TECHNIQUES DE L'INGÉNIEUR

Réf. : **AF86 V1**

Calcul matriciel

Date de publication :
10 octobre 1998

Cet article est issu de : **Sciences fondamentales | Mathématiques**

par **Gérard DEBEAUMARCHÉ, Danièle LINO**

Calcul matriciel

par **Gérard DEBEAUMARCHÉ**
Ancien élève de l'École normale supérieure de Cachan
Professeur de mathématiques spéciales au lycée Clemenceau de Reims

et **Danièle LINO**
Ancienne élève de l'École normale supérieure de Sèvres
Professeur de mathématiques spéciales au lycée Clemenceau de Reims

De très nombreux problèmes issus des mathématiques ou de leurs applications conduisent à l'étude (et à la résolution) de systèmes d'équations linéaires. Le système (S) :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1p}\,x_p = b_1 \\ \quad\vdots \qquad\qquad\qquad \vdots \\ a_{n1}x_1 + \ldots + a_{np}\,x_p = b_n \end{cases}$$

(où les inconnues sont les nombres $x_1, ..., x_p$ et où les nombres a_{ij} et b_i sont donnés dans $\mathbb{K}$) se note aussi, sous forme matricielle :

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ou de façon plus ramassée :

$$AX = B$$

(cette égalité ne signifiant rien d'autre que les n égalités du système (S)).

L'exemple ci-après illustre cette situation.

Considérons l'équation différentielle $y'' + q(x)\, y = f(x)$, que l'on retrouve en résistance des matériaux avec des conditions aux limites de la forme $y(a) = A$, $y(b) = B$, les fonctions q et f étant continues sur [a, b]. Ce problème aux limites admet une solution unique y, dont on peut chercher des approximations de la façon suivante. On établit un maillage du segment [a, b] en subdivisant celui-ci en $n + 1$ sous-segments égaux dont les extrémités sont notées $a = x_0$, $x_1, ..., x_n, x_{n+1} = b$. On sait, d'après la formule de Taylor, que :

$$\begin{cases} y(x+h) = y(x) + hy'(x) + \dfrac{h^2}{2}\, y''(x) + 0(h^2) \\ y(x-h) = y(x) - hy'(x) + \dfrac{h^2}{2}\, y''(x) + 0(h^2) \end{cases}$$

Donc une valeur approchée de y''(x) est (par addition) :

$$y''(x) \approx \frac{y(x+h) - 2y(x) + y(x-h)}{h^2}.$$

Ici, on prendra donc $h = \dfrac{b-a}{n+1}$ et il vient :

$$y''(x_i) \approx \frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2}.$$

L'équation différentielle $y'' + q(x)\, y = f(x)$ donne ainsi, lorsqu'on l'écrit aux point $x_1, x_2, ..., x_n$ du maillage :

$$\frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2} + q(x_i)\, y(x_i) = f(x_i).$$

En fait, vu l'approximation faite sur y'', on n'obtient pas les nombres $y(x_i)$ mais des valeurs approchées y_i de ceux-ci. Et les équations précédentes s'écrivent matriciellement :

$$\begin{pmatrix} -2 + h^2 q(x_1) & 1 & & (0) \\ 1 & -2 + h^2 q(x_2) & 1 & \\ & \ddots & \ddots & 1 \\ (0) & & 1 & -2 + h^2 q(x_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} -A + h^2 f(x_1) \\ h^2 f(x_2) \\ \vdots \\ -B + h^2 f(x_n) \end{pmatrix}$$

On est ainsi ramené à résoudre un système d'équations linéaires (de grande taille dans les applications pratiques).

Dans cet article, on se propose d'étudier la notion de matrice afin d'être à même, notamment, d'étudier des systèmes d'équations linéaires du type précé-

dent, mais aussi des systèmes d'équations différentielles que l'on peut rencontrer en physique lorsque des oscillateurs sont couplés, ou en cinétique chimique, etc., et d'autres problèmes plus fins intervenant dans toutes les applications des mathématiques.

1. Matrices d'une application linéaire. Opérations sur les matrices

On se propose d'introduire la notion de matrice à partir de celle d'application linéaire, puis de définir les opérations sur les matrices à partir de celles sur les applications linéaires. On étudiera ensuite la structure de certains de ces ensembles de matrices.

Dans la suite, tous les espaces vectoriels sont considérés sur un même corps $\mathbb{K}$ et sont de dimension finie.

1.1 Matrice d'une application linéaire

Dans ce paragraphe, on suppose les espaces vectoriels $\boldsymbol{E}$ et $\boldsymbol{F}$ de dimension p et n, rapportés à deux bases $\mathfrak{B}_{\boldsymbol{E}} = (e_1, ..., e_p)$, $\mathfrak{B}_{\boldsymbol{F}} = (f_1, ..., f_n)$ fixées une fois pour toutes.

On rappelle que la donnée d'une application linéaire de $\boldsymbol{E}$ dans $\boldsymbol{F}$ équivaut à la donnée de l'image de la base $\mathfrak{B}_{\boldsymbol{E}} = (e_1, ..., e_p)$ de $\boldsymbol{E}$, c'est-à-dire à la donnée des composantes des vecteurs $(f(e_1), ..., f(e_p))$ sur la base $\mathfrak{B}_{\boldsymbol{F}} = (f_1, ..., f_n)$ de $\boldsymbol{F}$.

Soit une application linéaire u de $\boldsymbol{E}$ dans $\boldsymbol{F}$. On pose :

$$u(e_j) = \sum_{i=1}^{n} u_{i,j}\, f_i \quad \text{pour} \quad 1 \leqslant j \leqslant p\,.$$

La définition de l'application u équivaut donc à la donnée des $n\,p$ scalaires $u_{i,j}$ que l'on peut organiser sous forme d'un tableau de p colonnes à n éléments, la $j^{\text{ième}}$ colonne étant constituée des n composantes du vecteur $u(e_j)$ dans la base $(f_1, ..., f_n)$:

$$\begin{pmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,p} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,p} \\ \vdots & \vdots & & \vdots \\ u_{n,1} & u_{n,2} & \cdots & u_{n,p} \end{pmatrix}$$

Cela conduit aux définitions suivantes.

Définition 1. On appelle matrice de type (n, p) sur le corps $\mathbb{K}$, ou matrice à n lignes et p colonnes, tout tableau $U = (u_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ de $n\,p$ éléments de $\mathbb{K}$ ainsi disposés en p colonnes à n éléments ou en n lignes à p éléments.

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble de ces matrices de type (n, p) à coefficients dans $\mathbb{K}$.

Définition 2. A toute application linéaire u de $\boldsymbol{E}$ (rapporté à une base $\mathfrak{B}_{\boldsymbol{E}}\,(e_1, ..., e_p)$) dans $\boldsymbol{F}$ (rapporté à une base $\mathfrak{B}_{\boldsymbol{F}}\,(f_1, ..., f_n)$), on fait correspondre une unique matrice de type (n, p) appelée matrice de u et notée $M(u\,;\mathfrak{B}_{\boldsymbol{E}}, \mathfrak{B}_{\boldsymbol{F}})$, ou $M(u)$ s'il n'y a pas d'ambiguïté sur les bases.

On se propose maintenant d'exploiter la notion de matrice d'une application linéaire afin d'obtenir les coordonnées d'un vecteur $y = u(x)$ en fonction de celle du vecteur x.

Proposition 1.

Avec les notations précédentes, si X et Y désignent les matrices colonnes des composantes de deux vecteurs x et y appartenant respectivement à $\boldsymbol{E}$ et $\boldsymbol{F}$, la relation $y = u(x)$ équivaut à la relation $Y = M(u)\,X$.

Preuve ◊ Posons $y = \sum_{i=1}^{n} y_i\, f_i$ et $x = \sum_{j=1}^{p} x_j\, e_j$.

En posant $M(u) = (m_{i,j})$, on a l'égalité :

$$y = u(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{p} m_{i,j}\, x_j \right) f_i\,,$$

ce qui équivaut à :

$$y_i = \sum_{j=1}^{p} m_{i,j}\, x_j \quad \text{pour} \quad 1 \leqslant i \leqslant n\,,$$

c'est-à-dire à $Y = M(u)\,X$. ◊

1.2 Somme et produit par un scalaire

Dans ce paragraphe, on suppose les espaces vectoriels de dimension p et n rapportés à deux bases :

$$\mathfrak{B}_{\boldsymbol{E}} = (e_1, ..., e_p) \quad \text{et} \quad \mathfrak{B}_{\boldsymbol{F}} = (f_1, ..., f_n)\,.$$

On se propose de définir deux opérations sur $\mathcal{M}_{n,p}(\mathbb{K})$: la somme et le produit par un scalaire.

Soient u et v deux applications linéaires définies sur $\boldsymbol{E}$ à valeurs dans $\boldsymbol{F}$. On pose, pour $1 \leqslant j \leqslant p$:

$$u(e_j) = \sum_{i=1}^{n} u_{i,j} f_i \quad \text{et} \quad v(e_j) = \sum_{i=1}^{n} v_{i,j} f_i\,.$$

On a alors :

$$(u+v)(e_j) = \sum_{i=1}^{n} (u_{i,j} + v_{i,j}) f_i \quad \text{et} \quad \lambda u\,(e_j) = \sum_{i=1}^{n} (\lambda u_{i,j}) f_i\,.$$

Il en résulte que la matrice de l'application linéaire $(u + v)$ est la matrice $(u_{i,j} + v_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ et que la matrice de l'application linéaire λu est la matrice $(\lambda u_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$. Cela conduit à la définition suivante.

Définition 3. Soient deux matrices de type (n, p) :

$$U = (u_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} \text{ et } V = (v_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$$

On appelle :

— **somme** des matrices $U, \; V \in \mathcal{M}_{n,p}(\mathbb{K})$, la matrice $U + V \in \mathcal{M}_{n,p}(\mathbb{K})$ dont le terme général est $u_{ij} + v_{ij}$;

— **produit** par un scalaire λ de la matrice U la matrice $\lambda U \in \mathcal{M}_{n,p}(\mathbb{K})$ de terme général λu_{ij}.

Proposition 2.

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ muni de la somme et de la multiplication par un scalaire est un espace vectoriel de dimension np sur le corps $\mathbb{K}$, et l'application $u \to M(u)$ constitue un isomorphisme de $\mathcal{L}(\boldsymbol{E}, \boldsymbol{F})$ dans $\mathcal{M}_{n,p}(\mathbb{K})$.

On a donc $\forall (u, v) \in \mathcal{L}(\boldsymbol{E}, \boldsymbol{F})$ et $\forall \lambda \in \mathbb{K}$:

$$M(u + v) = M(u) + M(v)$$

$$M(\lambda u) = \lambda \, M(u).$$

Preuve ◊ On vérifie sans peine que $\mathcal{M}_{n,p}(\mathbb{K})$ est un espace vectoriel et que l'application $u \to M(u)$ est bijective d'après la définition 2.

De plus, par définition des opérations matricielles, on a :

$$M(u + v) = M(u) + M(v)$$

et

$$M(\lambda u) = \lambda \, M(u),$$

d'où le résultat donné. ◊

Définition 4. Pour tout entier $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant p$, on appelle **matrice élémentaire** $E_{i,j}$, toute matrice à n lignes et p colonnes dont tous les éléments sont nuls, sauf celui se trouvant sur la $i^{\text{ième}}$ ligne et le $j^{\text{ième}}$ colonne qui vaut 1.

L'ensemble des np matrices élémentaires $(E_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$. C'est la **base canonique** de $\mathcal{M}_{n,p}(\mathbb{K})$.

Pour toute matrice $M = (m_{i,j})$ appartenant à $\mathcal{M}_{n,p}(\mathbb{K})$, on a donc :

$$M = \sum_{i=1}^{n} \sum_{j=1}^{p} m_{i,j} \, E_{i,j}.$$

1.3 Produit de deux matrices

Dans ce paragraphe, on suppose les espaces vectoriels $\boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}$ de dimension p, q, n rapportés à trois bases $(e_1, ..., e_p)$, $(f_1, ..., f_q)$ et $(g_1, ..., g_n)$ fixées une fois pour toutes.

Soient deux applications linéaires u, définie sur $\boldsymbol{E}$ à valeurs dans $\boldsymbol{F}$, et v, définie sur $\boldsymbol{F}$ à valeurs dans $\boldsymbol{G}$. La composée $v \circ u$ est une application linéaire définie sur $\boldsymbol{E}$ à valeurs dans $\boldsymbol{G}$.

Étudions la matrice de $v \circ u$ en fonction des matrices :

$$V = (v_{i,k})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant k \leqslant q}}$$

et

$$U = (u_{k,j})_{\substack{1 \leqslant k \leqslant q \\ 1 \leqslant j \leqslant p}}$$

de v et u relativement aux bases définies.

On a :

$$u(e_j) = \sum_{k=1}^{q} u_{k,j} \, f_k \quad 1 \leqslant j \leqslant p \quad ; \quad v(f_k) = \sum_{i=1}^{n} v_{i,k} \, g_i \quad 1 \leqslant k \leqslant q.$$

Il en résulte :

$$v \circ u \, (e_j) = \sum_{i=1}^{n} \left(\sum_{k=1}^{q} v_{i,k} \; u_{k,j} \right) g_i \quad 1 \leqslant j \leqslant p.$$

On en déduit que la $k^{\text{ième}}$ composante du vecteur $v \circ u \, (e_j)$ dans la base $(g_1, g_2, ..., g_n)$ vaut $\sum_{k=1}^{q} v_{i,k} \; u_{k,j}$.

Cela conduit à la définition suivante.

Définition 5. Soient deux matrices :

$$V = (v_{i,k})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant k \leqslant q}} \text{ de type } (n, q)$$

et

$$U = (u_{k,j})_{\substack{1 \leqslant k \leqslant q \\ 1 \leqslant j \leqslant p}} \text{ de type } (q, p).$$

La matrice produit VU est la matrice de type (n, p) dont le terme général vaut :

$$\sum_{k=1}^{q} v_{i,k} \; u_{k,j} \text{ pour } 1 \leqslant i \leqslant n \text{ et } 1 \leqslant j \leqslant p.$$

Le produit VU de deux matrices U et V n'a de sens que si le nombre de colonnes de V est égal au nombre de lignes de U.

Proposition 3.

Pour tout couple d'applications linéaires $u \in \mathcal{L}(\boldsymbol{E}, \boldsymbol{F})$, $v \in \mathcal{L}(\boldsymbol{F}, \boldsymbol{G})$, on a :

$$M(v \circ u) = M(v) \times M(u).$$

De plus, si $\boldsymbol{E} = \boldsymbol{F}$, on a donc $M(u^n) = (M(u))^n$ pour tout entier naturel n.

Preuve. ◊ L'égalité $M(v \circ u) = M(v) \times M(u)$ résulte immédiatement de la définition du produit de deux matrices. ◊

Proposition 4.

Soient $U \in \mathcal{M}_{n,p}(\mathbb{K})$, $V \in \mathcal{M}_{p,q}(\mathbb{K})$ et $W \in \mathcal{M}_{q,r}(\mathbb{K})$. On a :

$$(U \cdot V) W = U (V \cdot W),$$

que l'on note $U \cdot V \cdot W$.

Preuve. ◊ Le résultat se déduit de l'associativité de la composition des applications linéaires en posant :

$$U = M(u) \; ; \; V = M(v) \; ; \; W = M(w). \quad ◊$$

1.4 Matrices carrées

Dans ce paragraphe, $\boldsymbol{E}$ est un espace vectoriel de dimension n rapporté à une base $(e_1, ..., e_n)$.

On s'intéresse aux endomorphismes de cet espace vectoriel $\boldsymbol{E}$. Les matrices de tels endomorphismes appartiennent à $\mathcal{M}_{n,n}(\mathbb{K})$, noté $\mathcal{M}_n(\mathbb{K})$.

Définition 6. On appelle matrice carrée toute matrice dont le nombre de colonnes est égal à celui des lignes.

On peut toujours définir le produit de deux matrices carrées de même taille. Cette opération définit une nouvelle loi interne sur l'ensemble des matrices carrées à n lignes et n colonnes.

Proposition 5.

L'ensemble $\mathcal{M}_n(\mathbb{K})$ muni des lois $+$, $\cdot$, $\times$ est une algèbre unitaire et non commutative, ce qui signifie que :

a) $(\mathcal{M}_n(\mathbb{K}), +, \cdot)$ *est un espace vectoriel ;*

b) $(\mathcal{M}_n(\mathbb{K}), +, \times)$ *est un anneau non commutatif* (*si* $n > 1$) *d'élément neutre* $I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$;

c) $\forall \lambda \in \mathbb{K}, \forall M, N \in \mathcal{M}_n(\mathbb{K})$:

$$\lambda(MN) = (\lambda M)N = M(\lambda N).$$

Preuve. ◊

a) Cette propriété a été établie en proposition 2.

b) On vérifie en effet que la multiplication est associative et distributive par rapport à l'addition.

c) On vérifie élémentairement cette propriété. ◊

On s'intéresse maintenant à l'étude des éléments inversibles de $\mathcal{M}_n(\mathbb{K})$.

Définition 7. On dit qu'une matrice carrée U à n lignes et n colonnes est inversible s'il existe une matrice V vérifiant :

$$UV = VU = I_n.$$

L'ensemble des matrices inversibles de l'algèbre $\mathcal{M}_n(\mathbb{K})$ est un groupe, appelé **groupe linéaire** et noté $GL_n(\mathbb{K})$.

- On fait les remarques suivantes. Une matrice carrée U est inversible si, et seulement si, c'est la matrice d'un automorphisme de ***E***.
- Si U est inversible, la matrice V est unique. On la note U^{-1}.
- L'égalité $UV = I_n$ implique $VU = I_n$, et donc $V = U^{-1}$.

Ce résultat, propre à la dimension finie, s'obtient en observant que l'application $M \to MU$ est alors injective, donc surjective, ce qui établit l'inversibilité à gauche de U. Et si l'on a $UV = I_n$, $WU = I_n$; on a donc en multipliant la première égalité par W à gauche :

$$V = W = U^{-1}.$$

2. Changement de bases

Étant donné une application linéaire f d'un espace vectoriel dans un autre, un objectif est de trouver des bases relativement auxquelles la matrice de f a une forme simple : diagonale, diagonale par blocs...

A cet effet, on est souvent conduit à effectuer des changements de bases dans un espace vectoriel.

C'est ce que l'on étudie particulièrement dans ce paragraphe, en explicitant notamment les formules de changement de bases pour les composantes d'un vecteur et les matrices d'une application linéaire.

2.1 Matrices de passage

Définition 8. On considère un espace vectoriel ***E***, de dimension n, et deux bases de celui-ci :

$$\mathfrak{B} = (e_1, \ldots, e_n)$$

et
$$\mathfrak{B}' = (e_1', \ldots, e_n').$$

On appelle matrice de passage de la base $\mathfrak{B}$ à la base $\mathfrak{B}'$ la matrice carrée d'ordre n dont les éléments de la $j^{\text{ième}}$ colonne sont les composantes de e_j' dans la base $\mathfrak{B} = (e_1, \ldots, e_n)$.

On note cette matrice $P_{\mathfrak{B} \to \mathfrak{B}'}$.

On a donc :

$$P_{\mathfrak{B} \to \mathfrak{B}'} = \begin{array}{c} \begin{pmatrix} p_{1,1} & \cdots & p_{1,n} \\ p_{2,1} & & p_{2,n} \\ \vdots & & \vdots \\ p_{n,1} & \cdots & p_{n,n} \end{pmatrix} \\ \begin{matrix} \uparrow & \quad & \uparrow \\ e_1' & & e_n' \end{matrix} \end{array} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}$$

On notera, au vu des définitions précédentes, que la matrice de passage de la base $\mathfrak{B}$ à la base $\mathfrak{B}'$ n'est autre que la matrice de l'application identité de ***E*** (rapporté à la base $\mathfrak{B}'$) dans ***E*** (rapporté à la base $\mathfrak{B}$).

Proposition 6.

*Soient $\mathfrak{B}$ et $\mathfrak{B}'$ deux bases de l'espace vectoriel **E**. Alors la matrice de passage de la base $\mathfrak{B}'$ à la base $\mathfrak{B}$ est l'inverse de la matrice de passage de la base $\mathfrak{B}$ à la base $\mathfrak{B}'$; autrement dit, on a* :

$$P_{\mathfrak{B}' \to \mathfrak{B}} = (P_{\mathfrak{B} \to \mathfrak{B}'})^{-1}.$$

Preuve. ◊

La matrice de passage de $\mathfrak{B}$ à $\mathfrak{B}'$ est en effet inversible, puisqu'il s'agit de la matrice de l'application identité comme on l'a indiqué dans la remarque précédente.

De plus $P_{\mathfrak{B}' \to \mathfrak{B}} P_{\mathfrak{B} \to \mathfrak{B}'} = I_n$ puisque :

$$P_{\mathfrak{B}' \to \mathfrak{B}} P_{\mathfrak{B} \to \mathfrak{B}'} = \text{Mat}(\text{Id}, \mathfrak{B}, \mathfrak{B}')\ \text{Mat}(\text{Id}, \mathfrak{B}', \mathfrak{B})$$
$$= \text{Mat}(\text{Id}, \mathfrak{B}, \mathfrak{B}) = I_n.$$ ◊

Proposition 7.

*Soient $\mathfrak{B}$ et $\mathfrak{B}'$ deux bases de l'espace vectoriel **E**. Alors les matrices colonnes X et X' des composantes d'un vecteur x de **E** vérifient l'égalité* :

$$X = PX' \quad \text{où} \quad P = P_{\mathfrak{B} \to \mathfrak{B}'}.$$

Preuve. ◊ Avec les notations précédentes on a :

— d'une part :

$$x = \sum_{i=1}^{n} x_i e_i$$

— d'autre part :

$$x = \sum_{j=1}^{n} x'_j \, e'_j$$

mais il résulte de la définition de la matrice de passage que :

$$e'_j = \sum_{i=1}^{n} p_{i,j} \, e_i$$

et donc que :

$$x = \sum_{j=1}^{n} x'_j \, e'_j = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{i,j} \, x'_j \right) e_i = \sum_{i=1}^{n} x_i \, e_i$$

ou encore que $x_i = \sum_{j=1}^{n} p_{i,j} \, x'_j$. ◊

2.2 Matrices équivalentes. Matrices semblables

Proposition 8.

*On considère une application linéaire f d'un espace vectoriel **E** dans un espace vectoriel **F**. On désigne par :*

— $\mathfrak{B}_{\boldsymbol{E}}$ et $\mathfrak{B}'_{\boldsymbol{E}}$ *deux bases de **E**, et l'on note P la matrice de passage de* $\mathfrak{B}_{\boldsymbol{E}}$ à $\mathfrak{B}'_{\boldsymbol{E}}$;

— $\mathfrak{B}_{\boldsymbol{F}}$ et $\mathfrak{B}'_{\boldsymbol{F}}$ *deux bases de **F**, et l'on note Q la matrice de passage de* $\mathfrak{B}_{\boldsymbol{F}}$ et $\mathfrak{B}'_{\boldsymbol{F}}$;

— *M* et *M'* *les matrices de f relatives aux bases* $\mathfrak{B}_{\boldsymbol{E}}$ et $\mathfrak{B}_{\boldsymbol{F}}$ *d'une part,* $\mathfrak{B}_{\boldsymbol{E}}'$ et $\mathfrak{B}'_{\boldsymbol{F}}$ d'autre part.

On a l'égalité :

$$M' = Q^{-1} MP .$$

Preuve. ◊ Désignons par X, X' les matrices colonnes d'un vecteur x de $\boldsymbol{E}$ et par Y, Y' les matrices colonnes d'un vecteur y de $\boldsymbol{F}$.

On a, d'une part :

$$X = PX' \quad \text{et} \quad Y = QY' .$$

On a, d'autre part :

$$Y = MX \ \text{ et } \ Y' = M'X' .$$

Par conséquent, on dispose des égalités :

$$Y' = Q^{-1} MP\, X' \ \text{ et } \ Y' = M'X' .$$

Il en résulte que $M' = Q^{-1} MP$ en donnant à X' les valeurs

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

ce qui prouve que les vecteurs colonnes des deux matrices M' et $Q^{-1} MP$ sont égales. p

En particulier si f est un endomorphisme, on obtient le résultat suivant.

Proposition 9.

*Soit f un endomorphisme de **E**.*

On désigne par :

— $\mathfrak{B}_{\boldsymbol{E}}$ et $\mathfrak{B}'_{\boldsymbol{E}}$ *deux bases de **E** et l'on note P la matrice de passage de* $\mathfrak{B}_{\boldsymbol{E}}$ à $\mathfrak{B}'_{\boldsymbol{E}}$;

— *M et M' les matrices de f relatives aux bases* $\mathfrak{B}_{\boldsymbol{E}}$ et $\mathfrak{B}'_{\boldsymbol{E}}$.

On a l'égalité :

$$M' = P^{-1} MP .$$

Proposition 10.

Deux matrices M et M' appartenant à $\mathcal{M}_{n,p}(\mathbb{K})$ *sont les matrices d'une même application linéaire f d'un espace de dimension p dans un espace de dimension n si, et seulement si :*

$$\exists P \in GL_p(\mathbb{K}), \exists Q \in GL_n(\mathbb{K}) \text{ tel que } M' = Q^{-1} MP .$$

Preuve. ◊ Simple conséquence de la proposition 8. ◊

Définition 9. Deux matrices M et M' de $\mathcal{M}_{n,p}(\mathbb{K})$ sont dites **équivalentes** si elles sont les matrices d'une même application linéaire d'un espace de dimension p dans un espace de dimension n, c'est-à-dire si :

$$\exists P \in GL_p(\mathbb{K}), \exists Q \in GL_n(\mathbb{K}) \text{ tel que } M' = Q^{-1} MP .$$

Il s'agit clairement d'une relation réflexive, symétrique et transitive, donc d'une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Proposition 11.

Deux matrices M et M' appartenant à $\mathcal{M}_n(\mathbb{K})$ *sont les matrices d'un même endomorphisme f d'un espace de dimension n si, et seulement si :*

$$\exists P \in GL_p(\mathbb{K}), \text{ tel que } M' = P^{-1} MP .$$

Preuve. ◊ Simple conséquence de la proposition 9. ◊

Définition 10. Deux matrices M et M' de $\mathcal{M}_n(\mathbb{K})$ sont dites **semblables** si elles sont les matrices d'un même endomorphisme d'un espace de dimension n, c'est-à-dire si :

$$\exists P \in GL_n(\mathbb{K}), \text{ tel que } M' = P^{-1} MP .$$

Il s'agit clairement d'une relation réflexive, symétrique et transitive, donc d'une relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$.

3. Rang d'une matrice

Définition 11. On appelle rang d'une matrice M de $\mathcal{M}_{n,p}(\mathbb{K})$ le rang de la famille de ses p vecteurs colonnes dans $\mathbb{K}^n$.

Il s'agit donc de la dimension du sous-espace vectoriel de $\mathbb{K}^n$ engendré par ses p vecteurs colonnes.

Proposition 12.

*Soit f une application linéaire de **E** vers **F** dont la matrice relativement à un couple de bases* $\mathfrak{B}, \mathfrak{B}'$ *est M. Alors, le rang de f est égal au rang de M. Autrement dit, la dimension de l'image de f est égale au rang d'une quelconque de ses matrices.*

Preuve. ◊ Soit $\mathfrak{B}$ une base de $\boldsymbol{E}$. Le rang de f est la dimension de son image, c'est-à-dire le rang de la famille des p vecteurs images par f d'une base de $\boldsymbol{E}$, $\mathfrak{B}$ par exemple. Le rang de la matrice M est le rang des colonnes de M.

L'égalité du rang de f et du rang de M est une conséquence immédiate du lemme suivant. ◊

Lemme 1. Soient $\boldsymbol{E}$ et $\boldsymbol{E}'$ deux espaces de même dimension n rapportés à deux bases :

$$\mathfrak{B} = (e_1, \ldots, e_n), \mathfrak{B}' = (e'_1, \ldots, e'_n).$$

Si $(v_1, \ldots, v_p)$ et $(v'_1, \ldots, v'_p)$ sont deux familles de vecteurs de $\boldsymbol{E}$ et $\boldsymbol{E}'$ dont les composantes sont respectivement les mêmes dans les bases $\mathfrak{B}$ et $\mathfrak{B}'$, alors ces deux familles ont même rang.

Preuve. ◊ Soit f l'isomorphisme de $\boldsymbol{E}$ dans $\boldsymbol{E}'$ transformant la base $\mathfrak{B}$ en la base $\mathfrak{B}'$ et donc les vecteurs $(v_1, \ldots, v_p)$ en les vecteurs $(v'_1, \ldots, v'_p)$. Alors f induit une application linéaire évidemment bijective de Vect $(v_1, v_2, \ldots, v_p)$ sur Vect $(v'_1, v'_2, \ldots, v'_p)$ et donc les deux familles $(v_1, \ldots, v_p)$ et $(v'_1, \ldots, v'_p)$ ont même rang. ◊

4. Matrices équivalentes

Dans ce paragraphe, on se propose de regrouper en classes les matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ susceptibles de représenter une même application linéaire par rapport à différentes bases de ces espaces. On établira notamment que deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ appartiennent à une même classe si, et seulement si, elles ont même rang.

Proposition 13.

Soit f une application linéaire de $\boldsymbol{E}$ *(de dimension p) dans* $\boldsymbol{F}$ *(de dimension n). Elle est de rang r si, et seulement si, il existe des bases de* $\boldsymbol{E}$ *et de* $\boldsymbol{F}$ *relativement auxquelles la matrice de f est :*

$$J_r = \left(\begin{array}{ccc|c} 1 & & (0) & \\ & \ddots & & (0) \\ (0) & & 1 & \\ \hline (0) & & & (0) \end{array}\right)$$

Preuve. ◊ Prouvons que si f est de rang r, il existe des bases de $\boldsymbol{E}$ et de $\boldsymbol{F}$ telles que la matrice de f relativement à ces bases est J_r.

Il résulte de théorème du rang (cf. article [AF 85] *Algèbre linéaire*) que le noyau de f est de dimension $p - r$. On choisit une base $(e_{r+1}, \ldots, e_p)$ du noyau de f que l'on complète en une base $(e_1, \ldots, e_r, e_{r+1}, \ldots, e_p)$ de $\boldsymbol{E}$. La famille $(f(e_1), \ldots, f(e_r))$ engendre clairement l'image et est, de plus, libre, car si $\sum_{i=1}^{r} \lambda_i f(e_i) = 0$, alors $\sum_{i=1}^{r} \lambda_i e_i$ est élément du noyau et donc élément de :

$$\text{Vect}(e_1, \ldots, e_r) \cap \text{Vect}(e_{r+1}, \ldots, e_p) = \{0\},$$

ce qui implique $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$.

On peut donc compléter cette famille $(f(e_1), \ldots, f(e_r))$ en une base de $\boldsymbol{F}$ et il est alors immédiat de constater que la matrice de f relativement aux bases ainsi construites est J_r.

Puisque le rang d'une application linéaire est le rang de sa matrice, le rang de f est le rang de J_r et donc vaut r. ◊

Proposition 14.

Soient M et M′ deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.

Alors M est équivalente à M′ si, et seulement si, elles ont même rang.

Preuve. ◊ Si M et M' sont équivalentes, on sait qu'elle représentent une même application linéaire relativement à deux couples de bases et, donc, que leur rang commun est celui d'une telle application linéaire.

Réciproquement, si M et M' ont même rang r, elles sont équivalentes à J_r et par transitivité, équivalentes entre elles. ◊

L'objectif suivant est de classifier les matrices semblables (comme on l'a fait pour les matrices équivalentes à l'aide du rang).

Il s'agit là d'un problème nettement plus difficile qui conduit à la théorie de la réduction des endomorphismes ou des matrices carrées. Il est facile de constater que deux matrices semblables ont clairement même rang, même déterminant, mêmes valeurs propres. Ces conditions ne sont toutefois pas suffisantes pour affirmer la similitude de deux matrices. A titre d'exemple, dans le cas très particulier des matrices de type (1,1), on vérifiera que :

— il y a deux classes d'équivalence de matrices (1,1) : la matrice (0) et toutes les autres qui sont équivalentes ;

— il y a autant de classes de similitude de matrices (1,1) que de nombres réels, puisque deux matrices (a) et (b) sont semblables si, et seulement si, il existe p non nul tel que :

$$(b) = (p)^{-1}(a)(p) ;$$

autrement dit si $a = b$.

5. Algorithme du pivot de Gauss. Conséquences

On expose, dans ce paragraphe, l'algorithme du pivot de Gauss. Celui-ci permet :

— d'une part, d'avoir un algorithme simple et numériquement efficace de résolution des systèmes et d'inversion des matrices carrées ;

— d'autre part, de préciser des générateurs simples du groupe linéaire.

5.1 Manipulations élémentaires sur les matrices

Dans ce paragraphe, on désigne par $M = (m_{i,j})$ une matrice appartenant à $\mathcal{M}_{n,p}(\mathbb{K})$. Nous décrivons maintenant ce qu'il est convenu d'appeler manipulations élémentaires sur les lignes de cette matrice.

■ **Ajout à une ligne d'une autre ligne multipliée par un scalaire** λ

Ajouter à la première ligne λ fois la seconde s'interprète matriciellement de la façon suivante :

$$\begin{pmatrix} 1 & \lambda & \cdots & \\ & & (0) & \\ 1 & 0 & & \\ & \ddots & \ddots & 0 \\ (0) & & & 1 \end{pmatrix} \begin{pmatrix} m_{1,1} & \cdots & m_{1,p} \\ & & \\ m_{2,1} & \cdots & m_{2,p} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,p} \end{pmatrix} = \begin{pmatrix} m_{1,1} + \lambda m_{2,1} & \cdots & m_{1,p} + \lambda m_{2,p} \\ & & \\ m_{2,1} & \cdots & m_{2,p} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,p} \end{pmatrix}$$

Plus généralement la prémultiplication de la matrice M par la **matrice** dite de **translation**

$$T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$$

(où $E_{i,j}$ désigne la matrice de la base canonique où l'élément 1 figure en position i, j) donne la matrice obtenue en ajoutant à la $i^{\text{ème}}$ ligne de M λ fois sa $j^{\text{ième}}$ ligne.

■ **Multiplication d'une ligne par un scalaire λ non nul**

Multiplier la première ligne par un scalaire λ s'interprète matriciellement de la façon suivante :

$$\begin{pmatrix} \lambda & & & \\ & 1 & & (0) \\ & & \ddots & \\ (0) & & & 1 \end{pmatrix}\begin{pmatrix} m_{1,1} & \dots & m_{1,p} \\ m_{2,1} & \dots & m_{2,p} \\ \vdots & & \vdots \\ m_{n,1} & \dots & m_{n,p} \end{pmatrix} = \begin{pmatrix} \lambda m_{1,1} & \dots & \lambda m_{1,p} \\ m_{2,1} & \dots & m_{2,p} \\ \vdots & & \vdots \\ m_{n,1} & \dots & m_{n,p} \end{pmatrix}$$

Plus généralement la prémultiplication de la matrice M par la **matrice** dite **de dilatation ou d'affinité** Diag $(1, \dots, 1, \lambda, 1, \dots, 1)$ (cette notation représente la matrice diagonale de $\mathcal{M}_n(\mathbb{K})$ dont les éléments diagonaux sont λ en $i^{\text{ième}}$ position et 1 ailleurs), donne la matrice obtenue en multipliant la $i^{\text{ième}}$ ligne de M par λ.

■ **Échange de deux lignes**

Échanger les deux premières lignes de la matrice M s'interprète matriciellement de la façon suivante :

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \dots & \dots & & 0 & 1 \end{pmatrix}\begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,p} \\ m_{2,1} & m_{2,2} & \dots & m_{2,p} \\ \vdots & & & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,p} \end{pmatrix} = \begin{pmatrix} m_{2,1} & m_{2,2} & \dots & m_{2,p} \\ m_{1,1} & m_{1,2} & \dots & m_{1,p} \\ \vdots & & & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,p} \end{pmatrix}$$

Plus généralement, on peut obtenir l'échange de deux lignes quelconques de M par prémultiplication par une certaine matrice.

En fait, cette dernière manipulation élémentaire peut se ramener aux deux précédentes.

Ainsi, l'échange des deux premières lignes que nous notons L_1 et L_2 peut se réaliser comme suit :

— à L_1 on ajoute L_2 ; les deux premières lignes sont alors $L_1 + L_2$ et L_2 ;

— à la seconde ligne, on retranche la première ; les deux premières lignes sont $L_1 + L_2$ et $-L_1$;.

— à la première on ajoute la seconde ; les deux premières lignes sont L_2 et $-L_1$;

— on multiplie la deuxième ligne par -1 ; les deux premières lignes sont L_2 et L_1.

Cette suite de trois transvections et d'une dilatation a effectivement permis l'échange des deux lignes.

Proposition 15.

Les manipulations élémentaires sur les lignes d'une matrice ne modifient pas le rang de celle-ci.

Preuve. ◊ On sait que deux matrices équivalentes M et $M' = Q^{-1} MP$ (avec Q et P inversibles) ont même rang. On rappelle que M est une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. Il suffit de choisir dans ce qui précède $Q^{-1} = T_{i,j}(\lambda)$ ou Diag $(1, \dots, 1, \lambda, 1, \dots, 1)$ et $P = I_p$ pour constater que chacune des manipulations élémentaires précédentes a laissé invariant le rang de la matrice. ◊

Proposition 16.

Les inverses des matrices de transvections $T_{i,j}(\lambda)$ et de dilatation Diag$(1, \dots, 1, \lambda, 1, \dots, 1)$ *sont les matrices* :

$$T_{i,j}(-\lambda) \text{ et Diag}\left(1, \dots, 1, \frac{1}{\lambda}, 1, \dots, 1\right).$$

Preuve. ◊ Vérification immédiate. ◊

Tout ce qui a été dit sur les lignes vaut également pour les colonnes, à condition de remplacer les prémultiplications par des postmultiplications.

5.2 Diverses formes de l'algorithme du pivot de Gauss

5.2.1 Résolution d'un système de *n* équations à *n* inconnues

On résoud ce système par mise à la forme triangulaire de la matrice M du système $MX = B$.

On se propose, à l'aide de manipulations élémentaires sur les lignes, de ramener le système d'équations $MX = B$ à un système $M'X = B'$ où la matrice M' est triangulaire. La résolution d'un système triangulaire de taille n étant aisée et s'effectuant en n^2 opérations, le problème est ainsi résolu.

On décrit maintenant la suite finie de manipulations élémentaires sur les lignes permettant de passer du système $MX = B$ au système $M'X = B'$.

■ **Étape 1.**

Par échange de lignes, on amène en position (1,1) un élément non nul, appelé premier pivot, et si possible de module maximal, parmi les éléments de la première colonne de M. Si tous les éléments de cette première colonne sont nuls, le système n'a pas de solution unique et l'on arrête l'algorithme. Sinon, le choix d'un élément maximal en module permet de minimiser les erreurs numériques lors de la division par celui-ci.

En retranchant un certain nombre de fois la première ligne aux suivantes, on obtient donc un système de la forme $M^{(1)} X = B^{(1)}$ où $B^{(1)}$ est obtenue à partir de B par les mêmes manipulations élémentaires que celles faites sur M :

$$\begin{pmatrix} m_{1,1}^{(1)} & m_{1,2}^{(1)} & \dots & m_{1,n}^{(1)} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & m_{n,2}^{(1)} & \dots & m_{n,n}^{(1)} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ \vdots \\ \vdots \\ b_n^{(1)} \end{pmatrix}$$

Naturellement, les matrices $M^{(1)}$ et $B^{(1)}$ pourraient s'obtenir à partir de M et de B à l'aide de prémultiplications par les matrices des manipulations élémentaires effectuées.

■ **Étape 2.**

Par échange éventuel de lignes, on amène en position (2,2) un élément non nul, appelé deuxième pivot, et si possible de module maximal, parmi les éléments $m_{2,2}^{(1)}, m_{3,2}^{(1)}, \dots, m_{n,2}^{(1)}$ de la deuxième colonne de $M^{(1)}$. Si tous ces éléments sont nuls, le système n'a pas de solution unique et l'on arrête l'algorithme.

En retranchant un certain nombre de fois la deuxième ligne aux suivantes, on obtient donc un système de la forme $M^{(2)} X = B^{(2)}$ où $B^{(2)}$ est obtenu à partir de $B^{(1)}$ par les mêmes manipulations élémentaires que celles faites sur $M^{(1)}$:

$$\begin{pmatrix} m_{1,1}^{(1)} & m_{1,2}^{(1)} & \dots & m_{1,n}^{(1)} \\ 0 & m_{2,2}^{(2)} & & m_{2,n}^{(2)} \\ \vdots & 0 & & \vdots \\ 0 & 0 & \dots & m_{n,n}^{(2)} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

■ En poursuivant ainsi, ou bien le système n'admet pas de solution unique, ou bien on est arrivé à un système de la forme $M^{(n)} X = B^{(n)}$ où la matrice $M^{(n)}$ est triangulaire, ses éléments diagonaux étant les pivots successifs de l'algorithme.

■ **Décompte du nombre d'opérations**

A l'étape 1, pour $2 \leqslant i \leqslant n$, on retranche la première ligne multipliée par $\dfrac{a_{i,1}}{a_{1,1}}$ à la $i^{\text{ième}}$ ligne ; on effectue donc une division, n multiplications et n soustractions, soit $2n+1$ opérations pour chacune des lignes soit en tout $(2n+1)(n-1)$ opérations.

De même, à l'étape 2, on a effectué $(2n-1)(n-2)$ opérations.

De façon plus générale, à l'étape k, on fait $(2n-2k+3)(n-k)$ opérations.

5.2.2 Inversibilité et calcul de l'inverse d'une matrice carrée

On décrit ici une variante de l'algorithme précédent permettant de détecter l'inversibilité éventuelle d'une matrice carrée et de calculer son inverse. Décrivons maintenant la suite des opérations effectuées sur les lignes (on pourrait aussi procéder sur les colonnes).

■ **Étape 1.**

Par échange de lignes, on amène en position (1,1) un élément non nul, appelé premier pivot, et si possible de module maximal, parmi les éléments de la première colonne de M. Si la première colonne de M est nulle on arrête l'algorithme, la matrice M n'étant pas inversible.

On divise ensuite la première ligne par l'élément en position (1,1) et l'on retire aux $n-1$ dernières lignes un certain nombre de fois la première de manière à y faire apparaître des zéros.

A la fin de cette première étape, on obtient la matrice suivante :

$$M^{(1)} = \begin{pmatrix} 1 & m_{1,2}^{(1)} & \dots & m_{1,n}^{(1)} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & m_{n,2}^{(1)} & \dots & m_{n,n}^{(1)} \end{pmatrix}$$

La matrice $M^{(1)}$ pourrait s'obtenir à partir de la matrice M à l'aide de prémultiplications par les matrices des manipulations élémentaires effectuées.

■ **Étape 2.**

Par échange éventuel de lignes parmi les $n-1$ dernières, on amène en position (2,2) un élément non nul appelé deuxième pivot, et si possible de module maximal, parmi les éléments $m_{2,2}^{(1)}, m_{3,2}^{(1)}, \dots, m_{n,2}^{(1)}$ de la deuxième colonne de $M^{(1)}$. Si ces éléments sont nuls, la matrice $M^{(1)}$ (ainsi que la matrice M) n'est pas inversible et on arrête l'algorithme.

On divise ensuite la deuxième ligne par l'élément en position (2,2), puis on retire aux $(n-1)$ autres lignes un certain nombre de fois la deuxième ligne de manière à y faire apparaître des zéros.

A la fin de cette deuxième étape, on obtient la matrice suivante :

$$M^{(2)} = \begin{pmatrix} 1 & 0 & m_{1,3}^{(2)} & \dots & m_{1,n}^{(2)} \\ 0 & 1 & \vdots & & \vdots \\ \vdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & m_{n,3}^{(2)} & \dots & m_{n,n}^{(2)} \end{pmatrix}$$

La matrice $M^{(2)}$ pourrait se déduire de la matrice $M^{(1)}$ (donc de la matrice M) à l'aide de prémultiplications par les matrices des manipulations élémentaires effectuées.

■ En poursuivant ainsi, on obtient, dans le cas où M est inversible, la matrice $M^{(n)} = I_n$ et on a l'égalité :

$$E_1\, E_2\, \dots\, E_p\, M = I_n$$

où les matrices $E_1 \dots E_p$ sont les matrices des manipulations élémentaires effectuées au cours des n étapes successives. De l'égalité précédente, on déduit l'écriture de M^{-1} :

$$M^{-1} = E_1\, E_2\, \dots\, E_p = E_1\, E_2\, \dots\, E_p\, I_n\,.$$

Ce qui signifie que si l'on fait sur I_n les mêmes opérations que celles ayant été faites sur M (pour ramener M à I_n), alors on obtient M^{-1}.

Pour inverser M par cet algorithme, on effectue encore un nombre d'opérations en $0(n^3)$ (c'est-à-dire majoré par Cn^3).

Exemple d'inversion de $M = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix}$

On effectue à droite, sur la matrice I_3 simultanément les mêmes opérations que sur M.

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1 \end{array} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$

$$L_3 \leftarrow L_3 - L_2 \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

$$L_1 \leftarrow L_1 + \frac{1}{2}\, L_3 \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \begin{pmatrix} 1/2 & -1/2 & 1/2 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

$$L_3 \leftarrow -\frac{1}{2}\, L_3 \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1/2 & -1/2 & 1/2 \\ -2 & 1 & 0 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} = M^{-1}$$

5.2.3 Générateurs du groupe linéaire

On se propose maintenant de démontrer que toute matrice inversible peut s'écrire comme le produit de matrices de transvections et de dilatations.

Reprenons l'algorithme de calcul de l'inverse d'une matrice. Si M est une matrice inversible, on peut, par prémultiplication par des matrices de transvections et de dilatations, écrire que :

$$E_1\, \dots\, E_p\, M = I$$

ou encore $$M = E_p^{-1}\, \dots\, E_1^{-1}.$$

Comme l'inverse d'une matrice de dilatation est une matrice de dilatation et l'inverse d'une matrice de transvection une matrice de transvection, alors toute matrice inversible est un produit de matrices de transvection et de dilatation. Cela nous permet d'énoncer la proposition suivante.

Proposition 17.

Le groupe des matrices inversibles $GL_n(\mathbb{K})$ *est engendré par l'ensemble des matrices de dilatations, c'est-à-dire de la forme* diag(1, ..., 1, λ, 1, ..., 1) *où* $\lambda \in \mathbb{K}^*$ *, et des matrices de transvection, c'est-à-dire des matrices de la forme* $I_n + \lambda\, E_{i,j}$ où $i \neq j$ et $\lambda \in \mathbb{K}$.

6. Déterminants

6.1 Généralités

On se propose de définir une application permettant d'étudier la liberté d'une famille de n vecteurs dans un espace de dimension n et de caractériser par exemple les bases, les matrices inversibles, etc. A cet effet, on cherche une application $\phi : \boldsymbol{E}^n \to \mathbb{K}$ s'annulant sur les familles liées.

L'application ϕ pourrait donc vérifier les propriétés suivantes :

① ϕ doit être nulle si deux vecteurs égaux figurent dans la famille, c'est-à-dire si :

$$\forall (V_1, \dots, V_n) \in E^n \quad (V_i = V_j \text{ avec } i \neq j) \Rightarrow (\phi(V_1, \dots, V_n) = 0) \;;$$

② ϕ doit être nulle lorsqu'un des vecteurs est nul.

Cela est réalisé si ϕ est linéaire en V_1, en V_2, ..., en V_n, c'est-à-dire si ϕ est n-linéaire.

Étudions donc les formes n-linéaires alternées sur un espace de dimension n.

Définition 12. Soit $\boldsymbol{E}$ un $\mathbb{K}$-espace vectoriel de dimension n. On appelle forme n-linéaire sur $\boldsymbol{E}$ toute application $\phi : E^n \to \mathbb{K}$ vérifiant :

$$E \to \mathbb{K} \;;\; V_1 \to \phi(V_1, V_2, \dots, V_n) \text{ est linéaire ;}$$
$$\vdots$$
$$E \to \mathbb{K} \;;\; V_n \to \phi(V_1, V_2, \dots, V_n) \text{ est linéaire.}$$

On dit que la forme n-linéaire ϕ est, pour $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$ et $i \neq j$:

— **alternée** si :

$$\forall (V_1, \dots, V_n) \in E^n \quad (V_i = V_j) \Rightarrow \phi(V_1, \dots, V_n) = 0 \;;$$

— **antisymétrique** si :

$$\forall (V_1, \dots, V_n) \in E^n \Rightarrow$$
$$\phi(V_1, \dots, V_i, \dots, V_j, \dots) = -\phi(V_1, \dots, V_j, \dots, V_i, \dots).$$

On prouve que sur un $\mathbb{K}$-espace vectoriel (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), ϕ est une forme n-linéaire alternée si, et seulement si, elle est antisymétrique.

Rappel.

On appelle **permutation** toute bijection d'un ensemble $\boldsymbol{E}$ sur lui-même. L'ensemble des permutations de l'ensemble $\{1, 2, \dots, n\}$ sur lui-même est noté σ_n.

La composée de deux permutations est appelé **produit**.

Parmi les permutations, celles échangeant deux éléments, i et j par exemple, sont appelées **transpositions**.

On démontre que toute permutation σ de σ_n se décompose en un produit de p transpositions. Cette décomposition n'est pas unique, mais si l'on décompose une permutation en un produit de p transpositions ou de q transpositions, les entiers p et q ont même parité.

Compte tenu de cette propriété, si une permutation σ de σ_n se décompose en un produit de p transpositions, on définit la **signature** de σ comme l'entier $(-1)^p$ que l'on note $\varepsilon(\sigma)$. Par exemple, la signature d'une transposition vaut -1, la signature du produit de deux permutations σ_1 et σ_2, $\varepsilon(\sigma_1 \sigma_2)$, vaut $\varepsilon(\sigma_1)\,\varepsilon(\sigma_2)$, etc.

Proposition 18.

Soient $\boldsymbol{E}$ un $\mathbb{K}$-espace vectoriel de dimension n et ϕ une forme n-linéaire alternée sur $\boldsymbol{E}$. Alors ϕ vérifie la propriété suivante :

$$\forall \sigma \in \sigma_n, \forall (V_1, \dots, V_n) \in E^n \;;$$

$$\phi(V_{\sigma(1)}, \dots, V_{\sigma(n)}) = \varepsilon(\sigma)\,\phi(V_1, \dots, V_n) \;.$$

Preuve. ◊ Puisque la forme n-linéaire ϕ est alternée, donc antisymétrique, on a pour toute transposition τ (permutant deux indices de l'ensemble $\{1, 2, \dots, n\}$) et tout n-uplet $(V_1, \dots, V_n)$ de $\boldsymbol{E}$:

$$\phi(V_{Z(1)}, \dots, V_{Z(n)}) = -\phi(V_1, \dots, V_n).$$

Mais il a été rappelé que toute permutation se décompose en produit $\sigma = Z_p \circ Z_{p-1} \circ \dots \circ Z_1$ de transpositions.

On a donc, en écrivant p fois l'égalité précédente :

$$\phi(V_{\sigma(1)}, \dots, V_{\sigma(n)}) = (-1)^p\,\phi(V_1, \dots, V_n)$$
$$= \varepsilon(\sigma)\,\phi(V_1, \dots, V_n)$$

où $\varepsilon(\sigma)$ désigne la signature de la permutation σ. ◊

6.2 Étude des formes n-linéaires alternées sur E (dim $E = n$)

On se propose maintenant d'étudier l'expression d'une forme n-linéaire alternée f sur l'espace vectoriel $\boldsymbol{E}$ rapporté à une base $\mathfrak{B} = (e_1, \dots, e_n)$.

Soient $V_1, V_2, \dots, V_n$ n vecteurs de $\boldsymbol{E}$, donnés par les expressions suivantes :

$$V_1 = m_{1,1}e_1 + \dots \quad \dots + m_{n,1}e_n$$
$$\vdots$$
$$V_n = m_{1,n}e_1 + \dots \quad \dots + m_{n,n}e_n$$

En utilisant la propriété de n-linéarité de f, on obtient :

$$f(V_1, \dots, V_n) = f\left(\sum_{i_1} m_{i_{1,1}} e_{i_1}, \dots, \sum_{i_n} m_{i_{n,n}} e_{i_n}\right)$$
$$= \sum_{i_1}\sum_{i_2}\dots\sum_{i_n} m_{i_{1,1}}\, m_{i_{2,2}} \dots m_{i_{n,n}}\; f(e_{i_1}, \dots, e_{i_n}).$$

La forme n-linéaire f est alternée, ce qui implique la nullité de $f(e_{i_1}, \dots, e_{i_n})$ dès que deux des indices $i_1, i_2, \dots, i_n$ sont égaux.

Dans le développement de $f(V_1, \dots, V_n)$, on ne s'intéresse donc qu'au cas où les indices $i_1, \dots, i_n$ sont tous différents et forment une permutation de $1, \dots, n$:

$$\sigma(1) = i_1,\ \sigma(2) = i_2,\ \dots,\ \sigma(n) = i_n \quad \text{où} \quad \sigma \in \sigma_n.$$

Puisque l'application f est alternée, on obtient :

$$f(e_{i_1}, \dots, e_{i_n}) = f(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$
$$= \varepsilon(\sigma)\, f(e_1, \dots, e_n)$$

et $m_{i_{1,1}} \dots m_{i_{n,n}}\; f(e_{i_1}, \dots, e_{i_n}) = \varepsilon(\sigma)\; m_{\sigma(1),1} \dots m_{\sigma(n),n}\; f(e_1, \dots, e_n)$

En reprenant l'expression développée de $f(V_1, \dots, V_n)$, on a l'égalité suivante :

$$f(V_1, \dots, V_n) = \sum_{\sigma \in \sigma_n} \varepsilon(\sigma)\; m_{\sigma(1),1} \dots m_{\sigma(n),n}\; f(e_1, \dots, e_n)$$
$$= \left[\sum_{\sigma \in \sigma_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i),i}\right] f(e_1, \dots, e_n).$$

Si l'on pose $f(e_1, \ldots, e_n) = C$, on obtient :

$$f(V_1, \ldots, V_n) = C\left(\sum_{\sigma \in \sigma_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i), i}\right).$$

En particulier, si $C = 1$, on a prouvé qu'il existe, au plus, une forme n-linéaire alternée θ vérifiant :

$$\theta(e_1, \ldots, e_n) = 1.$$

Réciproquement, on montre que l'application :

$$\theta : (V_1, \ldots, V_n) \rightarrow \sum_{\sigma \in \sigma_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i), i},$$

est n-linéaire et alternée.

Pour conclure, on peut donc énoncer la proposition suivante.

Proposition 19.

*Soit **E** un espace vectoriel de dimension n, rapporté à une base* $\mathfrak{B} = (e_1, \ldots, e_n)$.

a *Il existe une, et une seule, forme n-linéaire alternée prenant la valeur 1 sur la base* $\mathfrak{B}$. *C'est l'application déterminant dans la base* $\mathfrak{B}$ *ou* $\det_{\mathfrak{B}}$, *qui associe aux n vecteurs* $V_1, \ldots, V_n$ *de **E*** :

$$\begin{aligned} V_1 &= m_{1,1}e_1 + \ldots \quad \ldots + m_{n,1}e_n \\ &\vdots \\ V_n &= m_{1,n}e_1 + \ldots \quad \ldots + m_{n,n}e_n \end{aligned}$$

le scalaire $\det_{\mathfrak{B}}(V_1, \ldots, V_n) = \sum_{\sigma \in \sigma_n} \varepsilon(\sigma)\, m_{\sigma(1),1}, \ldots, m_{\sigma(n),n}$

encore noté :

$$\begin{vmatrix} m_{1,1} & \ldots & m_{1,n} \\ \vdots & & \\ m_{n,1} & \ldots & m_{n,n} \end{vmatrix}.$$

b *L'ensemble des formes n-linéaires alternées sur **E** est une droite vectorielle dirigée par cette application* $\det_{\mathfrak{B}}$. *En particulier, deux formes n-linéaires alternées non nulles sur **E** sont proportionnelles.*

Remarquons que l'on a construit une application qui répond à l'objectif fixé de caractériser les bases de ***E***.

Proposition 20.

*Soit **E** un espace vectoriel de dimension n, rapporté à une base* $\mathfrak{B} = (e_1, \ldots, e_n)$. *Alors une famille* $(V_1, \ldots, V_n)$ *forme une base si, et seulement si,* $\det_{\mathfrak{B}}(V_1, \ldots, V_n)$ *est non nul.*

Preuve. ◊ Soit $\mathcal{V} = (V_1, \ldots, V_n)$ une famille de n vecteurs de ***E***.

a) On veut prouver que si $\mathcal{V}$ est une base de ***E***, $\det_{\mathfrak{B}}(V_1, \ldots, V_n)$ est non nul.

Les deux formes n-linéaires alternées $\det_{\mathfrak{B}}$ et $\det_{\mathcal{V}}$ sont non nulles ; elles sont donc proportionnelles. Il existe donc un scalaire λ vérifiant :

$$\det_{\mathcal{V}} = \lambda \det_{\mathfrak{B}}.$$

Comme $\det_{\mathcal{V}}(V_1, \ldots, V_n) = \lambda \det_{\mathfrak{B}}(V_1, \ldots, V_n)$

et $\det_{\mathcal{V}}(V_1, \ldots, V_n) = 1$, $\det_{\mathfrak{B}}(V_1, \ldots, V_n)$ est non nul.

b) Nous voulons maintenant prouver que, si $\det_{\mathfrak{B}}(V_1, \ldots, V_n)$ est non nul, la famille $\mathcal{V}$ est une base de ***E***. Il suffit de prouver qu'elle est libre, puisqu'elle est constituée de n vecteurs d'un espace de dimension n.

Supposons la famille $(V_1, \ldots, V_n)$ liée, l'un des vecteurs, V_n par exemple, est combinaison linéaire des autres. Le vecteur V_n s'écrit $\lambda_1 V_1 + \ldots + \lambda_{n-1} V_{n-1}$ et l'on obtient :

$$\begin{aligned} \det_{\mathfrak{B}}(V_1, \ldots, V_n) &= \det_{\mathfrak{B}}(V_1, \ldots, V_{n-1}, \lambda_1 V_1 + \ldots + \lambda_{n-1} V_{n-1}) \\ &= \sum_{i=1}^{n-1} \lambda_i \det_{\mathfrak{B}}(V_1, \ldots, V_i, \ldots, V_{n-1}, V_i) \\ &= 0 \end{aligned}$$

(puisque $\det_{\mathfrak{B}}$ est une forme n-linéaire alternée).

On aboutit à une contradiction, la famille $\mathcal{V}$ est donc libre et constitue une base de ***E***. ◊

6.3 Déterminant d'une matrice carrée et d'un endomorphisme

On rappelle que la base canonique de l'espace vectoriel $\mathbb{K}^n$ est la famille $(e_1, \ldots, e_n)$ de $\mathbb{K}^n$ où e_i est le n-uplet dont tous les éléments sont nuls, sauf le $i^{\text{ième}}$ qui est égal à 1, c'est-à-dire :

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0).$$

Soit $A = (a_{i,j})$ un élément de $\mathcal{M}_n(\mathbb{K})$; l'expression de la $k^{\text{ième}}$ colonne C_k de A dans la base canonique de $\mathbb{K}^n$ est :

$$C_k = \sum_{i=1}^{n} a_{i,k}\, e_i.$$

Définition 13. On appelle déterminant de la matrice $A \in \mathcal{M}_n(\mathbb{K})$, le déterminant des n vecteurs colonnes $C_1, \ldots, C_n$ de A dans la base canonique :

$$\mathfrak{B} = (e_1, \ldots, e_n) \text{ de } \mathbb{K}^n.$$

On peut donc écrire, en notant $\det = \det_{\mathfrak{B}}$:

$$\det A = \det(C_1, \ldots, C_n) = \det(Ae_1, \ldots, Ae_n).$$

Proposition 21.

Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$. *Alors* :

$$\det(AB) = \det A \det B.$$

Preuve. ◊ Par définition :

$$\det(AB) = \det(ABe_1, \ldots, ABe_n)$$

ou encore :

$$\det(AB) = \det(AV_1, \ldots, AV_n)$$

si l'on pose :

$$V_1 = Be_1, \ldots, V_n = Be_n.$$

Étudions l'application :

$$\phi : \mathbb{K}^n \rightarrow \mathbb{K} \; ; (X_1, \ldots, X_n) \rightarrow \det(AX_1, \ldots, AX_n)$$

Il est clair que c'est une forme n-linéaire alternée donc proportionnelle à l'application det (qui est l'application déterminant dans la base canonique de $\mathbb{K}^n$).

Il existe donc un scalaire λ vérifiant :

$$\phi(X_1, ..., X_n) = \lambda \det(X_1, ..., X_n)$$

soit : $\det(AX_1, ..., AX_n) = \lambda \det(X_1, ..., X_n)$.

Comme $\det(Ae_1, ..., Ae_n) = \det A$ et $\det(e_1, ..., e_n) = 1$ ($(e_1, ..., e_n)$ étant la base canonique de $\mathbb{K}^n$ et det le déterminant de cette base), on a obtenu :

$$\lambda = \det A$$

et $\det(AV_1, ..., AV_n) = \det A \ \det(V_1, ..., V_n)$.

Faisant alors $V_1 = Be_1, ..., V_n = Be_n$, on a :

$$\det(ABe_1, ..., ABe_n) = \det A \ \det(Be_1, ..., Be_n)$$

ce qui signifie par définition que :

$$\det AB = \det A \ \det B. \qquad \lozenge$$

Proposition 22.

Soit I_n la matrice identité de $\mathcal{M}_n(\mathbb{K})$. Alors :

$$\det I_n = 1.$$

Soit A un élément de $\mathcal{M}_n(\mathbb{K})$. Alors A est inversible si, et seulement si :

$$\det A \neq 0 \ \text{ et } \ \det A^{-1} = \frac{1}{\det A}.$$

Preuve. ◊

- Par définition :

$$\det I_n = \det(I_n e_1, ..., I_n e_n) = \det(e_1, ..., e_n)$$

donc : $\det I_n = 1$.

- De même, par définition :

$$\det A = \det(Ae_1, ..., Ae_n).$$

Ce déterminant est non nul si, et seulement si, la famille $(Ae_1, ..., Ae_n)$ est une base de $\mathbb{K}^n$, c'est-à-dire si la famille des colonnes de A $(C_1, ..., C_n)$ est une base de $\mathbb{K}^n$, donc si A est inversible. On a donc prouvé que :

$$\det A \neq 0$$

si, et seulement si, A est inversible.

Si maintenant A est inversible :

$$AA^{-1} = I_n$$

et $$\det(AA^{-1}) = \det A \ \det A^{-1} = 1$$

ce qui donne :

$$\det A^{-1} = \frac{1}{\det A}. \qquad \lozenge$$

Transposée d'une matrice et déterminant

Si A est une matrice à n lignes et n colonnes dont les éléments sont notés $(a_{i,j})$, on appelle **matrice transposée** de A, la matrice à n lignes et n colonnes, notée ${}^tA$, dont les éléments $(a'_{k,\ell})$ valent :

$$a'_{k,\ell} = a_{\ell,k}.$$

Cela revient à dire que l'élément qui figure sur la $k^{\text{ième}}$ ligne et la $\ell^{\text{ième}}$ colonne de ${}^tA$ est celui qui se trouve sur la $\ell^{\text{ième}}$ ligne et la $k^{\text{ième}}$ colonne de A. En particulier, les colonnes de la matrice ${}^tA$ sont les lignes de la matrice A. On établit la proposition suivante.

Proposition 23.

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors :

$$\det A = \det {}^tA.$$

Le déterminant de la matrice A qui, par définition, est le déterminant dans la base canonique de $\mathbb{K}^n$ de ses colonnes, est donc aussi le déterminant dans la base canonique de $\mathbb{K}^n$ de ses lignes.

Nous allons maintenant définir la notion de déterminant d'un endomorphisme.

Proposition 24.

Si deux matrices A et B appartenant à $\mathcal{M}_n(\mathbb{K})$ sont semblables, leurs déterminants sont égaux.

Preuve. ◊ Il existe une matrice $P \in GL_n(\mathbb{K})$ vérifiant $A = PBP^{-1}$ et l'on a :

$$\det A = \det P \ \det B \det P^{-1} = \det B. \qquad \lozenge$$

Définition 14. Soient ***E*** un espace vectoriel de dimension n et f un endomorphisme de ***E***. On appelle déterminant de f, le déterminant de la matrice de f dans une base quelconque de ***E***.

Cette définition a un sens puisque les matrices de f dans des bases différentes sont semblables.

En utilisant les propriétés des matrices, on établit, sans difficulité, la proposition suivante.

Proposition 25.

Si Id *est l'application identité de $\mathcal{L}(\boldsymbol{E})$* : $\det \mathrm{Id} = 1$.

*Soient f et g deux endomorphismes de **E**. Alors* :

$$\det f \circ g = \det f \ \det g.$$

De plus, f est un automorphisme si, et seulement si, $\det f \neq 0$ *et, dans ce cas* :

$$\det f^{-1} = \frac{1}{\det f}.$$

6.4 Calcul des déterminants

On se propose maintenant de donner des méthodes de calcul des déterminants d'une matrice.

En utilisant les propriétés des formes n-linéaires alternées, on peut faire les remarques suivantes :

— multiplier la $i^{\text{ième}}$ colonne (ou la $i^{\text{ième}}$ ligne) d'une matrice par un scalaire λ revient à multiplier son déterminant par λ ;

— ajouter à la $i^{\text{ième}}$ colonne (ou $i^{\text{ième}}$ ligne) d'une matrice une combinaison linéaire des autres colonnes (ou des autres lignes) ne modifie pas le déterminant de cette matrice.

6.4.1 Calcul par éléments

6.4.1.1 Cas d'une matrice triangulaire

Soit $T = \begin{pmatrix} t_{1,1} & \cdots & t_{1,n} \\ & \ddots & \\ (0) & & t_{n,n} \end{pmatrix}$.

Alors :

$$\det T = \sum_{\sigma \in \sigma_n} \varepsilon(\sigma) \ t_{\sigma(1),1} \cdots t_{\sigma(n),n}.$$

Or le coefficient $t_{i,j}$ de la matrice T est nul dès que $i > j$.

Le produit $t_{\sigma(1),1}, ..., t_{\sigma(n),n}$, n est nul dès qu'il existe un indice i tel que $\sigma(i) > i$. Il faut donc ne conserver que les permutations

$\sigma \in \sigma_n$ vérifiant $\sigma(1) \leqslant 1, \ldots, \sigma(n) \leqslant n$, c'est-à-dire l'identité, et comme ε (Id) = 1, on a établi :

$$\det T = t_{1,1}\, t_{2,2} \ldots t_{n,n}.$$

Le déterminant d'une matrice triangulaire (et a fortiori diagonale) est donc égal au produit des éléments diagonaux.

6.4.1.2 Cas général. Développement suivant une ligne ou une colonne

■ **Si $n = 2$:**

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

$$\det A = \sum_{\sigma \in \sigma_2} \varepsilon(\sigma)\, a_{\sigma(1),1}, a_{\sigma(2),2}.$$

L'ensemble des permutations σ_2 est constitué de deux permutations : l'identité et la transposition (de signature – 1) échangeant (1,2).

On a donc :

$$\det A = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}\, a_{2,2} - a_{2,1}\, a_{1,2}.$$

■ **Si $n = 3$:**

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

$$\det A = \sum_{\sigma \in \sigma_2} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2}\, a_{\sigma(3),3}$$

Or σ_3 est constitué des six éléments :

— l'identité ;

— les trois transpositions échangeant (1,2), (1,3) et (2,3) ;

— les deux permutations circulaires c_1 définie par $c_1(1) = 2$, $c_1(2) = 3$, $c_1(3) = 1$ et c_2 définie par $c_2(1) = 3$, $c_2(2) = 1$, $c_2(3) = 2$.

La signature de c_1 et c_2 vaut 1, car ces deux permutations sont produits de deux transpositions.

On obtient donc :

$$\det A = a_{1,1}\, a_{2,2}\, a_{3,3} + a_{2,1}\, a_{3,2}\, a_{1,3} + a_{3,1}\, a_{1,2}\, a_{2,3} - a_{1,1} a_{2,3}\, a_{3,2} - a_{3,1}\, a_{2,2}\, a_{1,3} - a_{2,1}\, a_{1,2}\, a_{3,3},$$

que l'on peut écrire :

$$\det A = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{2,1} \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + a_{3,1} \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{vmatrix}$$

Dans ce cas, on dit que l'on a développé le déterminant de A suivant la première colonne de A. On peut également écrire :

$$\det A = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2} \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3} \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix}$$

Dans ce cas, on dit que l'on a développé le déterminant de A suivant la première ligne de A.

Nous allons maintenant donner la formule généralisant ce résultat.

■ **Si n est quelconque :**

$$A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \\ a_{n,1} & \ldots & a_{n,n} \end{pmatrix}$$

On appelle :

— mineur $\Delta_{i,j}$ de $a_{i,j}$, le déterminant de la matrice à $n - 1$ lignes et $n - 1$ colonnes obtenu en supprimant la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne de A ;

— cofacteur de $m_{i,j}$, le scalaire $(-1)^{i+j}\Delta_{i,j}$.

On établit le résultat suivant.

Proposition 26.

Soit A un élément de $\mathcal{M}_n(\mathbb{K})$.

On a les égalités suivantes :

a) *Pour* $1 \leqslant j \leqslant n$:

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

C'est le développement du déterminant de A suivant la $j^{\text{ième}}$ colonne de A.

b) *Pour* $1 \leqslant i \leqslant n$:

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

C'est le développement du déterminant de A suivant la $i^{\text{ième}}$ ligne de A.

6.4.2 Calcul par blocs

Soit M une matrice à n lignes et n colonnes s'écrivant :

$$M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$$

où $A \in \mathcal{M}_p(\mathbb{K})$, $C \in \mathcal{M}_q(\mathbb{K})$ $(p + q = n)$.

On établit que :

$$\det M = \det A \det B.$$

On peut généraliser cette formule :

$$M = \left(\begin{array}{c|c|c|c} A_1 & * & \ldots & * \\ \hline 0 & A_2 & & \\ \hline & & & * \\ \hline 0 & & 0 & A_p \end{array}\right)$$

où $A_i \in \mathcal{M}_{n_i}(\mathbb{K})$ $(n_1 + \ldots + n_p = n)$, * signifiant qu'il y a des valeurs.

On établit que :

$$\det M = \det A_1 \ldots \det A_p.$$

6.4.3 Exemples

6.4.3.1 Calcul du déterminant de Van der Monde

On pose :

$$V(a_0, a_1, \ldots, a_n) = \begin{pmatrix} 1 & a_0 & a_0^2 & \ldots & a_0^n \\ 1 & a_1 & a_1^2 & \ldots & a_1^n \\ \vdots & & & & \\ 1 & a_n & a_n^2 & \ldots & a_n^n \end{pmatrix}$$

où $a_0, \ldots, a_n$ sont n scalaires quelconques.

On veut prouver que :

$$V(a_0, a_1, \ldots, a_n) = \prod_{0 \leqslant i < j \leqslant n} (a_j - a_i).$$

Ainsi ce déterminant est non nul si, et seulement si, $a_0, a_1, \ldots, a_n$ sont distincts.

On retranche à la $n^{\text{ième}}$ colonne a_n fois la $(n-1)^{\text{ième}}$, puis à la $(n-1)^{\text{ième}}$ colonne a_n fois la $(n-2)^{\text{ième}}$ et ainsi de suite.

On obtient :

$$V(a_0, a_1, \ldots, a_n) = \begin{vmatrix} 1 & (a_0 - a_n) & \ldots & a_0^{n-1}(a_0 - a_n) \\ \vdots & & & \\ 1 & (a_{n-1} - a_n) & \ldots & a_{n-1}^{n-1}(a_{n-1} - a_n) \\ 1 & 0 & \ldots & 0 \end{vmatrix}$$

On développe ensuite suivant la dernière ligne et l'on obtient :

$$\begin{aligned} V(a_0, a_1, \ldots, a_n) &= (-1)^{n+1} \begin{vmatrix} a_0 - a_n & (a_0 - a_n)a_0 & \ldots & (a_0 - a_n)\, a_0^{n-1} \\ \vdots & & & \\ a_{n-1} - a_n & (a_{n-1} - a_n)a_{n-1} & \ldots & (a_{n-1} - a_n)\, a_{n-1}^{n-1} \end{vmatrix} \\ &= (-1)^{n+1}(a_0 - a_n)(a_1 - a_n)\ldots(a_{n-1} - a_n)\; V(a_0, \ldots, a_{n-1}) \\ &= (a_n - a_{n-1})(a_n - a_{n-2})\ldots(a_n - a_0)\; V(a_0, \ldots, a_{n-1}). \end{aligned}$$

Comme $V(a_0, a_1) = \begin{vmatrix} 1 & a_0 \\ 1 & a_1 \end{vmatrix} = (a_1 - a_0)$, on a donc prouvé par récurrence que :

$$V(a_0, \ldots, a_n) = \prod_{0 \leqslant i < j \leqslant n} (a_j - a_i).$$

6.4.3.2 Calcul de $\begin{vmatrix} \boldsymbol{a+b} & \boldsymbol{a} & & (\boldsymbol{0}) \\ \boldsymbol{b} & \boldsymbol{a+b} & \ddots & \\ & \boldsymbol{b} & \ddots & \boldsymbol{a} \\ (\boldsymbol{0}) & & \boldsymbol{b} & \boldsymbol{a+b} \end{vmatrix}$

On pose :

$$\Delta_n = \begin{vmatrix} a+b & a & & (0) \\ b & a+b & a & \\ & b & \ddots & a \\ (0) & & b & a+b \end{vmatrix}$$

Δ_n étant le déterminant d'une matrice à n lignes et n colonnes.

On développe Δ_n suivant la première ligne :

$$\Delta_n = (a+b)\Delta_{n-1} - b \begin{bmatrix} a & & & (0) \\ b & a+b & a & \\ & (0) & \ddots & a \\ & & b & a+b \end{bmatrix}$$

et l'on développe le dernier déterminant de l'égalité ci-dessus suivant la 1re ligne.

On obtient la relation :

$$\Delta_n = (a+b)\,\Delta_{n-1} - ab\,\Delta_{n-2}.$$

Comme $(\Delta_n)_{n \in \mathbb{N}}$ vérifie la relation de récurrence :

$$\Delta_n = (a+b)\,\Delta_{n-1} - ab\,\Delta_{n-2},$$

alors Δ_n s'écrit : $\lambda\, a^n + \mu\, b^n$.

On a :

$$\Delta_1 = a + b \text{ et } \Delta_2 = (a+b)^2 - ab,$$

et donc :

$$\lambda = \frac{a}{a-b}, \mu = \frac{-b}{a-b}.$$

D'où :

— si $a \neq b$:

$$\Delta_n = \frac{a^n - b^n}{a-b}.$$

— si $a = b$, on trouve :

$$\Delta_n = n a^{n-1}.$$

On peut faire tendre b vers a et utiliser le fait que le déterminant est une fonction continue de ses coefficients donc de la variable b.

6.5 Application des déterminants

Les deux premières propriétés ci-dessous résultent de ce qui a déjà été établi précédemment.

- **Caractérisation des bases**

Une famille de n vecteurs d'un espace de dimension n est une base si, et seulement si, le déterminant de ses n vecteurs dans une base est non nul.

- **Caractérisation des matrices inversibles**

Une matrice carrée est inversible si, et seulement si, son déterminant est non nul.

- **Formules de Cramer**

Proposition 27.

Soient A une matrice de $\mathcal{M}_n(\mathbb{K})$, inversible, et B un vecteur de $\mathbb{K}^n$. Le système $AX = B$ a une solution unique :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

donnée, pour $1 \leqslant i \leqslant n$, par :

$$x_i = \frac{\det(V_1, \ldots, V_{i-1}, B, V_{i+1}, \ldots, V_n)}{\det(V_1, \ldots, V_n)}$$

où $(V_1, \ldots, V_n)$ sont les vecteurs colonnes de A.

Preuve. ◊ On a :

$$\sum_{k=1}^{n} x_k V_k = B$$

et $\det(V_1, ..., V_{i-1}, B, V_{i+1}, ..., V_x)$

$$= \det\left(V_1, ..., V_{i-1}, \sum_{k=1}^{n} x_k V_k, V_{i+1}, ..., V_n\right)$$

$$= \sum_{k=1}^{n} x_k \det(V_1, ..., V_{i-1}, V_k, V_{i+1}, ..., V_1).$$

Comme $\det(V_1, ..., V_{i-1}, V_k, V_{i+1}, ..., V_n)$ est nul si $k \neq i$, on obtient :

$$\det(V_1, ..., V_{i-1}, B, V_{i+1}, ..., V_n) = x_i \det(V_1, ..., V_n). \qquad \lozenge$$

Calcul de l'inverse d'une matrice

Soit M une matrice de $\mathcal{M}_n(\mathbb{K})$.

On note $\mathrm{Com}(M) = ((-1)^{i+j}\Delta_{ij})$ (comatrice de M) la matrice des cofacteurs de la matrice M. Nous allons démontrer l'égalité suivante :

$$M \cdot {}^t\mathrm{Com}(M) = {}^t\mathrm{Com}(M) \cdot M = \det M \cdot I_n.$$

On note $\alpha_{i,j}$ le coefficient d'indice (i,j) du produit $M \cdot {}^t\mathrm{Com}(M)$. Alors :

$$\alpha_{i,j} = \sum_{k=1}^{n} m_{i,k}(-1)^{j+k}\Delta_{j,k}.$$

Si $i = j$, $\alpha_{i,i} = \sum_{k=1}^{n} m_{i,k}(-1)^{i+k}\Delta_{i,k} = \det M$.

Cette dernière égalité correspond au développement suivant la $i^{\text{ième}}$ ligne.

Si $i \neq j$, notons M' la matrice déduite de M en remplaçant le $j^{\text{ième}}$ ligne par la $i^{\text{ième}}$ ligne, ce qui signifie que M' a deux lignes égales la $i^{\text{ième}}$ et la $j^{\text{ième}}$.

Si l'on développe le déterminant de M' suivant la $j^{\text{ième}}$ ligne, on obtient :

$$\det M' = \sum_{k=1}^{n} m_{i,k}(-1)^{j+k}\Delta_{j,k} = 0.$$

On a prouvé que :

$$M \cdot {}^t\mathrm{Com}(M) = \det M \cdot I_n.$$

On obtient de même :

$${}^t\mathrm{Com}(M) \cdot M = \det M \cdot I_n.$$

On peut énoncer la proposition suivante.

Proposition 28.

Soient M une matrice de $\mathcal{M}_n(\mathbb{K})$, et $\mathrm{Com}(M)$ la matrice des cofacteurs de M ou comatrice de M. Alors :

$$M \cdot {}^t\mathrm{Com}(M) = {}^t\mathrm{Com}(M) \cdot M = \det M \cdot I_n.$$

Soit M une matrice inversible de $\mathcal{M}_n(\mathbb{K})$; l'égalité précédente s'écrit :

$$M \cdot \frac{1}{\det M} \cdot {}^t\mathrm{Com}(M) = \frac{1}{\det M} \cdot {}^t\mathrm{Com}(M) \cdot M = I_n$$

et l'on obtient l'expression de l'inverse de M sous la forme :

$$M^{-1} = \frac{1}{\det M} \cdot {}^t\mathrm{Com}(M).$$

Caractérisation du rang d'une matrice

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. Le rang de M est la taille du plus grand déterminant extrait non nul.

Pour qu'une matrice M soit de rang r, il faut et il suffit que :

— il existe un déterminant extrait d'ordre r non nul ;

— tous les déterminants extraits d'ordre strictement supérieur à r soient nuls.

7. Réduction des endomorphismes et des matrices carrées

En pratique, le corps de base $\mathbb{K}$ sera $\mathbb{R}$ ou $\mathbb{C}$.

7.1 Éléments propres d'un endomorphisme et d'une matrice carrée

7.1.1 Généralités

On rappelle les définitions suivantes.

Définition 15. Soit f un endomorphisme. Un scalaire λ est valeur propre de f si, et seulement si :

a) il existe un vecteur non nul v tel que $f(v) = \lambda v$;

b) le noyau de l'endomorphisme $f - \lambda\,\mathrm{Id}$ n'est pas réduit au vecteur nul ;

c) l'endomorphisme $f - \lambda\,\mathrm{Id}$ n'est pas injectif ;

d) $\det(f - \lambda\,\mathrm{Id}) = 0$.

L'équivalence de ces propriétés résulte simplement de ce qui a été développé auparavant.

Définition 16. Soit f un endomorphisme.

On dit qu'un vecteur non nul v est un vecteur propre de f s'il existe un scalaire λ tel que $f(v) = \lambda v$.

On dit que v est un vecteur propre associé à λ ; l'ensemble des vecteurs propres associés à λ auquel on adjoint le vecteur nul est le sous-espace propre $\mathrm{Ker}(f - \lambda\,\mathrm{Id})$ associé à λ.

Une méthode (au moins théorique) de recherche des valeurs propres λ d'un endomorphisme f est donc la recherche des racines λ de l'équation en x :

$$\det(f - x\,\mathrm{Id}) = 0$$

à laquelle on s'intéresse maintenant. A cet effet, on introduit la matrice $M = (m_{i,j})$ de f relativement à une base de $\boldsymbol{E}$. On sait que le déterminant de $f - x\,\mathrm{Id}$ est celui de sa matrice $M - xI_n$ relativement à n'importe quelle base de $\boldsymbol{E}$.

L'équation précédente devient donc :

$$\det(f - x\,\mathrm{Id}) = \begin{vmatrix} m_{1,1} - x & m_{1,2} & \dots & m_{1,n} \\ m_{2,1} & m_{2,2} - x & \dots & m_{2,n} \\ \vdots & & & \\ m_{n,1} & \dots & \dots & m_{n,n} - x \end{vmatrix}$$

Le déterminant d'une matrice s'obtenant par somme et produit de ses coefficients, $P_f(x) = \det(f - x\,\mathrm{Id})$ est donc un polynôme en x que l'on peut écrire ainsi en notant $m_{i,j}(x)$ le coefficient d'indice i,j de la matrice $M - x I_n$:

$$P_f(x) = (m_{1,1} - x)(m_{2,2} - x)\dots(m_{n,n} - x) + \sum_{\sigma \in \mathfrak{S}_n - \mathrm{Id}} \varepsilon(\sigma)\, m_{\sigma(1),1}(x)\ \dots\ m_{\sigma(n),n}(x)$$

Dans le développement du déterminant, si $\sigma \neq \mathrm{Id}$, il existe un indice j tel que $\sigma(j) \neq j$ et donc au moins un autre indice i tel que $\sigma(i) \neq i$ (car σ est bijective de {1, 2, ..., x} dans lui-même) ; le degré en x des termes $\varepsilon(\sigma)\, m_{\sigma(1),1}(x)\ \dots\ m_{\sigma(n),n}(x)$, pour $\sigma \neq \mathrm{Id}$, est au plus $n - 2$.

a) Étudions le degré de P_f. Le terme de plus haut degré du polynôme $P_f(x)$ est obtenu au développement $(m_{1,1} - x)\ \dots\ (m_{n,n} - x)$ et vaut donc $(-1)^n x^n$.

b) Le coefficient constant du polynôme $P_f(x)$ s'obtient en faisant $x = 0$ et vaut donc $P_f(0) = \det f$.

c) Le coefficient de x^{n-1} est obtenu en développant $(m_{1,1} - x)\ \dots\ (m_{n,n} - x)$. Son coefficient vaut $(-1)^{n-1}(m_{1,1} + \dots + m_{n,n})$. Ce nombre, obtenu en faisant la somme des éléments diagonaux de M, s'appelle la trace de M. On vérifie son invariance par changement de bases ($\mathrm{tr}\, P^{-1}MP = \mathrm{tr}\, M$), et on l'appelle aussi trace de l'endomorphisme f.

Le coefficient de x^{n-1} est donc $(-1)^{n-1}\,\mathrm{tr}\,(f)$.

De cela résulte la définition suivante.

Définition 17. Soit f un endomorphisme d'un espace $\boldsymbol{E}$ de dimension finie. On appelle polynôme caractéristique de f le polynôme :

$$P_f(x) = \det(f - x\,\mathrm{Id}) = (-1)^n[x^n - \mathrm{tr}\,(f)x^{n-1} + \dots + (-1)^n \det f].$$

On sait que les valeurs propres d'un endomorphisme sont les racines de son polynôme caractéristique. On pose la définition suivante.

Définition 18. Soit f un endomorphisme de $\boldsymbol{E}$. Les valeurs propres de f sont les racines de son polynôme caractéristique P_f, et l'on convient d'appeler **ordre de multiplicité** d'une valeur propre son **ordre de multiplicité** en tant que racine du polynôme caractéristique de f.

On sait qu'un polynôme de degré n admet au plus n racines (y compris en tenant compte des ordres de multiplicité). Il en résulte qu'un endormorphisme d'un espace de dimension n admet au plus n valeurs propres (y compris en tenant compte des ordres de multiplicité).

Définition 19. On appelle **endomorphisme scindé** tout endomorphisme dont le polynôme caractéristique est scindé.

Lorsque le corps de base est celui des complexes, tout endomorphisme est scindé.

Si $\lambda_1, \lambda_2, \dots, \lambda_n$ désignent les valeurs propres distinctes ou non d'un endomorphisme scindé d'un espace vectoriel de dimension n, on a :

$$P_f(x) = (-1)^n[x^n - \mathrm{tr}\, f\, x^{n-1} + \dots + (-1)^n \det f] = (-1)^n (x - \lambda_1)\dots(x - \lambda_n).$$

Cela donne par identification :

$$\lambda_1 + \dots + \lambda_n = \mathrm{tr}\, f \quad ; \quad \lambda_1 \dots \lambda_n = \det f.$$

On se propose maintenant de comparer, pour un endomorphisme donné, l'ordre de multiplicité d'une valeur propre et la dimension du sous-espace propre correspondant.

Proposition 29.

Soient f un endomorphisme d'un espace vectoriel $\boldsymbol{E}$ *de dimension finie et* λ *une valeur de f. Alors, l'ordre de multiplicité de* λ *est supérieur ou égal à la dimension du sous-espace propre associé à la valeur propre* λ.

Preuve. ◊ Notons $\boldsymbol{E}_\lambda$ le sous-espace propre associé à la valeur propre λ. Soit $(e_1, \dots, e_p)$ une base de $\boldsymbol{E}_\lambda$; la dimension de $\boldsymbol{E}_\lambda$ vaut donc p.

On complète cette base en une base $\mathfrak{B}$ de $\boldsymbol{E}$ $(e_1, \dots, e_n)$.

La matrice de f dans la base $\mathfrak{B}$ s'écrit :

$$M = \begin{array}{c} \\ p\updownarrow \\ \\ \end{array} \begin{array}{|cc|c|} \multicolumn{2}{c}{\overleftrightarrow{\ \ p\ \ }} & \multicolumn{1}{c}{} \\ \lambda & (0) & * \\ (0) & \lambda & \\ \hline \multicolumn{2}{|c|}{(0)} & M' \end{array}$$

Le polynôme caractéristique f vaut :

$$P_f(x) = \det(M - x I_n) = (\lambda - x)^p \det(M' - x I_{n-p}).$$

Comme $(\lambda - x)^p$ divise $P_f(x)$, λ est racine d'ordre p au moins de P_f et donc l'ordre de multiplicité de la valeur propre λ est supérieur ou égal à la dimension p du sous-espace propre associé à λ. ◊

Notons que tout ce qui a été défini dans ce paragraphe pour un endomorphisme peut l'être également, de façon immédiate, pour une matrice carrée : valeurs propres, vecteurs propres, sous-espaces propres, polynôme caractéristique, etc.

Exemples.

① Soit la matrice :

$$M_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \theta \notin \pi\,\mathbb{Z}.$$

Le polynôme caractéristique de M_θ est :

$$\det(M_\theta - x I_2) = \begin{vmatrix} \cos\theta - x & -\sin\theta \\ \sin\theta & \cos\theta - x \end{vmatrix} = (\cos\theta - x)^2 + \sin^2\theta.$$

Dans $\mathbb{C}$, le polynôme caractéristique est scindé :

$$P_{M_\theta}(x) = (x - e^{i\theta})(x - e^{-i\theta})$$

et les valeurs propres de M_θ sont $e^{i\theta}$ et $e^{-i\theta}$.

Dans $\mathbb{R}$, M_θ n'a pas de valeurs propres.

7.1.2 Matrices symétriques réelles

On donne maintenant un exemple, fondamental par ses applications, de matrices réelles dont on sait a priori que les valeurs propres sont réelles.

Proposition 30.

Les valeurs propres d'une matrice symétrique réelle sont nécessairement réelles.

Preuve. ◊ Comme c'était le cas dans l'exemple ① pour la matrice M_θ, les valeurs propres d'une matrice réelle sont a priori à chercher dans le corps $\mathbb{C}$ des nombres complexes. Soit donc λ une valeur propre (réelle ou complexe) d'une matrice symétrique réelle d'ordre n notée M. Il existe donc dans $\mathbb{C}^n$ un vecteur non nul X tel que :

$$MX = \lambda X.$$

On notera $\bar{\lambda}$ le conjugué (dans $\mathbb{C}$) du nombre complexe λ et $\bar{X}$ la matrice conjuguée de X c'est-à-dire dont les éléments sont les conjuguées des éléments de X.

On multiplie cette égalité à gauche par la matrice ligne ${}^t\bar{X}$

$${}^t\bar{X}\,MX = \lambda\,{}^t\bar{X}\,X = \lambda\sum_{k=1}^{n}|x_k|^2.$$

Conjuguons cette égalité. Comme M est réelle, on a $\bar{M} = M$, d'où :

$${}^tX\,M\,\bar{X} = \bar{\lambda}\sum_{k=1}^{n}|x_k|^2.$$

Transposons cette égalité. Comme M est symétrique, on a ${}^tM = M$, d'où :

$${}^t\bar{X}\,MX = \bar{\lambda}\sum_{k=1}^{n}|x_k|^2.$$

En comparant cette égalité à la première, on a :

$$\lambda\sum_{k=1}^{n}|x_k|^2 = \bar{\lambda}\sum_{k=1}^{n}|x_k|^2.$$

Puisque le vecteur propre X est non nul, la somme $\sum_{k=1}^{n}|x_k|^2$ est non nulle. Il en résulte que $\lambda = \bar{\lambda}$ et donc que la valeur propre λ est nécessairement réelle. ◊

Exemple : cherchons les valeurs propres de la matrice symétrique réelle :

$$M = \begin{pmatrix} m & & (1) \\ & \ddots & \\ (1) & & m \end{pmatrix}.$$

Le polynôme caractéristique est :

$$\det(M - x\,I_n) = \begin{vmatrix} m-x & & (1) \\ & \ddots & \\ (1) & & m-x \end{vmatrix} = \begin{vmatrix} m+n-1-x & & (1) \\ \vdots & m-x & \\ m+n-1-x & (1) & m-x \end{vmatrix}$$

par addition de toutes les colonnes à la première.

Poursuivons en retranchant la première ligne à toutes les suivantes. On a :

$$\det(M - x\,I_n) = \begin{vmatrix} m+n-1-x & 1 & \dots & 1 \\ 0 & m-x-1 & 0 & \dots 0 \\ \vdots & 0 & \ddots & \\ & \vdots & \ddots & 0 \\ 0 & 0 & 0 & m-x-1 \end{vmatrix}$$

$$= (m+n-1-x)(m-x-1)^{n-1}$$

Ainsi donc cette matrice a pour valeurs propres :

$x = m + n - 1$ avec un ordre de multiplicité égal à 1 ;

$x = m - 1$ avec un ordre de multiplicité égal à $n - 1$.

7.2 Trigonalisation et théorème de Hamilton-Cayley

Définition 20. Soit f un endomorphisme d'un espace vectoriel de dimension finie $\boldsymbol{E}$. On dit que f est trigonalisable s'il existe une base de $\boldsymbol{E}$ dans laquelle la matrice de f est triangulaire, c'est-à-dire da la forme :

$$T = \begin{pmatrix} t_{1,1} & t_{1,2} & \dots & t_{1,n} \\ 0 & t_{2,2} & \dots & \\ & \ddots & \ddots & \\ 0 & & 0 & t_{n,n} \end{pmatrix}.$$

On dit qu'une matrice carrée M d'ordre n est trigonalisable s'il existe un changement de bases la rendant triangulaire, c'est-à-dire s'il existe une matrice inversible d'ordre n P telle que $T = P^{-1}\,MP$ soit triangulaire.

Proposition 31.

Un endomorphisme f (ou une matrice carrée M) est trigonalisable si, et seulement si, son polynôme caractéristique est scindé.

Preuve. ◊ Supposons f trigonalisable. En utilisant une base dans laquelle la matrice de f est une matrice triangulaire T, on a :

$$\det(f - x\,\mathrm{Id}) = \begin{vmatrix} t_{1,1}-x & t_{1,2} & \dots & t_{1,n} \\ & \ddots & & \\ (0) & & & t_{n,n}-x \end{vmatrix} = \prod_{k=1}^{n}(t_{k,k}-x).$$

Ce polynôme est bien scindé. De plus, on remarque que les valeurs propres de f ne sont autres que les éléments diagonaux $t_{1,1}, t_{2,2}, \dots, t_{n,n}$ de la matrice T.

■ Réciproquement, supposons l'endomorphisme f scindé. Il admet donc au moins une valeur propre λ.

Notons v un vecteur propre associé, et complétons celui-ci en une base $(v, v_2, \dots, v_n)$ de $\boldsymbol{E}$. La matrice de f dans cette base est alors :

$$M_n = \left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & M_{n-1} & \\ 0 & & v & \end{array}\right)$$

Si $n = 2$, cette matrice est déjà triangulaire.

Sinon, on raisonne par récurrence sur n en supposant le résultat vrai au rang $n - 1$. L'égalité ci-dessus donne :

$$\det(f - x\,\mathrm{Id}) = \det(M_n - x\,I_n) = (\lambda - x)\det(M_{n-1} - x\,I_{n-1}).$$

L'hypothèse étant que f est scindé, il est clair que $M_{n-1} - x\,I_{n-1}$ est scindé et donc, d'après l'hypothèse de récurrence, M_{n-1} est trigonalisable. Il existe donc une matrice inversible P_{n-1} telle que :

$$P_{n-1}^{-1}\,M_{n-1}\,P_{n-1} = T_{n-1}$$

est triangulaire.

On a alors en posant :

$$P_n = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P_{n-1} & \\ 0 & & & \end{array}\right) \text{ et donc } P_n^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P_{n-1}^{-1} & \\ 0 & & & \end{array}\right)$$

$$P_n^{-1} M_n P_n = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P_{n-1}^{-1} & \\ 0 & & & \end{array}\right)\left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & M_{n-1} & \\ 0 & & & \end{array}\right)\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P_{n-1} & \\ 0 & & & \end{array}\right)$$

$$= \left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & T_{n-1} & \\ 0 & & & \end{array}\right).$$

Comme T_{n-1} est triangulaire, cette matrice est elle-même triangulaire, ce qui achève la démonstration. ◊

Théorème de Hamilton-Cayley

Un endomorphisme f (ou une matrice carrée M) annule son polynôme caractéristique. Autrement dit, si :

$$P_f(x) = a_n x^n + \dots + a_1 x + a_0$$

alors :

$$P_f(f) = a_n f^n + \dots + a_1 f + a_0 \text{ Id} = 0$$

ou, matriciellement, si $P_M(x) = a_n x^n + \dots + a_1 x + a_0$:

$$P_M(M) = a_n M^n + \dots + a_1 M + a_0 I_n = 0.$$

Preuve. ◊ Supposons, tout d'abord, que le corps $\mathbb{K}$ soit celui des nombres complexes et soit f un endomorphisme du $\mathbb{C}$ espace $\boldsymbol{E}$. Le polynôme caractéristique de f est scindé et il existe donc une base $(e_1, \dots, e_n)$ de $\boldsymbol{E}$ dans laquelle la matrice de f est une matrice triangulaire :

$$T = \begin{pmatrix} t_{1,1} & t_{1,2} & \dots & t_{1,n} \\ & t_{2,2} & & \\ (0) & & \ddots & \\ & & & t_{n,n} \end{pmatrix}.$$

Dans cette base, on a donc pour $1 \leqslant k \leqslant n$:

$$f(e_k) = t_{1,k} e_1 + \dots + t_{k,k} e_k$$

où $f(e_k) - t_{k,k} e_k \in \text{Vect}(e_1, \dots, e_{k-1})$.

De plus, le polynôme caractéristique de f est :

$$P_f(x) = \det(f - x \text{ Id}) = \det(T - x I_n) = (t_{1,1} - x) \dots (t_{n,n} - x).$$

Nous montrons maintenant le théorème en prouvant que $P_f(f)$ est nul sur $\boldsymbol{E}$, c'est-à-dire sur la base $(e_1, \dots, e_n)$:

$$\forall k \in \{1, 2 \dots, n\} \quad P_f(f)(e_k) = (t_{1,1} \text{ Id} - f) \circ \dots \circ (t_{n,n} \text{ Id} - f)(e_k) = 0$$

En effet, les polynômes en f figurant dans ce produit commutent ; il suffit donc de prouver que $(t_{1,1} \text{ Id} - f) \circ \dots \circ (t_{k,k} \text{ Id} - f)(e_k) = 0$, ce qui est vrai puisque :

— si $k = 1$, $(t_{1,1} \text{ Id} - f)(e_1) = 0$;
— si $k = 2$,

$$(t_{1,1} \text{ Id} - f) \circ (t_{2,2} \text{ Id} - f)(e_2) \in (t_{1,1} \text{ Id} - f)(\text{Vect}(e_1)) = \{0\}.$$

Par un raisonnement par récurrence immédiat, on obtient :

$$(t_{1,1} \text{ Id} - f) \circ \dots \circ (t_{k,k} \text{ Id} - f)(e_k) \in (t_{1,1} \text{ Id} - f)$$
$$\circ \dots \circ (t_{k-1,k-1} \text{ Id} - f)(\text{Vect}(e_1, e_2, e_{k-1})) = \{0\}.$$

Ainsi donc $P_f(f) = 0$ et donc $P_M(M) = 0$ pour toute matrice carrée complexe.

Toute matrice carrée réelle pouvant être considérée comme une matrice carrée complexe, on a donc a fortiori $P_M(M) = 0$ pour toute matrice carrée réelle et donc $P_f(f) = 0$ pour tout endomorphisme réel.

Exemple : soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On a :

$$P_M(x) = x^2 - (a + d)x + (ad - bc) = x^2 - \text{tr}(M)\, x + \det(M).$$

Le théorème de Hamilton-Cayley indique donc que toute matrice carrée d'ordre 2 vérifie la relation :

$$M^2 - \text{tr}(M) M + \det(M)\, I_2 = 0.$$

7.3 Diagonalisation des endomorphismes et des matrices carrées

Définition 21. Soit f un endomorphisme d'un espace vectoriel de dimension finie $\boldsymbol{E}$. On dit que f est diagonalisable s'il vérifie l'une des conditions équivalentes suivantes.

① Il existe une base de vecteurs propres de f, c'est-à-dire une base dans laquelle la matrice de f est diagonale.

On vérifie sans peine que les valeurs propres de f sont les éléments diagonaux de la matrice.

② L'espace $\boldsymbol{E}$ est somme directe des sous-espaces propres de f.

L'équivalence de ces deux points est immédiate.

① ⇒ ② S'il existe une base de vecteurs propres de f, alors la somme directe des sous-espaces propres de f contient une base de $\boldsymbol{E}$, donc est égale à $\boldsymbol{E}$.

② ⇒ ① Si $\boldsymbol{E}$ est somme directe des sous-espaces propres de f, on obtient une base de vecteurs propres de f en rassemblant une base de chacun des sous-espaces propres de f.

Définition 22. On dit qu'une matrice carrée M, d'ordre n, est diagonalisable s'il existe un changement de base la rendant diagonale, c'est-à-dire s'il existe une matrice inversible P, d'ordre n, telle que :

$$D = P^{-1} MP$$

soit diagonale.

7.3.1 Diagonalisation dans un cas élémentaire

On donne maintenant une condition suffisante simple de diagonalisation, s'appuyant sur le résultat suivant dont on rappelle l'énoncé : « Les vecteurs propres associés à des valeurs propres distinctes d'un endomorphisme ou d'une matrice forment une famille libre, et les sous-espaces propres associés à des valeurs propres distinctes d'un endomorphisme ou d'une matrice sont en somme directe ».

Rappelons l'idée de la démonstration, qui se fait par récurrence sur le nombre n de vecteurs propres considérés.

Pour $n = 1$, un vecteur propre v_1 constitue une famille libre puisqu'il est non nul.

Si le résultat est vrai au rang $n - 1$, alors il est vrai au rang n puisque (en notant $v_1, \ldots, v_n$ les vecteurs propres associés aux valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$), on a :

$$\alpha_1 v_1 + \ldots + \alpha_{n-1} v_{n-1} + \alpha_n v_n = 0 .$$

Donc, en composant par f et en tenant compte de $f(v_i) = \lambda_i v_i$:

$$\alpha_1 \lambda_1 v_1 + \ldots + \alpha_{n-1} \lambda_{n-1} v_{n-1} + \alpha_n \lambda_n v_n = 0$$

d'où, par différence, en multipliant la première égalité par λ_n :

$$\alpha_1(\lambda_n - \lambda_1) v_1 + \ldots + \alpha_{n-1}(\lambda_n - \lambda_{n-1}) v_{n-1} = 0 .$$

L'hypothèse de récurrence donne alors :

$$\alpha_1(\lambda_n - \lambda_1) = \ldots = \alpha_{n-1}(\lambda_n - \lambda_{n-1}) = 0$$

et puisque les valeurs propres sont distinctes :

$$\alpha_1 = \ldots \alpha_{n-1} = 0 .$$

Il reste alors, dans la première égalité :

$$\alpha_n v_n = 0$$

et, comme un vecteur propre n'est pas nul, $\alpha_n = 0$, ce qui conclut la démonstration.

Proposition 32.

Tout endomorphisme f d'un espace de dimension n (ou toute matrice carrée d'ordre n) admettant n valeurs propres distinctes est diagonalisable.

Preuve. ◊ En effet, en choisissant n vecteurs propres $v_1, \ldots, v_n$ associés à ces n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$, on a, d'après le résultat précédent, une famille libre de n vecteurs propres, donc une base de vecteurs propres. ◊

On passe maintenant à des conditions nécessaires et suffisantes de diagonalisation.

Proposition 33.

Un endomorphisme f (ou une matrice carrée M) est diagonalisable si, et seulement si, elle vérifie les deux conditons suivantes :

a) *le polynôme caractéristique de f (ou de M) est scindé* ;

b) *pour toute valeur propre de λ de f (ou de M), la dimension du sous-espace propre associé à λ est égal à l'ordre de multiplicité de la valeur propre λ*.

La condition **a)** est nécessairement vérifiée lorsque $\mathbb{K} = \mathbb{C}$; par contre, dans le cas réel, elle signifie que les valeurs propres de f (ou de M) sont réelles.

La condition **b)** signifie que les sous-espaces propres sont, à chaque fois, de dimension maximale. En effet, on sait que, pour toute valeur propre λ, la dimension de $\mathrm{Ker}(f - \lambda\,\mathrm{Id})$ est inférieure ou égale à l'ordre de multiplicité de la valeur propre λ.

Preuve. ◊ Soit $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de f dans le corps $\mathbb{K}$ et $r_1, \ldots, r_p$ leurs ordres de multiplicité. L'endomorphisme f est diagonalisable si, et seulement si, $\boldsymbol{E}$ est la somme directe de ses sous-espaces propres, autrement dit si, et seulement si :

$$\dim \boldsymbol{E} = \sum_{i=1}^{p} \dim \mathrm{Ker}(f - \lambda_i \mathrm{Id}) .$$

Or, on a pour $1 \leqslant i \leqslant p$:

$$\dim \mathrm{Ker}(f - \lambda_i \mathrm{Id}) \leqslant r_i .$$

Par conséquent :

$$\sum_{i=1}^{p} \dim \mathrm{Ker}(f - \lambda_i \mathrm{Id}) \leqslant \sum_{i=1}^{p} r_i \leqslant n = \dim \boldsymbol{E} .$$

En effet, la somme des ordres de multiplicité des valeurs propres, c'est-à-dire des racines du polynôme caractéristique, est inférieure ou égale au degré de celui-ci qui est égal à $n = \dim \boldsymbol{E}$.

Ainsi, f est diagonalisable si, et seulement si, les inégalités :

$$\sum_{i=1}^{p} \dim \mathrm{Ker}(f - \lambda_i \mathrm{Id}) \leqslant \sum_{i=1}^{p} r_i \leqslant n = \dim \boldsymbol{E} ,$$

sont des égalités, ce qui équivaut à :

a) $\sum_{i=1}^{p} r_i = n = \dim \boldsymbol{E}$, ce qui signifie que :

$$P_f(x) = (\lambda_1 - x)^{r_1} (\lambda_2 - x)^{r_2} \ldots (\lambda_p - x)^{r_p}$$

autrement dit, le polynôme caractéristique est scindé dans $\mathbb{K}$.

b) $\dim \mathrm{Ker}(f - \lambda_i \mathrm{Id}) = r_i$ pour $1 \leqslant i \leqslant p$, ce qui achève la démonstration. ◊

On donne maintenant une seconde caractérisation des endomorphismes (ou matrices carrées diagonalisables), s'appuyant sur les deux résultats préliminaires établis ci-dessous.

Proposition 34.

Soient P un polynôme de $\mathbb{K}[X]$ *et f un endomorphisme d'un* $\mathbb{K}$*-espace vectoriel* $\boldsymbol{E}$*. Si P (f) = 0 (on dit que le polynôme P annule f), les valeurs propres de f sont nécessairement racines de P.*

Preuve. ◊ Posons $P(x) = a_p x^p + \ldots + a_1 x + a_0$.

Pour toute valeur propre λ de f existe un vecteur non nul v tel que $f(v) = \lambda v$. Par conséquent :

$$f^2(v) = \lambda f(v) = \lambda^2 v$$

et, par récurrence immédiate :

$$f^k(v) = \lambda^k v$$

pour tout nombre entier vectoriel k. Puisque $P(f) = 0$, on a :

$$\begin{aligned} 0 = P(f)(v) &= a_p f^p(v) + \ldots + a_1 f(v) + a_0 v \\ &= (a_p \lambda^p + \ldots + a_1 \lambda + a_0) v . \end{aligned}$$

Donc : $P(\lambda) = a_p \lambda^p + \ldots + a_1 \lambda + a_0 = 0$. ◊

Lemme des noyaux

Soient P_1, P_2, P_n, n polynômes deux à deux premiers entre eux et f un endomorphisme de $\boldsymbol{E}$. On a :

$$\mathrm{Ker}\, P_1 \ldots P_n(f) = \mathrm{Ker}\, P_1(f) \oplus \ldots \oplus \mathrm{Ker}\, P_n(f) .$$

En particulier, si $P_1 \ldots P_n$ annule f, alors :

$$\boldsymbol{E} = \mathrm{Ker}\, P_1(f) \oplus \ldots \oplus \mathrm{Ker}\, P_n(f) .$$

Preuve. ◊

■ Raisonnons par récurrence sur n, en commençant par le cas $n = 2$.

On dispose des deux polynôme P_1, P_2, qui sont premiers entre eux. D'après l'égalité de Bezout, existent deux polynômes U_1 et U_2 tels que :

$$U_1 P_1 + U_2 P_2 = 1 .$$

Substituons f à x :

$$U_1(f) \circ P_1(f) + U_2(f) \circ P_2(f) = \mathrm{Id} ,$$

et donc, pour tout vecteur v de $\boldsymbol{E}$:

$$U_1(f) \circ P_1(f)(v) + U_2(f) \circ P_2(f)(v) = v . \qquad \text{(A)}$$

Nous montrons maintenant que :

$$\mathrm{Ker}\ P_1P_2(f) = \mathrm{Ker}\ P_1(f) \oplus \mathrm{Ker}\ P_2(f) .$$

a) Vérifions que :

$$\mathrm{Ker}\ P_1(f) \cap \mathrm{Ker}\ P_2(f) = \{0\} .$$

L'égalité (A) donne en effet $v = 0$ dès lors que $P_1(f)(v) = 0$ et $P_2(f)(v) = 0$.

b) Établissons l'inclusion suivante :

$$\mathrm{Ker}\ P_1P_2(f) \subset \mathrm{Ker}\ P_1(f) + \mathrm{Ker}\ P_2(f) .$$

Pour tout vecteur v appartenant à $\mathrm{Ker}\ P_1P_2(f)(v)$, il est clair que $P_2(f)(v)$ appartient à $\mathrm{Ker}\ P_1(f)$ et $P_1(f)(v)$ appartient à $\mathrm{Ker}\ P_2(f)$.

L'égalité (A) exprime donc bien v comme somme de deux vecteurs appartenant respectivement à $\mathrm{Ker}\ P_2(f)$ et $\mathrm{Ker}\ P_1(f)$.

c) Établissons l'inclusion suivante :

$$\mathrm{Ker}\ P_1(f) + \mathrm{Ker}\ P_2(f) \subset \mathrm{Ker}\ P_1P_2(f) .$$

Soient v_1, v_2 deux vecteurs appartenant à $\mathrm{Ker}\ P_1(f), \mathrm{Ker}\ P_2(f)$; alors $v_1 + v_2$ appartient à $\mathrm{Ker}\ P_1P_2(f)$ car :

$$P_1P_2(f)(v_1 + v_2) = P_2(f) \circ P_1(f)(v_1) + P_1(f) \circ P_2(f)(v_2) = 0 .$$

Cela achève la démonstration dans le cas $n = 2$.

■ Supposons le résultat vrai au rang $n - 1$ et démontrons-le au **rang *n***.

Puisque $P_1, P_2, ..., P_n$, sont deux à deux premiers entre eux, P_n est donc premier avec P_1, avec P_2, ..., avec P_{n-1}, par conséquent avec leur produit $P_1P_2 \dots P_{n-1}$.

Le résultat obtenu pour $n = 2$ donne alors :

$$\mathrm{Ker}\ P_1 \dots P_n(f) = \mathrm{Ker}\ P_1 \dots P_{n-1}(f) \oplus \mathrm{Ker}\ P_n\ (f)$$

et d'après l'hypothèse de récurrence :

$$\mathrm{Ker}\ P_1 \dots P_n(f) = \mathrm{Ker}\ P_1(f) \oplus \dots \oplus \mathrm{Ker}\ P_n\ (f) . \qquad \Diamond$$

7.3.2 Caractérisation des endomorphismes ou des matrices diagonisables

Nous passons maintenant à la seconde condition nécessaire et suffisante de diagonalisation.

Proposition 35.

Un endomorphisme f (ou une matrice carrée M) est diagonalisable si, et seulement si, il annule un polynôme scindé dont les racines sont simples.

Preuve. ◊

■ Notons $\lambda_1, \lambda_2, ..., \lambda_p$ les valeurs propres distinctes de f. Si f est diagonalisable, on a :

$$\boldsymbol{E} = \mathrm{Ker}\ (f - \lambda_1\ \mathrm{Id}) \oplus \dots \oplus \mathrm{Ker}\ (f - \lambda_p\ \mathrm{Id}) .$$

Puisque $\lambda_1, \lambda_2, ..., \lambda_p$ sont distinctes, les polynômes $(x - \lambda_1)$, $(x - \lambda_2)$, ..., $(x - \lambda_p)$ sont deux à deux premiers entre eux et le lemme des noyaux donne :

$$\boldsymbol{E} = \mathrm{Ker}\ [(f - \lambda_1\ \mathrm{Id}) \circ \dots \circ (f - \lambda_p\ \mathrm{Id})] .$$

Ainsi, $(f - \lambda_1\ \mathrm{Id}) \circ \dots \circ (f - \lambda_p\ \mathrm{Id})$ est l'endomorphisme nul, ce qui revient à dire que f annule le polynôme scindé à racines simples $P(x) = (x - \lambda_1)\dots(x - \lambda_p)$.

■ Réciproquement, supposons que f annule le polynôme scindé à racines simples $P(x) = (x - \lambda_1)\dots(x - \lambda_q)$.

Comme les polynômes $(x - \lambda_1)\dots(x - \lambda_q)$ sont deux à deux premiers entre eux, le lemme des noyaux donne de même :

$$\boldsymbol{E} = \mathrm{Ker}\ P(f) = \mathrm{Ker}\ (f - \lambda_1\ \mathrm{Id}) \oplus \dots \oplus \mathrm{Ker}\ (f - \lambda_q\ \mathrm{Id}) .$$

Parmi ces sous-espaces, certains peuvent être réduits au vecteur nul et peuvent donc être supprimés de la somme. Quitte à changer la numérotation, il reste alors :

$$\boldsymbol{E} = \mathrm{Ker}\ (f - \lambda_1\ \mathrm{Id}) \oplus \dots \oplus \mathrm{Ker}\ (f - \lambda_p\ \mathrm{Id}) .$$

Ces scalaires $\lambda_1, ..., \lambda_p$ sont, bien entendu, des valeurs propres de f puisque les noyaux $\mathrm{Ker}\ (f - \lambda_k\ \mathrm{Id})$ sont non réduits au vecteur nul pour $1 \leqslant k \leqslant p$.

Réciproquement toute valeur propre de f est nécessairement racine du polynôme P et figure donc dans la liste $\lambda_1, ..., \lambda_p$. Ainsi donc, l'égalité précédente établit que $\boldsymbol{E}$ est somme directe des sous-espaces propres de f, donc que f est diagonalisable. ◊

Exemples.

① On sait qu'un projecteur p vérifie la relation $p^2 = p$. Il annule donc le polynôme scindé à racines simples $x(x - 1)$. A ce titre, il est diagonalisable et ses valeurs propres sont au plus 0 et 1. On retrouve donc le fait que la matrice d'un projecteur dans une base convenable est donc, si r est l'ordre de multiplicité de la valeur propre 1 :

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) .$$

On procèderait de même pour une symétrie avec le polynôme annulateur $x^2 - 1$.

② Reprenons la matrice $M = \begin{pmatrix} m & (1) \\ (1) & m \end{pmatrix}$.

Nous avons que les valeurs propres nécessairement réelles de cette matrice symétrique réelle sont : $m + n - 1$ avec la multiplicité 1 et $m - 1$ avec la multiplicité $n - 1$.

Il est clair que $v_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est vecteur propre associé à $\lambda = m + n - 1$.

Les vecteurs propres associés à $\lambda = m - 1$ vérifient la relation :

$$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (m-1) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ou, ce qui revient au même à : $x_1 + x_2 + \dots + x_n = 0$.

Il s'agit là de l'équation d'un hyperplan de $\mathbb{R}^n$, et le sous-espace propre associé à $\lambda = m - 1$ est de dimension $n - 1$.

La matrice M est donc diagonalisable, puisque la somme des dimensions des deux sous-espaces propres est n, et la matrice suivante est une matrice de passage de la base canonique à une base de vecteurs propres :

$$P = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ \vdots & -1 & & & \\ \vdots & & -1 & & (0) \\ \vdots & (0) & & \ddots & \\ 1 & & & & -1 \end{pmatrix} .$$

La matrice dans cette base de vecteurs propres est donc :

$$P^{-1}MP = \begin{pmatrix} m+n-1 & & \\ & m-1 & (0) \\ (0) & & \ddots \; m-1 \end{pmatrix}.$$

Cette égalité s'établissant bien entendu sans aucun calcul.

③ Considérons la matrice ligne non nulle $L = [\ell_1, ..., \ell_n]$ et la matrice colonne non nulle $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$.

Formons la matrice carrée M définie par :

$$M = CL = \begin{pmatrix} c_1\ell_1 & \dots & c_1\ell_n \\ \vdots & & \\ c_n\ell_1 & \dots & c_n\ell_n \end{pmatrix}.$$

Cette matrice est non nulle, donc son rang est au moins 1, et toutes ses colonnes sont proportionnelles à C, donc son rang est au plus 1. Ainsi :

$$\text{rang } M = 1 \text{ et } \dim \operatorname{Ker} M = n-1.$$

Le noyau Ker M n'est autre que le sous-espace propre associé à la valeur propre 0 dont l'ordre de multiplicité est, on le sait, supérieur ou égal à dim Ker $M = n - 1$; 0 est donc au moins $n - 1$ fois valeur propre de M.

Pour obtenir la dernière valeur propre, il suffit de se souvenir que la somme des valeurs propres n'est autre que la trace de la matrice qui vaut ici :

$$\operatorname{tr}(M) = \sum_{k=1}^{n} c_k\ell_k = LC.$$

Deux cas peuvent donc se présenter.

a) Si $LC \neq 0$, M est diagonalisable et semblable à la matrice :

$$D = \begin{pmatrix} LC & & (0) \\ & 0 & \\ (0) & & \ddots \; 0 \end{pmatrix}.$$

En effet, le sous-espace propre associé à 0 est de dimension $n - 1$ et le sous-espace propre associé à LC est la droite dirigée par C.

b) Si $LC = 0$, M n'est pas diagonalisable, sinon, elle serait semblable à la matrice nulle, donc nulle (car $P^{-1}\,0\,P = 0$), ce qui n'est pas.

7.4 Diagonalisation par blocs

Définition 23. Soit f un endomorphisme d'un espace vectoriel de dimension finie $\boldsymbol{E}$. On dit que f est diagonalisable s'il vérifie l'une des conditions équivalentes suivantes.

① Il existe une base $(e_1, ..., e_n)$ de $\boldsymbol{E}$ dans laquelle la matrice est diagonale par blocs c'est-à-dire de la forme :

$$M = \left(\begin{array}{c|c|c|c} M_1 & 0 & & 0 \\ \hline 0 & M_2 & & 0 \\ \hline & & & \\ \hline 0 & & 0 & M_p \end{array}\right).$$

② L'espace $\boldsymbol{E}$ est somme directe de sous-espaces stables par f.

L'équivalence de ces deux points est immédiate.

① ⇒ ② En notant $r_1, r_2, ..., r_p$, les dimensions des blocs figurant dans la matrice ci-dessus, on voit que les images des vecteurs $e_1, e_2, ..., e_{r_1}$, sont des combinaisons linéaires de $e_1, e_2, ..., e_{r_1}$; autrement dit, $\boldsymbol{E}_1 = \operatorname{Vect}(e_1, e_2, ..., e_{r_1})$ est stable par f.

On voit également, que $\boldsymbol{E}_2 = \operatorname{Vect}(e_{r_1+1}, e_{r_1+2}, ..., e_{r_1+r_2})$ est stable par f et, de même, avec des notations évidentes $\boldsymbol{E}_3, ..., \boldsymbol{E}_p$. La somme de ces sous-espaces stables $\boldsymbol{E}_1, \boldsymbol{E}_2, ..., \boldsymbol{E}_p$ est clairement directe et égale à $\boldsymbol{E}$.

② ⇒ ① Si $\boldsymbol{E}$ est somme directe des sous-espaces stables $\boldsymbol{E}_1, \boldsymbol{E}_2, ..., \boldsymbol{E}_p$, on a en réunissant des bases de $\boldsymbol{E}_1, \boldsymbol{E}_2, ..., \boldsymbol{E}_p$ une base de $\boldsymbol{E}$ dans laquelle la matrice de f est clairement diagonale par blocs.

Définition 24. On dit qu'une matrice carrée M d'ordre n est diagonalisable par blocs s'il existe un changement de base la rendant diagonale par blocs, c'est-à-dire s'il existe une matrice inversible P d'ordre n telle que :

$$D = P^{-1}MP$$

soit diagonale par blocs.

Avant de passer à l'énoncé du théorème central de ce paragraphe, nous rappelons le résultat suivant.

Proposition 36.

Soient f et g deux endomorphismes d'un espace vectoriel $\boldsymbol{E}$ *tels que* $f \circ g = g \circ f$. *Alors f laisse stable* Ker g *et* Im g.

En particulier, comme f commute avec tout polynôme P (f) en l'endomorphisme f, f laisse stable Ker $P(f)$ et Im $P(f)$.

Rappelons brièvement la preuve dans le cas de la stabilité de Ker g. Soit un vecteur v appartenant à Ker g. On a $g(v) = 0$, donc $f \circ g(v) = 0$ et également $g \circ f = 0$; ainsi $f(v)$ appartient à Ker g qui est bien stable par f.

Définition 25. Soient f un endomorphisme de $\boldsymbol{E}$ et λ une valeur propre de f d'ordre de multiplicité r. On appelle sous-espace caractéristique de f associé à λ le sous-espace $\operatorname{Ker}(f - \lambda\,\mathrm{Id})^r$. Il s'agit d'un sous-espace stable par f sur lequel l'endomorphisme induit par f est somme d'une homothétie de rapport λ et d'un endomorphisme nilpotent.

En effet, l'endomorphisme $(f - \lambda\,\mathrm{Id})^r$ est un polynôme en f, et donc f laisse stable son noyau. De plus, pour tout vecteur v appartenant à $\operatorname{Ker}(f - \lambda\,\mathrm{Id})^r$ on a :

$$(f - \lambda\,\mathrm{Id})^r(v) = 0.$$

Ainsi, l'endomorphisme induit par $(f - \lambda\,\mathrm{Id})$ sur $\operatorname{Ker}(f - \lambda\,\mathrm{Id})^r$ est nilpotent, ce qui justifie la définition précédente.

Comme on sait que, dans une base bien choisie, la matrice d'un endomorphisme nilpotent est triangulaire avec des zéros sur la diagonale, la matrice de l'endomorphisme induit par f sur le sous-espace caractéristique associé à λ a, dans une base convenable, la forme suivante :

$$\begin{pmatrix} \lambda & * & * \\ & \lambda & \\ & \ddots & * \\ (0) & & \lambda \end{pmatrix}.$$

Proposition 37.

Soit f un endomorphisme de $\boldsymbol{E}$. *L'endomorphisme f est scindé si, et seulement si, l'espace* $\boldsymbol{E}$ *est somme directe des sous-espaces* *caractéristiques de f. Dans ce cas, la matrice de f dans une base convenable de **E** a la forme :*

$$M = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ (0) & \lambda_1\end{matrix}} & & (0) \\ & \boxed{\begin{matrix}\lambda_2 & * \\ (0) & \lambda_2\end{matrix}} & \\ (0) & & \boxed{\begin{matrix}\lambda_p & * \\ (0) & \lambda_p\end{matrix}} \end{pmatrix}.$$

De plus, la dimension de chacun des sous-espaces caractéristiques est égale à l'ordre de multiplicité de la valeur propre associée.

Preuve. ◊

■ Notons $\lambda_1, ..., \lambda_p$, les valeurs propres distinctes de f et $r_1, ..., r_p$ leurs ordres de multiplicité. Le polynôme caractéristique de f est donc, si f est scindé :

$$P_f(x) = (\lambda_1 - x)^{r_1} ... (\lambda_p - x)^{r_p}.$$

Il résulte alors du théorème de Hamilton-Cayley que :

$$(\lambda_1 \,\mathrm{Id} - f)^{r_1} \circ ... \circ (\lambda_p \,\mathrm{Id} - f)^{r_p} = 0.$$

Comme $\lambda_1, ..., \lambda_p$ sont distinctes, les polynômes $(x - \lambda_1)^{r_1}, ..., (x - \lambda_p)^{r_p}$ sont deux à deux premiers entre eux et le lemme des noyaux donne :

$$\boldsymbol{E} = \mathrm{Ker}\,(f - \lambda_1 \,\mathrm{Id})^{r_1} \oplus ... \oplus \mathrm{Ker}\,(f - \lambda_p \,\mathrm{Id})^{r_p}.$$

L'espace est donc bien somme directe des sous-espaces caractéristiques de f.

■ Réciproquement, si l'espace est somme directe des sous-espaces caractéristiques de f, un choix convenable de bases dans chacun de ceux-ci (cf. remarque qui suit la définition des sous-espaces caractéristiques) donne par réunion une base de **E** dans laquelle la matrice de f est :

$$M = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ (0) & \lambda_1\end{matrix}} & & (0) \\ & \ddots & \\ (0) & & \boxed{\begin{matrix}\lambda_p & * \\ (0) & \lambda_p\end{matrix}} \end{pmatrix}.$$

La dimension de chacun des blocs est ici la dimension de chacun des sous-espaces caractéristiques, notée d_k. Le polynôme caractéristique de f est clairement scindé et égal à :

$$P_f(x) = \det(f - x\,\mathrm{Id}) = \det(M - xI_n) = \prod_{k=1}^{p} (\lambda_k - x)^{d_k}.$$

Il apparaît bien que d_k est l'ordre de multiplicité de la valeur propre λ_k ; autrement dit, la dimension du sous-espace caractéristique associé à la valeur propre λ_k est égale à l'ordre de multiplicité de la valeur propre λ_k. ◊

Exemple : considérons l'endomorphisme de $\mathbb{R}^4$ dont la matrice vaut :

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

Calculons le polynôme caractéristique de M.

$$P_f(x) = \det(M - x\,I_4)$$

$$= \begin{vmatrix} 1-x & -1 & 2 & -2 \\ 0 & -x & 1 & -1 \\ 1 & -1 & 1-x & 0 \\ 1 & -1 & 1 & -x \end{vmatrix}.$$

$$C_1 \leftarrow C_1 + C_2 = \begin{vmatrix} -x & -1 & 2 & -2 \\ -x & -x & 1 & -1 \\ 0 & -1 & 1-x & 0 \\ 0 & -1 & 1 & -x \end{vmatrix}$$

$$L_2 \leftarrow L_2 - L_1 = \begin{vmatrix} -x & -1 & 2 & -2 \\ 0 & -x+1 & -1 & 1 \\ 0 & -1 & 1-x & 0 \\ 0 & -1 & 1 & -x \end{vmatrix}$$

$$\text{Développement suivant } L_1 = (-x)\begin{vmatrix} -x+1 & -1 & 1 \\ -1 & 1-x & 0 \\ -1 & 1 & -x \end{vmatrix}$$

$$C_1 \leftarrow C_2 + C_1 = (-x)\begin{vmatrix} -x & -1 & 1 \\ -x & 1-x & 0 \\ 0 & 1 & -x \end{vmatrix}$$

$$L_2 \leftarrow -L_1 + L_2 = (-x)\begin{vmatrix} -x & -1 & 1 \\ 0 & 2-x & -1 \\ 0 & 1 & -x \end{vmatrix}$$

Développement suivant

$$L_1 = x^2 \begin{vmatrix} 2-x & -1 \\ 1 & -x \end{vmatrix} = x^2\,(x^2 - 2x + 1) = x^2\,(x-1)^2.$$

La matrice M a deux valeurs propres double 0 et 1.

Déterminons une base de Ker $(M)^2$ et de Ker $(M - I_4)^2$.

a) Base de Ker $(M)^2$

Le vecteur $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ appartient à Ker M, si, et seulement si :

$$\begin{cases} x - y + 2z - 2t = 0 \\ z - t = 0 \\ x - y + z = 0 \\ x - y + z = 0 \end{cases}$$

soit $\begin{cases} x - y + z = 0 \\ z - 2t = 0 \\ z - t = 0 \end{cases}$ ou encore $\begin{cases} x = y \\ z = t = 0 \end{cases}$.

On peut choisir $e_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ comme base de Ker M.

Un vecteur $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ appartient à Ker M^2 si, et seulement si :

$$\begin{cases} x - y + 2z - 2t = \lambda \\ z - t = \lambda \\ x - y + z = 0 \\ x - y + z = 0 \end{cases}$$

Cela nous donne une base de Ker M^2 sous la forme :

$$e_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

avec $f(e_2) = e_1$ et $f(e_1) = 0$.

Exemple : b) Base de Ker $(M - I_4)^2$

On peut choisir :

$$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \text{ et } e_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

On a $f(e_4) = e_3 + e_4$ et $f(e_3) = e_3$.

La matrice de f dans la nouvelle base (e_1, e_2, e_3, e_4) s'exprime sous la forme :

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

et en posant $P = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ on a l'égalité suivante :

$$M = P\,T\,P^{-1}.$$