mathscr{C}' \arrow[r, "\mathrm{Id}_F, Q"'] & F, \mathscr{C} \end{tikzcd}$$

Démonstration Le diagramme ci-contre peut presque tenir lieu de preuve. L'égalité $u \circ \mathrm{Id}_E = \mathrm{Id}_F \circ u$ s'écrit matriciellement $AP = QA'$, i.e. $A' = Q^{-1}AP$.

✖ **Attention !** Il y a DEUX formules de changement de bases, une pour les vecteurs et une pour les applications linéaires, merci de ne pas les confondre !

Théorème (Changement de bases et matrice J_r) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives p et n et $u \in \mathscr{L}(E,F)$ de rang r. Alors pour une certaine base $\mathscr{B}$ de E et une certaine base $\mathscr{C}$ de F : $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u) = J_r$, où J_r est la matrice de taille $n \times p$ suivante : $J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Démonstration Soient $\mathscr{B} = (e_i)_{1 \leqslant j \leqslant p}$ une base de E et $\mathscr{C} = (f_i)_{1 \leqslant i \leqslant n}$ une base de F. Est-il possible d'imposer à ces bases que la matrice de u y soit J_r ?

- Pour que les $p-r$ dernières colonnes de $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ soient nulles, il faut et il suffit que les vecteurs $e_{r+1},\dots,e_p$ soient éléments de $\mathrm{Ker}\,u$ et linéairement indépendants. Comme $\mathrm{Ker}\,u$ est de dimension $p-r$ d'après le théorème du rang, nous n'avons qu'à choisir pour famille $(e_j)_{r+1\leqslant j\leqslant p}$ une base de $\mathrm{Ker}\,u$ et la compléter simplement en une base $\mathscr{B}$ de E.
- Pour que les r premières colonnes de $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$ soient ce qu'on veut, on peut poser $f_j=u(e_j)$ pour tout $j\in[\![1,r]\!]$, puis compléter en une base $\mathscr{C}$ de F, MAIS CE N'EST POSSIBLE QUE SI LA FAMILLE $(f_j)_{1\leqslant j\leqslant r}$ EST LIBRE. Or la famille $(e_j)_{1\leqslant j\leqslant r}$ engendre par construction un supplémentaire I de $\mathrm{Ker}\,u$ dans E, donc $u|_I$ est un isomorphisme de I sur $\mathrm{Im}\,u$. La famille $(f_j)_{1\leqslant j\leqslant r}=\big(u(e_j)\big)_{1\leqslant j\leqslant r}$ est ainsi libre.

Ces deux points garantissent bien l'existence de deux bases $\mathscr{B}$ et $\mathscr{C}$ pour lesquelles $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)=J_r$.

2.2 MATRICES ÉQUIVALENTES

Définition-théorème (Matrices équivalentes)

- **Définition :** Soient $A,B\in\mathscr{M}_{n,p}(\mathbb{K})$. On dit que B est *équivalente à* A s'il existe deux matrices $P\in\mathrm{GL}_p(\mathbb{K})$ et $Q\in\mathrm{GL}_n(\mathbb{K})$ inversibles pour lesquelles $B=Q^{-1}AP$.
- **Premier exemple fondamental :** Si B a été obtenue à partir de A après une série d'opérations élémentaires, alors B est équivalente à A.
- **Deuxième exemple fondamental :** Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, $\mathscr{C}$ et $\mathscr{C}'$ deux bases de F et $u\in\mathscr{L}(E,F)$. La matrice $\mathrm{Mat}_{\mathscr{B}',\mathscr{C}'}(u)$ est équivalente à la matrice $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

Démonstration Pour le premier exemple fondamental, se souvenir du fait que les opérations élémentaires peuvent être vues comme des multiplications par des matrices inversibles. Le deuxième exemple fondamental n'est qu'une reformulation du théorème de changement de base pour une application linéaire.

Théorème (Propriétés de la relation d'équivalence)

(i) **Relation d'équivalence :** La relation d'équivalence sur $\mathscr{M}_{n,p}(\mathbb{K})$ est une… relation d'équivalence — mais pas dans le même sens !

(ii) **Caractérisation par le rang :** Deux matrices de $\mathscr{M}_{n,p}(\mathbb{K})$ sont équivalentes si et seulement si elles ont le même rang. Cela revient aussi à dire que toute matrice de $\mathscr{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à J_r.

Démonstration

(i) Soient $A,B,C\in\mathscr{M}_{n,p}(\mathbb{K})$.

Réflexivité : A est équivalente à A car I_p et I_n sont inversibles et $A=I_n^{-1}AI_p$.

Symétrie : Si B est équivalente à A, i.e. $B=Q^{-1}AP$ pour certaines matrices $P\in\mathrm{GL}_p(\mathbb{K})$ et $Q\in\mathrm{GL}_n(\mathbb{K})$, alors P^{-1} et Q^{-1} sont inversibles et $A=\left(Q^{-1}\right)^{-1}BP^{-1}$, donc A est équivalente à B.

Transitivité : Si B est équivalente à A et C équivalente à B, i.e. $B=Q^{-1}AP$ et $C=Q'^{-1}BP'$ pour certaines matrices $P,P'\in\mathrm{GL}_p(\mathbb{K})$ et $Q,Q'\in\mathrm{GL}_n(\mathbb{K})$, alors PP' et QQ' sont inversibles et :

$$C=Q'^{-1}BP'=Q'^{-1}Q^{-1}APP'=(QQ')^{-1}A(PP'),\qquad \text{donc } C \text{ est équivalente à } A.$$

(ii) Soit $A\in\mathscr{M}_{n,p}(\mathbb{K})$ de rang r. Notons $\widehat{A}$ l'application linéaire canoniquement associée à A et $\mathscr{B}_n$ et $\mathscr{B}_p$ les bases canoniques respectives de $\mathbb{K}^n$ et $\mathbb{K}^p$. On rappelle que $\mathrm{Mat}_{\mathscr{B}_p,\mathscr{B}_n}\big(\widehat{A}\big)=A$. Or d'après un théorème précédent : $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}\big(\widehat{A}\big)=J_r$ pour certaines bases $\mathscr{B}$ de $\mathbb{K}^p$ et $\mathscr{C}$ de $\mathbb{K}^n$, donc si on pose $P=P_{\mathscr{B}_p}^{\mathscr{B}}$ et $Q=P_{\mathscr{B}_n}^{\mathscr{C}}$: $J_r=Q^{-1}AP$ par changement de base, donc A et J_r sont équivalentes.

Le résultat qui suit n'est pas tout à fait à sa place dans ce paragraphe sur les matrices équivalentes, mais il découle de la caractérisation précédente.

Théorème (Invariance du rang par transposition) Pour tout $A\in\mathscr{M}_{n,p}(\mathbb{K})$: $\mathrm{rg}\big(A^\top\big)=\mathrm{rg}(A)$.

Démonstration D'après le théorème précédent, A est équivalente à J_r pour $r = \text{rg}(A)$, donc par simple transposition dans la définition, $A^\top$ est équivalente à $J_r^\top$ — attention, J_r est de taille $n \times p$ tandis que $J_r^\top$ est de taille $p \times n$. Conclusion : $\text{rg}(A^\top) = \text{rg}(J_r^\top) = r = \text{rg}(J_r) = \text{rg}(A)$.

La définition et les résultats qui suivent n'ont rien à voir avec les matrices équivalentes, mais ils requièrent l'invariance du rang par transposition, donc nous ne pouvions pas les énoncer jusqu'ici.

Définition-théorème (Matrices extraites) Soient $A \in \mathscr{M}_{n,p}(\mathbb{K})$.

(i) **Définition :** On appelle *matrice extraite de A* toute matrice de la forme $\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_{p'}} \\ \vdots & & \vdots \\ a_{i_{n'} j_1} & \cdots & a_{i_{n'} j_{p'}} \end{pmatrix}$ où $i_1, \ldots, i_{n'}, j_1, \ldots, j_{p'}$ satisfont les inégalités $1 \leqslant i_1 < \ldots < i_{n'} \leqslant n$ et $1 \leqslant j_1 < \ldots < j_{p'} \leqslant p$.

(ii) **Rang d'une matrice extraite :** Pour toute matrice B extraite de A : $\text{rg}(B) \leqslant \text{rg}(A)$.

(iii) **Caractérisation du rang par les matrices carrées extraites :** Le rang de A est la taille maximale des matrices inversibles qu'on peut extraire de A.

Démonstration

(ii) La matrice B est obtenue à partir de A par supression d'un certain nombre de lignes et de colonnes. Notons B' la matrice intermédiaire obtenue quand on supprime seulement les colonnes. On passe de A à B' par une suppression de colonnes, puis de B' à B par une suppression de lignes. Le rang d'une matrice étant par définition le rang de la famille de ses colonnes : $\text{rg}(A) \geqslant \text{rg}(B') = \text{rg}(B'^\top) \geqslant \text{rg}(B^\top) = \text{rg}(B)$.

(iii) Il nous suffit d'établir l'équivalence suivante — avec $r \in \mathbb{N}^*$:

$$\text{rg}(A) \geqslant r \quad \iff \quad \text{On peut extraire de } A \text{ une matrice inversible de taille } r.$$

- Si on peut extraire de A une matrice inversible de taille r, alors d'après (ii) : $\text{rg}(A) \geqslant r$.
- Réciproquement, supposons $\text{rg}(A) \geqslant r$. On peut donc extraire de la famille des colonnes de A une famille libre de r vecteurs. Matriciellement, nous pouvons donc extraire de A une matrice $B \in \mathscr{M}_{n,r}(\mathbb{K})$ de rang r par suppression de certaines colonnes. En raisonnant de même, nous pouvons extraire de $B^\top$ une matrice $C \in \mathscr{M}_r(\mathbb{K})$ de rang r par suppression de certaines colonnes — i.e. par suppression de certaines lignes de B. La matrice $C^\top$ est finalement une matrice extraite de A inversible de taille r.

Exemple La matrice $\begin{pmatrix} 4 & 10 \\ 5 & 11 \end{pmatrix}$ est extraite de $\begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix}$ — on a retenu les lignes 1 et 2 et les colonnes 2 et 4.

Exemple $\text{rg}\begin{pmatrix} 1 & 1 & 2 & 0 \\ 2 & 3 & 1 & 0 \\ 1 & 0 & 3 & 1 \\ 3 & 0 & 0 & 2 \end{pmatrix} \geqslant 3$ car la matrice extraite $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ est inversible — pourquoi, d'ailleurs ?

2.3 Matrices semblables et trace d'un endomorphisme

Définition-théorème (Matrices semblables)

- **Définition :** Soient $A, B \in \mathscr{M}_n(\mathbb{K})$. On dit que B est *semblable à A (sur* $\mathbb{K}$*)* s'il existe une matrice $P \in \text{GL}_n(\mathbb{K})$ inversible pour laquelle $B = P^{-1}AP$.
- **Exemple fondamental :** Soient E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle, $u \in \mathscr{L}(E)$ et $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Les matrices $\text{Mat}_{\mathscr{B}}(u)$ et $\text{Mat}_{\mathscr{B}'}(u)$ sont alors semblables.

Démonstration Dans l'exemple fondamental, si on pose $P = P_{\mathscr{B}}^{\mathscr{B}'}$, alors $\text{Mat}_{\mathscr{B}'}(u) = P^{-1}\text{Mat}_{\mathscr{B}}(u)P$ par changement de base.

Attention ! On a vite fait de confondre équivalence et similitude. La relation de similitude **n'est définie que** pour des matrices **carrées**, et dans son exemple fondamental, on travaille avec des **endo**morphismes et on a les **mêmes bases au départ et à l'arrivée** — donc deux bases $\mathscr{B}$ et $\mathscr{B}'$ au lieu des quatre $\mathscr{B}$, $\mathscr{C}$, $\mathscr{B}'$ et $\mathscr{C}'$ de l'équivalence.

Théorème (Propriétés de la relation de similitude)

(i) **Relation d'équivalence :** La relation de similitude sur $\mathcal{M}_n(\mathbb{K})$ est une relation d'équivalence.

(ii) **Invariance du rang et de la trace par similitude :** Deux matrices semblables de $\mathcal{M}_n(\mathbb{K})$ ont même trace et même rang.

Démonstration

(i) Reprendre la preuve du résultat analogue pour les matrices équivalentes.

(ii) Soient $A, B \in \mathcal{M}_n(\mathbb{K})$. Si A et B sont semblables, disons $B = P^{-1}AP$ avec $P \in \mathrm{GL}_n(\mathbb{K})$, elles sont équivalentes, donc de même rang, et $\mathrm{tr}(B) = \mathrm{tr}\big(P^{-1}(AP)\big) = \mathrm{tr}\big((AP)P^{-1}\big) = \mathrm{tr}(A)$ d'après les propriétés de la trace.

Exemple Les matrices $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 3 \\ 2 & 2 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 1 & 1 & 3 \end{pmatrix}$ ne sont pas semblables car elles n'ont pas la même trace, mais elles sont équivalentes car on peut passer de l'une à l'autre par une opération élémentaire très simple — n'est-ce pas ?

Exemple Les matrices $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 4 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ sont semblables.

Démonstration Notons (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$ et f l'endomorphisme de $\mathbb{R}^3$ canoniquement associé à la matrice $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Comme : $f(e_1) = (0,0,0)$, $f(e_2) = e_1$ et $f(e_3) = e_1 + e_3$, la matrice de f dans la base (e_3, e_1, e_2) est la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, donc $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ sont semblables.

De la même manière : $f(e_1) = (0,0,0)$, $f(4e_2) = 4e_1$ et $f(2e_3) = 2e_1 + 2e_3$, donc la matrice de f dans la base $(e_1, 4e_2, 2e_3)$ est $\begin{pmatrix} 0 & 4 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, donc les matrices $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 0 & 4 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ sont semblables.

Définition (Trace d'un endomorphisme en dimension finie) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle. Pour tout $u \in \mathcal{L}(E)$, la trace de la matrice $\mathrm{Mat}_{\mathscr{B}}(u)$ ne dépend pas du choix de la base $\mathscr{B}$ de E choisie. On l'appelle *trace de u*, notée $\mathrm{tr}(u)$ ou $\mathrm{Tr}(u)$.

- **Linéarité :** Pour tous $u, v \in \mathcal{L}(E)$ et $\lambda, \mu \in \mathbb{K}$: $\mathrm{tr}(\lambda u + \mu v) = \lambda \mathrm{tr}(u) + \mu \mathrm{tr}(v)$.
- **Effet sur une composée :** Pour tous $u, v \in \mathcal{L}(E)$: $\mathrm{tr}(u \circ v) = \mathrm{tr}(v \circ u)$.

Démonstration Il nous suffit de justifier la définition. Or pour toutes bases $\mathscr{B}$ et $\mathscr{B}'$ de E, les matrices $\mathrm{Mat}_{\mathscr{B}}(u)$ et $\mathrm{Mat}_{\mathscr{B}'}(u)$ étant semblables, nous venons de voir qu'elles ont même trace.

Exemple Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et p un projecteur de E. Alors $\mathrm{tr}(p) = \mathrm{rg}(p)$.

Démonstration Nous avons vu dans un précédent exemple que si $\mathscr{B}$ est une base de E adaptée à la décomposition $E = \mathrm{Im}\, p \oplus \mathrm{Ker}\, p$ et si on pose $n = \dim E$ et $r = \mathrm{rg}(p)$, alors $\mathrm{Mat}_{\mathscr{B}}(p) = \begin{pmatrix} I_r & \\ & 0_{n-r} \end{pmatrix}$. A fortiori : $\mathrm{tr}(p) = \mathrm{tr}\big(\mathrm{Mat}_{\mathscr{B}}(p)\big) = \mathrm{tr}\begin{pmatrix} I_r & \\ & 0_{n-r} \end{pmatrix} = r = \mathrm{rg}(p)$.

3 Introduction à la diagonalisation d'un endomorphisme ou d'une matrice carrée

Les concepts de ce dernier paragraphe ne sont pas au programme de MPSI, il font partie du domaine de la *réduction* évoqué plus haut et vous les étudierez en deuxième année. Cependant, parce qu'ils apparaissent sous forme cachée dans de nombreux exercices de première année, je préfère vous les livrer en partie dès maintenant, cela ne pourra que vous aider à comprendre la logique des exercices en question.

3.1 Éléments propres d'un endomorphisme ou d'une matrice carrée

Définition (Valeurs propres, vecteurs propres, sous-espaces propres)

- **Cas des endomorphismes :** Soient $E \neq \{0_E\}$ un $\mathbb{K}$-espace vectoriel de dimension finie et $u \in \mathscr{L}(E)$. On appelle *valeur propre de u* tout scalaire $\lambda \in \mathbb{K}$ pour lequel $\mathrm{Ker}\big(u - \lambda \mathrm{Id}_E\big) \neq \{0_E\}$, i.e. pour lequel il existe un vecteur **NON NUL** $x \in E$ tel que $u(x) = \lambda x$.

  Un tel vecteur x est appelé un *vecteur propre de u (associé à la valeur propre λ)*. L'ensemble $\mathrm{Ker}\big(u - \lambda \mathrm{Id}_E\big)$ de ces vecteurs propres — avec le vecteur nul en plus — est quant à lui appelé le *sous-espace propre de u associé à la valeur propre λ*.

  Enfin, l'ensemble des valeurs propres de u est appelé le *spectre de u*.

- **Cas des matrices carrées :** Soit $A \in \mathscr{M}_n(\mathbb{K})$. On appelle *valeur propre de A* toute valeur propre de l'endomorphisme $\widehat{A}$ de $\mathbb{K}^n$ canoniquement associé à A, i.e. tout vecteur **NON NUL** $X \in \mathbb{K}^n$ pour lequel $AX = \lambda X$. Un tel vecteur X est appelé un *vecteur propre de A* et l'ensemble $\mathrm{Ker}(A - \lambda I_n)$ est appelé le *sous-espace propre de A associé à λ*. Enfin, le spectre de $\widehat{A}$ est aussi appelé le *spectre de A*.

Exemple La matrice $A = \begin{pmatrix} -1 & 3 & 3 \\ 3 & -1 & -3 \\ -3 & 3 & 5 \end{pmatrix}$ possède exactement deux valeurs propres, à savoir 2 et -1.

De plus : $\mathrm{Ker}(A - 2I_3) = \mathrm{Vect}\Big((1,1,0),(1,0,1)\Big)$ et $\mathrm{Ker}(A + I_3) = \mathrm{Vect}\Big((1,-1,1)\Big)$.

Démonstration Il s'agit seulement de résoudre pour tout $\lambda \in \mathbb{R}$ le système $AX = \lambda X$ d'inconnue $X \in \mathbb{R}^3$. On peut le faire matriciellement grâce à de simples opérations élémentaires sur les **LIGNES**. Fixons $\lambda \in \mathbb{R}$.

$$\mathrm{Ker}(A - \lambda I_3) = \mathrm{Ker}\begin{pmatrix} -(\lambda+1) & 3 & 3 \\ 3 & -(\lambda+1) & -3 \\ -3 & 3 & 5-\lambda \end{pmatrix} = \mathrm{Ker}\begin{pmatrix} -(\lambda+1) & 3 & 3 \\ 2-\lambda & 2-\lambda & 0 \\ -3 & 3 & 5-\lambda \end{pmatrix} \qquad L_2 \leftarrow L_1 + L_2.$$

Pour $\lambda = 2$, la deuxième équation sous-jacente disparaît :

$$\mathrm{Ker}(A - 2I_3) = \mathrm{Ker}\begin{pmatrix} -3 & 3 & 3 \\ -3 & 3 & 3 \end{pmatrix} = \Big\{(x,y,z) \in \mathbb{R}^3 \ \big| \quad x = y+z\Big\} = \Big\{(y+z,y,z) \ \big| \quad y,z \in \mathbb{R}\Big\} = \mathrm{Vect}\Big((1,1,0),(1,0,1)\Big).$$

Poursuivons nos calculs dans le cas où $\lambda \neq 2$:

$$\mathrm{Ker}(A - \lambda I_3) = \mathrm{Ker}\begin{pmatrix} -(\lambda+1) & 3 & 3 \\ 1 & 1 & 0 \\ -3 & 3 & 5-\lambda \end{pmatrix} = \mathrm{Ker}\begin{pmatrix} -(\lambda+4) & 0 & 3 \\ 1 & 1 & 0 \\ -6 & 0 & 5-\lambda \end{pmatrix} \quad \begin{matrix} L_1 \leftarrow L_1 - 3L_2 \\ L_3 \leftarrow L_3 - 3L_2 \end{matrix} \quad = \mathrm{Ker}\begin{pmatrix} -(\lambda+4) & 0 & 3 \\ 1 & 1 & 0 \\ (2-\lambda)(\lambda+1) & 0 & 0 \end{pmatrix} \quad L_3 \leftarrow 3L_3 - (5-\lambda)L_1.$$

Pour $\lambda = -1$, la troisième équation sous-jacente disparaît :

$$\mathrm{Ker}(A + I_3) = \mathrm{Ker}\begin{pmatrix} -3 & 0 & 3 \\ 1 & 1 & 0 \end{pmatrix} = \Big\{(x,y,z) \in \mathbb{R}^3 \ \big| \quad y = -x \quad \text{et} \quad z = x\Big\} = \mathrm{Vect}\Big((1,-1,1)\Big).$$

Enfin, si λ ne vaut ni 2 ni -1 : $\mathrm{Ker}(A - \lambda I_3) = \mathrm{Ker}\begin{pmatrix} -(\lambda+4) & 0 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \big\{(0,0,0)\big\}$ car la matrice obtenue est inversible.

Théorème (Familles de vecteurs propres) Soient E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et $u \in \mathscr{L}(E)$.

- **Finitude du spectre :** Le spectre de u est fini de cardinal inférieur ou égal à $\dim E$.
- **Liberté des familles de vecteurs propres :** Toute famille de vecteurs propres de u associée à des valeurs propres distinctes est libre. On obtient donc toujours une famille libre quand on concatène des bases associées à des sous-espaces propres distincts de u.

On dispose bien sûr d'un énoncé analogue sur le spectre et les sous-espaces propres d'une matrice carrée.

Démonstration Nous allons commencer par montrer de deux manières différentes que toute famille de vecteurs propres associée à des valeurs propres distinctes est libre.

- **Preuve n°1 à l'aide d'une matrice de Vandermonde :** Soient $\lambda_1, \ldots, \lambda_p$ des valeurs propres distinctes de u. Donnons-nous pour tout $i \in [\![1,p]\!]$ un vecteur propre x_i de u associé à la valeur propre λ_i et montrons que la famille $(x_1, \ldots, x_p)$ est libre. Soient $\alpha_1, \ldots, \alpha_p \in \mathbb{K}$ des scalaires pour lesquels $\alpha_1 x_1 + \ldots + \alpha_p x_p = 0_E$. Montrons que $\alpha_1 = \ldots = \alpha_p = 0_E$. Composons pour cela k fois par f la relation $\alpha_1 x_1 + \ldots + \alpha_p x_p = 0_E$ pour tout $k \in \mathbb{N}$: $\alpha_1 \lambda_1^k x_1 + \ldots + \alpha_p \lambda_p^k x_p = 0_E \qquad \bigstar_k$. Notons ensuite V la matrice de Vandermonde des scalaires $\lambda_1, \ldots, \lambda_n$: $V = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{pmatrix}$, inversible car $\lambda_1, \ldots, \lambda_p$ sont distincts. Les relations

$\bigstar_0, \ldots, \bigstar_{n-1}$ s'écrivent matriciellement : $\mathrm{Mat}_{\mathscr{B}}(\alpha_1 x_1, \ldots, \alpha_p x_p) \times V = 0$ dans toute base $\mathscr{B}$ de E, donc par inversibilité de V : $\mathrm{Mat}_{\mathscr{B}}(\alpha_1 x_1, \ldots, \alpha_p x_p) = 0$, i.e. $\alpha_1 x_1 = \ldots = \alpha_p x_p = 0_E$. Or $x_1, \ldots, x_p$ sont non nuls en tant que vecteurs propres, donc $\alpha_1 = \ldots = \alpha_p = 0$.

- **Preuve n°2 par récurrence sur le cardinal de la famille :**

  Initialisation : Par définition, tout vecteur propre est non nul, donc forme à lui seul une famille libre.

  Hérédité : Soit $p \in \mathbb{N}^*$. Faisons l'hypothèse que toute famille de p vecteurs propres de u associée à des valeurs propres distinctes est libre. Soit $(x_1, \ldots, x_{p+1})$ une famille de vecteurs propres de u associés à des valeurs propres distinctes $\lambda_1, \ldots, \lambda_{p+1}$. Montrons que $(x_1, \ldots, x_{p+1})$ est libre. Soient $\alpha_1, \ldots, \alpha_{p+1} \in \mathbb{K}$ des scalaires pour lesquels $\alpha_1 x_1 + \ldots + \alpha_{p+1} x_{p+1} = 0_E$. Composons par u : $\alpha_1 \lambda_1 x_1 + \ldots + \alpha_{p+1} \lambda_{p+1} x_{p+1} = 0_E$, puis tuons par soustraction le vecteur x_{p+1} : $\alpha_1 (\lambda_1 - \lambda_{p+1}) x_1 + \ldots + \alpha_p (\lambda_p - \lambda_{p+1}) x_p = 0_E$. La famille $(x_1, \ldots, x_p)$ étant libre par hypothèse de récurrence : $\alpha_i (\lambda_i - \lambda_{p+1}) = 0$ pour tout $i \in [\![1, p]\!]$, puis $\alpha_1 = \ldots = \alpha_p = 0$ car $\lambda_1, \ldots, \lambda_{p+1}$ sont distincts. Enfin, par non-nullité de x_{p+1} : $\alpha_{p+1} = 0$.

- À ce stade, le spectre de u est fini, car E étant de dimension finie, toute famille libre de E est finie. Nous pouvons donc noter $\lambda_1, \ldots, \lambda_p$ les valeurs propres de u et nous donner pour tout $i \in [\![1, p]\!]$ une base $\left(x_{i,1}, \ldots, x_{i,n_i}\right)$ de $\mathrm{Ker}\left(u - \lambda_i \mathrm{Id}_E\right)$. Nous allons montrer que la concaténation $\left(x_{1,1}, \ldots, x_{p,n_p}\right)$ de ces bases est libre. Soient $\lambda_{1,1}, \ldots, \lambda_{p,n_p} \in \mathbb{K}$ des scalaires pour lesquels $\lambda_{1,1} x_{1,1} + \ldots + \lambda_{p,n_p} x_{p,n_p} = 0_E$. Pour tout $i \in [\![1, p]\!]$, le vecteur $x_i = \lambda_{i,1} x_{i,1} + \ldots + \lambda_{i,n_i} x_{i,n_i}$ appartient à $\mathrm{Ker}\left(u - \lambda_i \mathrm{Id}_E\right)$, donc est soit nul, soit un vecteur propre de u associé à la valeur propre λ_i. Or $x_1 + \ldots + x_p = 0_E$, donc par liberté des familles de vecteurs propres associés à des valeurs propres distinctes : $x_1 = \ldots = x_p = 0_E$. Ainsi, pour tout $i \in [\![1, p]\!]$: $\lambda_{i,1} x_{i,1} + \ldots + \lambda_{i,n_i} x_{i,n_i} = 0_E$, donc par liberté de $\left(x_{i,1}, \ldots, x_{i,n_i}\right)$: $\lambda_{i,1} = \ldots = \lambda_{i,n_i} = 0$. ■

3.2 Diagonalisabilité d'un endomorphisme ou d'une matrice carrée

Définition-théorème (Endomorphisme/matrice carrée diagonalisable)

- **Cas des endomorphismes :** Soient E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et $u \in \mathscr{L}(E)$. On dit que u est *diagonalisable* si E possède une base constituée de vecteurs propres de u. Le cas échéant, si $\mathscr{B} = (x_1, \ldots, x_n)$ est une telle base et si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres associées : $\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.
- **Cas des matrices carrées :** Soit $A \in \mathscr{M}_n(\mathbb{K})$. On dit que A est *diagonalisable* (*sur* $\mathbb{K}$) si l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé est diagonalisable, i.e. si $\mathbb{K}^n$ possède une base constituée de vecteurs propres de A. Le cas échéant, si $(X_1, \ldots, X_n)$ est une telle base et si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres associées, alors $A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}$ où l'on a noté P la matrice de colonnes $X_1, \ldots, X_n$.

Démonstration Dans le cas matriciel, supposons A diagonalisable et notons $\widehat{A}$ l'endomorphisme de $\mathbb{K}^n$ canoniquement associé ainsi que $\mathscr{B}_n$ la base canonique de $\mathbb{K}^n$. Pour tout $i \in [\![1, n]\!]$: $\widehat{A}(X_i) = \lambda_i X_i$, donc la matrice de $\widehat{A}$ dans $(X_1, \ldots, X_n)$ est facile à écrire : $\mathrm{Mat}_{(X_1,\ldots,X_n)}\left(\widehat{A}\right) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Or $P = P_{\mathscr{B}_n}^{(X_1,\ldots,X_n)}$, donc par changement de base : $\mathrm{Mat}_{(X_1,\ldots,X_n)}\left(\widehat{A}\right) = P^{-1} \mathrm{Mat}_{\mathscr{B}_n}\left(\widehat{A}\right) P$, et finalement $A = P \,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\, P^{-1}$. ■

Exemple On pose $A = \begin{pmatrix} -1 & 3 & 3 \\ 3 & -1 & -3 \\ -3 & 3 & 5 \end{pmatrix}$. Pour tout $k \in \mathbb{N}$: $A^k = \begin{pmatrix} (-1)^k & a_k & a_k \\ a_k & (-1)^k & -a_k \\ -a_k & a_k & 2^k + a_k \end{pmatrix}$ avec $a_k = 2^k - (-1)^k$.

Démonstration **Idée de la preuve :** Parce que les puissances d'une matrice diagonale sont très faciles à calculer, nous allons commencer par montrer que A est diagonalisable.

- Nous avons calculé les valeurs propres et les sous-espaces propres de A dans un exemple précédent. En l'occurrence, pour $e_1 = (1, 1, 0)$, $e_2 = (1, 0, 1)$ et $e_3 = (1, -1, 1)$: $Ae_1 = 2e_1$, $Ae_2 = 2e_2$ et $Ae_3 = -e_3$. La famille (e_1, e_2, e_3) est libre car c'est une famille de vecteurs propres de A associés à des valeurs propres distinctes, c'est donc une base de $\mathbb{R}^3$ pour une raison de cardinal. Conclusion : A est diagonalisable.

  En outre, la matrice P de (e_1, e_2, e_3) dans la base canonique de $\mathbb{R}^3$ est inversible, et après calcul :
  $$P^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}, \qquad \text{puis } A = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1} \text{ par changement de base.}$$

- Finalement, pour tout $k \in \mathbb{N}$: $A^k = \left(P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1} \right)^k = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1} \times \ldots \times P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1}$
  $$= P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}^k P^{-1} = P \begin{pmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & (-1)^k \end{pmatrix} P^{-1} = \begin{pmatrix} (-1)^k & a_k & a_k \\ a_k & (-1)^k & -a_k \\ -a_k & a_k & 2^k + a_k \end{pmatrix} \quad \text{si on pose } a_k = 2^k - (-1)^k.$$