Algèbre bilinéaire et Intégration

par

Emmanuel Hebey

Université de Cergy-Pontoise

Année 2019-2020

Chapitre 3 Diagonalisation

On continue de ne considérer que le cas d'espaces vectoriels réels. Là encore (là surtout) une théorie analogue existe pour les espaces vectoriels complexes. On commence par traiter du point de vue des endomorphismes, pour traiter ensuite du point de vue des matrices. La question générale qui est posée est la suivante :

Question

Etant donné un endomorphisme f d'un $\mathbb{R}$-espace vectoriel E de dimension finie, parmi toutes les représentations possibles $M_{\mathcal{B}\mathcal{B}}(f)$ de f, où $\mathcal{B}$ est une base de E, y en a-t-il une qui soit plus pertinente, plus esthétique, plus facile à manier que les autres ?

Dans une théorie de diagonalisation, la représentation plus pertinente que l'on cherche est une représentation diagonale. La question devient donc : "parmi toutes les représentations possibles $M_{\mathcal{B}\mathcal{B}}(f)$ de f, y en a-t-il une qui soit diagonale" ?

1. Le cas des endomorphismes

Définition

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, et $f \in End(E)$ un endomorphisme de E. On dit que $\lambda \in \mathbb{R}$ est une valeur propre de f s'il existe un vecteur $u \in E$, $u \neq 0$, tel que

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

Dans ce cas, u est dit vecteur propre de f associé à la valeur propre λ. Si λ est une valeur propre de f, on note E_λ l'espace propre de f associé à λ, défini par $E_\lambda = \left\{ u \in E \ / \ f(u) = \lambda u \right\}$.

Proposition

E_λ est un sous espace vectoriel de E.

Preuve : Soient $x, y \in E_\lambda$ et $t \in \mathbb{R}$ quelconques, alors $f(x + ty) = f(x) + tf(y) = \lambda x + t\lambda y = \lambda(x + ty)$ et donc $x + ty \in E_\lambda$. CQFD.

Théorème

Un réel λ est valeur propre d'un endomorphisme f si et seulement si l'endomorphisme $f - \lambda Id_E$ n'est pas inversible. Dans ce cas,

$$E_\lambda = Ker\big(f - \lambda Id_E\big) \ ,$$

où $f - \lambda Id_E$ est l'endomorphisme de E défini pour tout $x \in E$ par $\big(f - \lambda Id_E\big)(x) = f(x) - \lambda x$.

Preuve : Il suffit de se souvenir qu'un endomorphisme $g \in \mathrm{End}(E)$ (ici $g = f - \lambda \mathrm{Id}_E$) est inversible si et seulement si il est injectif, et donc si et seulement si $\mathrm{Ker}(g) = \{0\}$. Et bien évidemment dire qu'il existe $u \neq 0$ tel que $f(u) = \lambda u$ équivaut à dire que $\mathrm{Ker}\big(f - \lambda \mathrm{Id}_E\big) \neq \{0\}$. Il est ensuite facile de voir que $E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$. CQFD.

Définition

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie n et $f \in End(E)$ un endomorphisme de E. Soit $\mathcal{B}$ une base de E. On appelle polynôme caractéristique de f le polynôme

$$P(X) = det\big(M_{\mathcal{B}\mathcal{B}}(f) - X Id_n\big)$$

pour tout $X \in \mathbb{R}$.

Théorème

Le polynôme caractéristique de f ne dépend pas du choix de la base $\mathcal{B}$. C'est un polynôme de degré n dont le terme dominant est $(-1)^n X^n$. Il est suivi du terme $(-1)^{n-1} tr(f) X^{n-1}$, où $tr(f)$ est la trace de l'endomorphisme f, et son terme constant vaut $det\big(M_{\mathcal{B}\mathcal{B}}(f)\big)$, une quantité qui ne dépend pas de $\mathcal{B}$.

Puisque la quantité ne dépend pas du choix de $\mathcal{B}$, notons $\det(f)$ pour $\det\big(M_{\mathcal{B}\mathcal{B}}(f)\big)$. On a alors

$$P(X) = (-1)^n X^n + (-1)^{n-1}\mathrm{tr}(f) X^{n-1} + a_{n-2}X^{n-2} + \cdots + a_1 X + \det(f)$$

pour tout $X \in \mathbb{R}$, où les $a_1, a_2, \ldots, a_{n-2}$ sont des réels qui ne dépendent que de f.

Preuve : (1) On montre que P ne dépend pas de $\mathcal{B}$ et donc que $\det\big(M_{\mathcal{B}\mathcal{B}}(f)\big) = P(0)$ ne dépend pas non plus de $\mathcal{B}$. Si $\mathcal{B}, \mathcal{B}'$ sont deux bases de E, et $\lambda \in \mathbb{R}$, alors

$$\det\big(M_{\mathcal{B}\mathcal{B}}(f) - \lambda \mathrm{Id}_n\big) = \det\big(M_{\mathcal{B}\mathcal{B}}(f - \lambda \mathrm{Id}_E)\big) \ ,$$
$$\det\big(M_{\mathcal{B}'\mathcal{B}'}(f) - \lambda \mathrm{Id}_n\big) = \det\big(M_{\mathcal{B}'\mathcal{B}'}(f - \lambda \mathrm{Id}_E)\big) \ .$$

On remarquera ici que $M_{\mathcal{B}\mathcal{B}}(\mathrm{Id}_E) = \mathrm{Id}_n$ pour toute base $\mathcal{B}$. Si $g \in \mathrm{End}(E)$ est un endomorphisme de E on a, cf. Chapitre 2, que

$$M_{\mathcal{B}'\mathcal{B}'}(g) = M_{\mathcal{B}\to\mathcal{B}'}^{-1} M_{\mathcal{B}\mathcal{B}}(g) M_{\mathcal{B}\to\mathcal{B}'} \ .$$

On en déduit que

Preuve suite :

$$\begin{aligned}\det\left(M_{\mathcal{B}'\mathcal{B}'}(g)\right) &= \det\left(M^{-1}_{\mathcal{B}\to\mathcal{B}'}\right)\det\left(M_{\mathcal{B}\mathcal{B}}(g)\right)\det\left(M_{\mathcal{B}\to\mathcal{B}'}\right)\\ &= \frac{1}{\det\left(M_{\mathcal{B}\to\mathcal{B}'}\right)}\det\left(M_{\mathcal{B}\mathcal{B}}(g)\right)\det\left(M_{\mathcal{B}\to\mathcal{B}'}\right)\\ &= \det\left(M_{\mathcal{B}\mathcal{B}}(g)\right)\ .\end{aligned}$$

D'où l'affirmation que ni P ni $\det\big(M_{\mathcal{B}\mathcal{B}}(f)\big)$ ne dépendent de $\mathcal{B}$.

(2) On montre que P est bien un polynôme de degré n et on calcule les coefficients a_n, a_{n-1} et a_0 de P. On a

$$P(X) = \sum_{\sigma\in\mathcal{P}_n} \varepsilon(\sigma)\big(a_{1\sigma(1)} - X\delta_{1\sigma(1)}\big)\ldots\big(a_{n\sigma(n)} - X\delta_{n\sigma(n)}\big)\ ,$$

où les a_{ij} sont les coefficients de $M_{\mathcal{B}\mathcal{B}}(f)$. Donc P est bien un polynôme de degré n. Le terme en X^n et aussi le terme en X^{n-1} sont forcément donnés par la permutation $\sigma =$ identité. En effet, sinon (si $\sigma \neq$ identité) alors au moins deux éléments bougent, donc il existe $i \neq j$ tels que $\sigma(i) \neq i$ et $\sigma(j) \neq j$, donc $\delta_{i\sigma(i)} = \delta_{j\sigma(j)} = 0$ pour ces deux i et j, et alors le produit correspondant $\big(a_{1\sigma(1)} - X\delta_{1\sigma(1)}\big)\ldots\big(a_{n\sigma(n)} - X\delta_{n\sigma(n)}\big)$ est au plus de degré $n-2$.

Preuve suite et fin : Une fois identifié que les termes d'ordres n et $n-1$ proviennent de $\sigma =$ identité on a que

$$\big(a_{1\sigma(1)} - X\delta_{1\sigma(1)}\big)\dots\big(a_{n\sigma(n)} - X\delta_{n\sigma(n)}\big) = \big(a_{11} - X\big)\dots\big(a_{nn} - X\big)$$

et on voit que le terme d'ordre n est $(-1)^n X^n$ tandis que le terme d'ordre $n-1$ est $(-1)^{n-1}\big(\sum_{i=1}^{n} a_{ii}\big)X^{n-1}$. Il est enfin facile de voir que le terme constant a_0 dans P est $P(0)$, qui vaut $\det\big(M_{\mathcal{B}\mathcal{B}}(f)\big)$. D'où le théorème. CQFD

Théorème

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie n, $f \in End(E)$ un endomrophisme de E et P son polynôme caractéristique. Alors $\lambda \in \mathbb{R}$ est valeur propre de f si et seulement si λ est racine de P (donc si et seulement si $P(\lambda) = 0$). En particulier, f a au plus n valeurs propres distinctes.

Preuve : On sait que λ est valeur propre de f si et seulement si $f - \lambda \mathrm{Id}_E$ n'est pas inversible. Etant donné $\mathcal{B}$ une base de E, un

Preuve suite : endomorphisme $g \in \mathrm{End}(E)$ est inversible si et seulement si $\det(M_{\mathcal{B}\mathcal{B}}(g)) \neq 0$. Donc $\lambda \in \mathbb{R}$ est valeur propre de f si et seulement si $\det(M_{\mathcal{B}\mathcal{B}}(f) - \lambda \mathrm{Id}_n) = 0$, ce qui revient à dire que λ est une racine de P. Un polynôme de degré n ayant au plus n racines distinctes, le théorème est démontré. CQFD.

Remarque : une grosse différence avec la théorie complexe est que $\mathbb{C}$ est un corps algébriquement clos. En d'autre termes, un polynôme se factorise toujours dans $\mathbb{C}$ et a donc toujours n racines (distinctes ou pas) dans $\mathbb{C}$, alors qu'un polynôme peut très bien n'avoir aucune racine dans $\mathbb{R}$ (c'est la cas de $P(X) = X^2 + 1$). Un endomrophisme peut du coup ne pas avoir de valeurs propres dans $\mathbb{R}$. Soit f l'endomorphisme de $\mathbb{R}^2$ dont la matrice de représentation $M_{\mathcal{B}\mathcal{B}}(f)$ dans la base canonique de $\mathbb{R}^2$ vaut

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

Cet endomorphisme a pour polynôme caractéristique $P(X) = X^2 + 1$, et il n'a donc pas de valeurs propres dans $\mathbb{R}$.

Soit $A = (a_{ij})$ une matrice. On dit que A est diagonale si $a_{ij} = 0$ dès que $i \neq j$. Donc A est une matrice diagonale si A ne comporte que des termes diagonaux...

Définition

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, et soit $f \in End(E)$ un endomorphisme de E. L'endomorphisme f est dit diagonalisable s'il existe une base $\mathcal{B}$ de E pour laquelle $M_{\mathcal{B}\mathcal{B}}(f)$ est une matrice diagonale.

Remarque : Si f est diagonalisable, alors le polynôme caractéristique P de f a toutes ses racines dans $\mathbb{R}$. En effet, en travaillant dans la base $\mathcal{B}$ qui diagonalise f, et si on note $\lambda_1, \ldots, \lambda_n$ les termes diagonaux de $M_{\mathcal{B}\mathcal{B}}(f)$, on récupère que

$$P(X) = (\lambda_1 - X) \ldots (\lambda_n - X) \ .$$

Donc P a bien toutes ses racines dans $\mathbb{R}$ et ces racines sont les termes diagonaux λ_i de f (qui peuvent très bien être égaux entre eux pour certains d'entre eux).

Théorème (Théorème fondamental de la théorie de la diagonalisation sur $\mathbb{R}$)

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, et soit $f \in End(E)$ un endomorphisme de E. Soit de plus $\lambda_1, \ldots, \lambda_k$ les valeurs propres distinctes de f (donc les racines distinctes dans $\mathbb{R}$ du polynôme caractéristique de f), et soit E_{λ_i}, $i = 1, \ldots, k$, les espaces propres correspondant. La somme $E_{\lambda_1} + \cdots + E_{\lambda_k}$ est toujours dirècte, donc on a toujours

$$E_{\lambda_1} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \ ,$$

et f est diagonalisable si et seulement si E est somme (dirècte) des E_{λ_i}, $i = 1, \ldots, k$, donc si et seulement si $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

On démontre un petit résultat préliminaire sur les sommes dirèctes avant de démontrer le théorème.

Lemme

Une somme de sous espaces vectoriels $E_1, \dots E_k$ est dirècte si et seulement si $dim\,(E_1 + \cdots + E_k) = \sum_{i=1}^k dim(E_i)$.

Preuve : Supposons tout d'abord que la somme est dirècte. Si $k = 2$ le résultat suit de ce qui a été dit dans le chapitre 1 puisque

$$\dim(E_1 + E_2) = \dim(E_1) + \dim(E_2) - \dim(E_1 \cap E_2) \ ,$$

et puisque la somme de E_1 et E_2 est dirècte si et seulement si $E_1 \cap E_2 = \{0\}$. On peut alors établir une preuve par récurrence en remarquant que si la somme des $E_1, \dots E_k$ est dirècte alors, en particulier, $E_k \cap (E_1 + \cdots + E_{k-1}) = \{0\}$ et donc la somme de $E_1 + \cdots + E_{k-1}$ et de E_k est dirècte. Par suite,

$$\dim\,(E_1 + \cdots + E_k) = \dim\,(E_1 + \cdots + E_{k-1}) + \dim(E_k) \ .$$

La récurrence sur k se met facilement en place à partir de cette relation.

Réciproquement supposons que

$$\dim(E_1 + \cdots + E_k) = \sum_{i=1}^{k} \dim(E_i) \ .$$

Soit $\mathcal{B}_i$ une base de E_i et soit (avec abus de notation) la famille

$$\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k) \ .$$

Cette famille est clairement génératrice pour $E_1 + \cdots + E_k$ (tout vecteur de $E_1 + \cdots + E_k$ se décompose en somme de vecteurs des E_i, qui eux-mêmes se décomposent dans les $\mathcal{B}_i$). Or, d'après la relation de départ, $\mathcal{B}$ a autant de vecteurs que la dimension de $E_1 + \cdots + E_k$. C'est donc une base de $E_1 + \cdots + E_k$. On en déduit facilement que tout vecteur de $E_1 + \cdots + E_k$ se décompose de façon unique en somme de vecteurs des E_i car sinon on aurait des décompositions différentes dans $\mathcal{B}$, ce qui est impossible par définition d'une base. CQFD

Preuve du théorème : (1) On commence par montrer que la somme des E_{λ_i} est dirècte. Sans perdre en généralité, on peut supposer que $k \geq 2$. On doit alors montrer que pour tout $i = 2, \ldots, k$,

$$E_{\lambda_i} \cap \big(\sum_{j<i} E_{\lambda_j}\big) = \{0\} \ .$$

On vérifie facilement que $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Si ce n'était pas le cas, il existerait $u \neq 0$ tel que $u \in E_{\lambda_1} \cap E_{\lambda_2}$. Or $u \in E_{\lambda_1}$ entraîne que $f(u) = \lambda_1 u$, tandis que $u \in E_{\lambda_2}$ entraîne que $f(u) = \lambda_2 u$. Comme $\lambda_1 \neq \lambda_2$, ces deux relations ne peuvent avoir lieu simultanément. La somme $E_{\lambda_1} \oplus E_{\lambda_2}$ est donc dirècte. On montre maintenant que la somme $E_{\lambda_1} \oplus E_{\lambda_2} \oplus E_{\lambda_3}$ est dirècte. Sans perdre en généralité on pourra supposer que

(i) soit $|\lambda_1| \leq |\lambda_2| < |\lambda_3|$,

(ii) soit $|\lambda_1| < |\lambda_2| = |\lambda_3|$.

On veut montrer que

$$E_{\lambda_3} \cap (E_{\lambda_1} + E_{\lambda_2}) = \{0\} \ .$$

Preuve suite : Soit u un vecteur dans l'intersection, donc $u \in E_{\lambda_3}$, et il existe $u_1 \in E_{\lambda_1}$ et $u_2 \in E_{\lambda_2}$ tels que $u = u_1 + u_2$. Alors $f(u) = f(u_1) + f(u_2)$, et donc $\lambda_3 u = \lambda_1 u_1 + \lambda_2 u_2$. En appliquant de nouveau f à cette égalité, et ainsi de suite, on arrive à $\lambda_3^k u = \lambda_1^k u_1 + \lambda_2^k u_2$ pour tout k entier. En particulier,

$$u = \left(\frac{\lambda_1}{\lambda_3}\right)^k u_1 + \left(\frac{\lambda_2}{\lambda_3}\right)^k u_2 \ .$$

Dans le cas (i), en faisant $k \to +\infty$ on obtient que $u = 0$. D'ou $E_{\lambda_3} \cap (E_{\lambda_1} + E_{\lambda_2}) = \{0\}$. Dans le cas (ii), en prenant $k = 2p$ et en faisant tendre $k \to +\infty$ on obtient que $u = u_2$. Or $E_{\lambda_2} \cap E_{\lambda_3} = \{0\}$ pour les mêmes raisons que $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Dans les deux cas on en déduit que $u = 0$. Donc

$$E_{\lambda_1} + E_{\lambda_2} + E_{\lambda_3} = E_{\lambda_1} \oplus E_{\lambda_2} \oplus E_{\lambda_3} \ .$$

La relation générale s'obtient de la même façon.

Preuve suite : (2) On montre que f est diagonalisable ssi

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$$

est somme dirècte des sous espaces propres de f. Cette condition est nécessaire dans la mesure où si f est diagonalisable, alors il existe une base $\mathcal{B}$ de E formée de vecteurs propres, donc de vecteurs dans la somme des E_{λ_i}. En effet, si on note $\mathcal{B} = (e_1, \ldots, e_n)$, et si on note λ'_i les valeurs propres de f pour $i = 1, \ldots, n$ (avec répetition des valeurs propres égales puisque $\{\lambda'_1, \ldots, \lambda'_n\} = \{\lambda_1, \ldots, \lambda_k\}$) alors la relation

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} \lambda'_1 & 0 & \ldots & 0 \\ 0 & \lambda'_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda'_n \end{pmatrix}$$

entraîne que $f(e_i) = \lambda'_i e_i$ pour tout i, et donc que les e_i sont tous des vecteurs propres de f. Donc ils sont tous dans la somme $E_{\lambda_1} + \cdots + E_{\lambda_k}$. D'où $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$ (et on sait que la somme est dirècte d'après ce qui a été dit plus haut).

Preuve suite et fin : On montre qu'à l'inverse, la condition est suffisante. Pour le voir, on considère $\mathcal{B}_i = (e_1^i, \dots, e_{m_i}^i)$ une base de E_{λ_i}, $i = 1, \dots, k$, et on pose

$$\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$$

la famille constitué des vecteurs de $\mathcal{B}_1$ suivis des vecteurs de $\mathcal{B}_2$, etc., suivis des vecteurs de $\mathcal{B}_k$. Puisque

$$E = E_{\lambda_1} + \cdots + E_{\lambda_k} \ ,$$

la famille $\mathcal{B}$ est une famille génératrice de E (tout vecteur de E s'écrit comme somme de vecteurs des E_{λ_i} qui eux-mêmes sécrivent comme combinaison linéaire de vecteurs de $\mathcal{B}$... de sorte que tout vecteur de E s'écrit comme combinaison linéaire de vecteurs de $\mathcal{B}$). Par ailleurs, puisque la somme est dirècte, $dimE = \sum_{i=1}^k dimE_{\lambda_i}$ de sorte que $\mathcal{B}$ a exactement $dimE$ vecteurs. Donc $\mathcal{B}$ est une base de E. Comme elle est constituée uniquement de vecteurs propres, f est de ce fait diagonalisable. D'où le théorème. CQFD

Corollaire

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n, et soit $f \in End(E)$ un endomorphisme de E. Soit de plus $\lambda_1, \ldots, \lambda_k$ les racines distinctes dans $\mathbb{R}$ du polynôme caractéristique P de f, et soit E_{λ_i}, $i = 1, \ldots, k$, les espaces propres correspondant. Alors f est diagonalisable si et seulement si

$$dimE = \sum_{i=1}^{k} dimE_{\lambda_i} \ .$$

En particulier, si P a n racines distinctes dans $\mathbb{R}$, avec $n = dimE$, alors f est diagonalisable.

Preuve : Puisque la somme est dirècte, cf le lemme précédent,

$$\dim\left(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}\right) = \sum_{i=1}^{k} \dim E_{\lambda_i} \ .$$

Si f est diagonalisable alors l'égalité des dimensions est vraie.

Preuve suite : Réciproquement si $\dim E = \sum_{i=1}^{k} \dim E_{\lambda_i}$, on en déduit que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Donc f est diagonalisble d'après le théorème précédent. Si maintenant $k = n$, sachant que $\dim E_{\lambda_i} \geq 1$ pour tout i, on trouve que $\sum_{i=1}^{n} \dim E_{\lambda_i} \geq n$. Comme on a aussi que

$$\sum_{i=1}^{n} \dim E_{\lambda_i} = \dim\big(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n}\big) \leq dimE$$

il suit que $\dim E = \sum_{i=1}^{n} \dim E_{\lambda_i}$, ce qui démontre le corollaire. En particulier, on a aussi dans ce cas que $\dim E_{\lambda_i} = 1$ pour tout $i = 1, \ldots, n$. CQFD

Lemme

Soit f un endomorphisme de E et $E_1, \ldots, E_k$ ses espaces propres. Alors f est diagonalisable si et seulement si pour toutes bases $\mathcal{B}_i$ de E_i, la famille $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ constituée des vecteurs de $\mathcal{B}_1$ suivis des vecteurs de $\mathcal{B}_2$ etc. est une base de E.

Preuve : Soient $\mathcal{B}_1, \ldots, \mathcal{B}_k$ des bases quelconques des $E_1, \ldots, E_k$.

Si $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ est une base de E alors la matrice $M_{\mathcal{B}\mathcal{B}}(f)$ de f dans $\mathcal{B}$ est forcément diagonale. Les λ_i (valeurs propres des E_i) sont répétées autant de fois qu'il y a de vecteurs dans les $\mathcal{B}_i$, et donc autant de fois que la dimension des E_i.

Si f est diagonalisable alors

$$E = E_1 \oplus \cdots \oplus E_k \ .$$

La famille $\mathcal{B}$ est clairement génératrice pour $E_1 \oplus \cdots \oplus E_k$. Elle est donc aussi génératrice pour E. Elle a autant de vecteurs que $\sum_{i=1}^{k} \dim(E_i)$ et donc (cf. le lemme précédent) autant de vecteurs que $\dim(E_1 \oplus \cdots \oplus E_k) = \dim(E)$. C'est donc une base de E. CQFD

Exercice : Soient E un $\mathbb{R}$-espace vectoriel de dimensions finie n et $f, g \in End(E)$ deux endomorphismes de E. On suppose que f est diagonalisable. Montrer que f et g commutent, à savoir $f \circ g = g \circ f$, si et seulement si les espaces propres de f sont stables par g, à savoir si et seulement si pour tout espace propre E_i de f on a $g(E_i) \subset E_i$.

Solution : Supposons que f et g commutent. Soit E_i un espace propre de f associé à une valeur propre λ_i. Soit $u \in E_i$. On a

$$g\left(f(u)\right) = \lambda_i g(u) = f\left(g(u)\right) \ .$$

Donc $v = g(u)$ appartient à E_i. Soit $g(E_i) \subset E_i$ puisque u est quelconque dans E_i. Si f et g commutent les espaces propres de f sont donc stables par g. Réciproquement supposons que les espaces propres de f sont stables par g. Puisque f est diagonalisable il existe une base $\mathcal{B}$ de E constituée de vecteurs propres. Disons que f a pour espaces propres $E_1, \ldots, E_p$ et écrivons (avec abus de langage) que $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ où les $\mathcal{B}_i$ sont des bases des E_i. Fixons i quelconque, notons λ_i la valeur propre pour E_i et notons $\mathcal{B}_i = (e_1^i, \ldots, e_k^i)$. Comme $g(E_i) \subset E_i$, il existe des α_{jm}, $j, m = 1, \ldots, k$, tels que pour tout $j = 1, \ldots, k$,

$$g(e_j^i) = \sum_{m=1}^{k} \alpha_{jm} e_m^i \ .$$

Par suite, pour tout $j = 1, \ldots, k$,

$$f\left(g(e_j^i)\right) = \sum_{m=1}^{k} \alpha_{jm} f(e_m^i) = \lambda_i \sum_{m=1}^{k} \alpha_{jm} e_m^i$$

tandis que

$$g\left(f(e_j^i)\right) = \lambda_i g(e_j^i) = \lambda_i \sum_{m=1}^{k} \alpha_{jm} e_m^i \ .$$

On voit donc que pour tout $i = 1, \ldots, p$ et tout $j = 1, \ldots, k$,

$$f\left(g(e^i_j)\right) = g\left(f(e^i_j)\right) \ .$$

En d'autres termes, il existe une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E ayant la propriété que pour tout $i = 1, \ldots, n$,

$$g \circ f(e_i) = f \circ g(e_i) \ .$$

Deux applications linéaires qui sont égales sur une base le sont sur tout l'espace. Donc $g \circ f = f \circ g$. $\square$

2. Ordre de multiplicité des racines et dimensions des espaces propres

Théorème

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, et soit $f \in End(E)$ un endomorphisme de E. Soit P le polynôme caractéristique de f et $\lambda \in \mathbb{R}$ une valeur propre de f. On suppose que l'ordre de multiplicité de λ en tant que racine de P est k, et donc que

$$P(X) = (X - \lambda)^k Q(X)$$

où Q est un polynôme de degré $n - k$ avec $Q(\lambda) \neq 0$. Soit E_λ l'espace propre associé à la valeur propre λ. Alors forcément $dimE_\lambda \leq k$.

On démontre ce résultat en montrant tout d'abord que le lemme suivant a lieu.

Lemme

Soit M une matrice carrée d'ordre n. On suppose que M est du type

$$M = \begin{pmatrix} \lambda Id_p & A \\ 0 & B \end{pmatrix}$$

où $p < n$, $\lambda \in \mathbb{R}$, Id_p est la matrice identité d'ordre p, A est une matrice $p \times (n-p)$, B est une matrice $(n-p) \times (n-p)$ et 0 est la matrice nulle $(n-p) \times p$. Alors $det(M) = \lambda^p det(B)$.

Solution : Pour "coder" la matrice M facilement fixons n et p (on sort donc d'un raisonnement mathématique général) et supposons par exemple que $n = 8$ et $p = 4$. Alors M est du type :

$$M = \begin{pmatrix} \lambda & 0 & 0 & 0 & a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & \lambda & 0 & 0 & a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & \lambda & 0 & a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & \lambda & a_{41} & a_{42} & a_{43} & a_{44} \\ 0 & 0 & 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & 0 & 0 & 0 & b_{31} & b_{32} & b_{33} & b_{34} \\ 0 & 0 & 0 & 0 & b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$$

En développant cette matrice suivant la première colonne on voit que

$$\det M = \lambda \det \begin{pmatrix} \lambda & 0 & 0 & a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & \lambda & 0 & a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & \lambda & a_{41} & a_{42} & a_{43} & a_{44} \\ 0 & 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & 0 & 0 & b_{31} & b_{32} & b_{33} & b_{34} \\ 0 & 0 & 0 & b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$$

On redéveloppe suivant la première colonne et on obtient que

$$\det M = \lambda^2 \det \begin{pmatrix} \lambda & 0 & a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & \lambda & a_{41} & a_{42} & a_{43} & a_{44} \\ 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & 0 & b_{31} & b_{32} & b_{33} & b_{34} \\ 0 & 0 & b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$$

Encore deux développements suivant la première colonne et on obtient que

$$\det M = \lambda^4 \det \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$$

Soit

$$\det(M) = \lambda^4 \det(B)$$

La preuve générale fonctionne de la même façon. On développe p fois suivant la première colonne pour obtenir que

$$\det(M) = \lambda^p \det(B)\ .$$

Une preuve mathématique rigoureuse par récurrence peut être mise en place. D'où le lemme. $\square$

On revient maintenant à la preuve du théorème. On suppose que λ est une racine d'ordre k du polynôme caractéristique. On veut montrer que $\dim E_\lambda \leq k$. On peut définir cet ordre de deux façons différentes : soit comme l'entier k pour lequel on peut écrire que $P(X) = (X - \lambda)^k Q(X)$ avec Q un polynôme réel de degré $n - k$ vérifiant que $Q(\lambda) \neq 0$, ou alors comme le plus grand entier p pour lequel on puisse écrire que $P(X) = (X - \lambda)^p Q(X)$ avec Q un polynôme réel de degré $n - p$. L'équivalence de ces deux définitions suit de la remarque que λ est racine d'un polynôme Q si et seulement si il existe un polynôme $\tilde{Q}$ avec $Q(X) = (X - \lambda)\tilde{Q}(X)$.

Preuve du théorème : Notons $p = \dim E_\lambda$. Si $p = n$ alors $E_\lambda = E$ et donc $f = \lambda \mathrm{Id}_E$ (i.e $f(x) = \lambda x$ pour tout $x \in E$). Mais alors $P(X) = (\lambda - X)^n$, donc $k = n$ et on a bien que $p \leq k$ (en fait $n = n$). On suppose maintenant que $p < n$. Soit $(e_1, \ldots, e_p)$ une base de E_λ. On la complète par des vecteurs $e_{p+1}, \ldots, e_n$ pour obtenir une base $\mathcal{B} = (e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ de E. On a alors

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} \lambda \mathrm{Id}_p & A \\ 0 & B \end{pmatrix}$$

où Id_p est la matrice identité d'ordre p, A est une matrice $p \times (n-p)$ et B est une matrice $(n-p) \times (n-p)$. La matrice $M_{\mathcal{B}\mathcal{B}}(f) - X\mathrm{Id}_n$ s'écrit alors

$$M_{\mathcal{B}\mathcal{B}}(f) - X\mathrm{Id}_n = \begin{pmatrix} (\lambda - X)\mathrm{Id}_p & A \\ 0 & B - X\mathrm{Id}_{n-p} \end{pmatrix}$$

Avec le lemme précédent : $P(X) = (\lambda - X)^p Q(X)$, où Q est le polynôme caractéristique de B. Donc $p \leq k$ puisqu'on peut voir k comme le plus grand de tels p. D'où le théorème. $\square$

On part avec les données suivantes :
E un $\mathbb{R}$-espace vectoriel de dimension n,
$\mathcal{B}$ une base de E, $f \in \mathrm{End}(E)$ et $A = M_{\mathcal{B}\mathcal{B}}(f)$.

$$\downarrow$$

Calcul du polynôme caractéristique
$$P(X) = det\big(A - X\mathrm{Id}_n\big)$$
de f où Id_n est la matrice identité $n \times n$

$$\downarrow$$

Recherche des racines de P dans $\mathbb{R}$

$\downarrow$ $\qquad\qquad\qquad\qquad$ $\downarrow$

Il y a des racines de P qui sont dans $\mathbb{C}$ mais pas dans $\mathbb{R}$. Alors f n'est pas diagonalisable.

P a n racines distinctes dans $\mathbb{R}$. Alors f est diagonalisable.

$\downarrow$

Sinon soient $\lambda_1, \ldots, \lambda_k$ les racines distinctes de P dans $\mathbb{R}$. On détermine les bases $\mathcal{B}_i$ des sous espaces propres définis pour $i = 1, \ldots, k$ par

$$E_{\lambda_i} = \mathrm{Ker}(f - \lambda_i \mathrm{Id}_E)$$

La famille $\tilde{\mathcal{B}} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ est une famille libre de E puisque la somme des E_{λ_i} est dirècte.

$\downarrow$ $\downarrow$

Si $\mathrm{Card}\tilde{\mathcal{B}} = n$, donc si $\tilde{\mathcal{B}}$ est une base de E, alors f est diagonalisable et $M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f)$ est une matrice diagonale dont la diagonale est constituée des λ_i.

Si $Card\tilde{\mathcal{B}} < n$, donc si $\tilde{\mathcal{B}}$ n'est pas une base de E, alors f n'est pas diagonalisable.

Soient E un $\mathbb{R}$-espace vectoriel de dimension 3, $\mathcal{B} = (e_1, e_2, e_3)$ une base de E, et $f \in \mathrm{End}(E)$ l'endomorphisme de E défini par

$$\begin{aligned} f\left(x_1 e_1 + x_2 e_2 + x_3 e_3\right) = {} & (2x_1 - 2x_2 + x_3)\, e_1 \\ & + (2x_1 - 3x_2 + 2x_3)\, e_2 - (x_1 - 2x_2)\, e_3 \end{aligned}$$

pour tous $x_1, x_2, x_3 \in \mathbb{R}$. Sa matrice de représentation dans $\mathcal{B}$ est donnée par

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} 2 & -2 & 1 \\ 2 & -3 & 2 \\ -1 & 2 & 0 \end{pmatrix} .$$

Son polynôme caractéristique est alors donné par

$$
\begin{aligned}
P(X) &= \det \begin{pmatrix} 2-X & -2 & 1 \\ 2 & -3-X & 2 \\ -1 & 2 & -X \end{pmatrix} \\
&= (2-X)\times(3+X)\times X + 2\times 2\times 1 + (-2)\times 2\times(-1) \\
&- 1\times(3+X)\times 1 - 2\times 2\times(2-X) - 2\times(-2)\times(-X) \\
&= -X(X-2)(X+3) - (X+3) \\
&= -(X+3)\times\left(X^2-2X+1\right) \\
&= -(X-1)^2(X+3)\ .
\end{aligned}
$$

L'endomorphisme f a donc deux valeurs propres qui sont -3 et 1. Soient E_{-3} et E_1 les espaces propres de f associés à ces valeurs propres. On a

$$
E_{-3} = \left\{ x = x_1 e_1 + x_2 e_2 + x_3 e_3 \ / \ f(x) = -3x \right\}
$$

et on écrit que $f(x) = -3x$ si et seulement si

$$\begin{pmatrix} 2 & -2 & 1 \\ 2 & -3 & 2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -3 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

On trouve pour système d'équations

$$\begin{cases} 2x_1 - 2x_2 + x_3 = -3x_1 \\ 2x_1 - 3x_2 + 2x_3 = -3x_2 \\ -x_1 + 2x_2 = -3x_3 \end{cases}$$

On a les équivalences suivantes

$$\begin{cases} 2x_1 - 2x_2 + x_3 = -3x_1 \\ 2x_1 - 3x_2 + 2x_3 = -3x_2 \\ -x_1 + 2x_2 = -3x_3 \end{cases} \Leftrightarrow \begin{cases} 5x_1 - 2x_2 + x_3 = 0 \\ 2x_1 + 2x_3 = 0 \\ -x_1 + 2x_2 + 3x_3 = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 4x_1 - 2x_2 = 0 \\ x_3 = -x_1 \\ -4x_1 + 2x_2 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_2 = 2x_1 \\ x_3 = -x_1 \end{cases} .$$

Donc

$$\begin{aligned} E_{-3} &= \left\{ x = x_1 e_1 + x_2 e_2 + x_3 e_3 \ / \ x_2 = 2x_1 \text{ et } x_3 = -x_1 \right\} \\ &= \left\{ x_1 e_1 + 2x_1 e_2 - x_1 e_3 \ / \ x_1 \in \mathbb{R} \right\} \\ &= \left\{ x_1 \left(e_1 + 2e_2 - e_3 \right) \ / \ x_1 \in \mathbb{R} \right\} \end{aligned}$$

et si on pose

$$\tilde{e}_1 = e_1 + 2e_2 - e_3$$

alors E_{-3} est la droite vectorielle de base $(\tilde{e}_1)$. On a de même que

$$E_1 = \left\{ x = x_1 e_1 + x_2 e_2 + x_3 e_3 \ / \ f(x) = x \right\}$$

et on écrit que $f(x) = x$ si et seulement si

$$\begin{pmatrix} 2 & -2 & 1 \\ 2 & -3 & 2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

On trouve pour système d'équations

$$\begin{cases} 2x_1 - 2x_2 + x_3 = x_1 \\ 2x_1 - 3x_2 + 2x_3 = x_2 \\ -x_1 + 2x_2 = x_3 \end{cases}$$

On a les équivalences suivantes

$$\begin{cases} 2x_1 - 2x_2 + x_3 = x_1 \\ 2x_1 - 3x_2 + 2x_3 = x_2 \\ -x_1 + 2x_2 = x_3 \end{cases} \Leftrightarrow \begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_1 - 4x_2 + 2x_3 = 0 \\ -x_1 + 2x_2 - x_3 = 0 \end{cases}$$

$$\Leftrightarrow \quad x_1 - 2x_2 + x_3 = 0 .$$

Donc

$$\begin{aligned} E_1 &= \left\{ x = x_1 e_1 + x_2 e_2 + x_3 e_3 \ / \ x_1 - 2x_2 + x_3 = 0 \right\} \\ &= \left\{ x_1 e_1 + x_2 e_2 - x_1 e_3 + 2x_2 e_3 \ / \ x_1, x_2 \in \mathbb{R} \right\} \\ &= \left\{ x_1 \left(e_1 - e_3 \right) + x_2 \left(e_2 + 2e_3 \right) \ / \ x_1, x_2 \in \mathbb{R} \right\} \end{aligned}$$

et si on pose

$$\tilde{e}_2 = e_1 - e_3 \quad \text{et} \quad \tilde{e}_3 = e_2 + 2e_3$$

alors $(\tilde{e}_2, \tilde{e}_3)$ est une famille génératrice de E_1. Il est facile de vérifier que cette famille est aussi libre car

$$\begin{aligned} \lambda \tilde{e}_2 + \mu \tilde{e}_3 = 0 &\Rightarrow \ \lambda e_1 + \mu e_2 + (2\mu - \lambda) e_3 = 0 \\ &\Rightarrow \ \lambda = \mu = 0 \end{aligned}$$

puisque (e_1, e_2, e_3) est une base. Donc $(\tilde{e}_2, \tilde{e}_3)$ est une base de E_1.

On en déduit que f a deux valeurs propres -3 et 1, que $\dim E_{-3} = 1$ et que $\dim E_1 = 2$. Comme $1 + 2 = 3$, on a que $E = E_{-3} + E_1$ et f est diagonalisable.

Si on note $\tilde{\mathcal{B}} = (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3)$, alors $\tilde{\mathcal{B}}$ est une base de E et on a que

$$M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f) = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

De plus

$$M_{\mathcal{B}\to\tilde{\mathcal{B}}} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ -1 & -1 & 2 \end{pmatrix} .$$

Pour inverser $M_{\mathcal{B}\to\tilde{\mathcal{B}}}$ on peut remarquer que

$$\begin{cases} x+y=X \\ 2x+z=Y \\ -x-y+2z=Z \end{cases} \Leftrightarrow \begin{cases} x+y=X \\ 2x+z=Y \\ 2z=X+Z \end{cases}$$

$$\Leftrightarrow \begin{cases} z=\frac{1}{2}X+\frac{1}{2}Z \\ x=-\frac{1}{4}X+\frac{1}{2}Y-\frac{1}{4}Z \\ y=\frac{5}{4}X-\frac{1}{2}Y+\frac{1}{4}Z \end{cases}$$

Donc :

$$M_{\mathcal{B}\to\tilde{\mathcal{B}}}^{-1} = \begin{pmatrix} -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ \frac{5}{4} & -\frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} ,$$

et on a que

$$M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f) = M_{\mathcal{B}\to\tilde{\mathcal{B}}}^{-1} M_{\mathcal{B}\mathcal{B}}(f) M_{\mathcal{B}\to\tilde{\mathcal{B}}} ,$$

soit encore

$$M_{\mathcal{B}\mathcal{B}}(f) = M_{\mathcal{B}\to\tilde{\mathcal{B}}} M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f) M_{\mathcal{B}\to\tilde{\mathcal{B}}}^{-1} .$$

Remarque : Il s'ensuit que pour n'importe quel entier k,

$$M_{\mathcal{B}\mathcal{B}}(f)^k = M_{\mathcal{B}\to\tilde{\mathcal{B}}} M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f)^k M_{\mathcal{B}\to\tilde{\mathcal{B}}}^{-1} , \qquad (1)$$

et comme

$$M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f)^k = \begin{pmatrix} (-3)^k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

on va pouvoir facilement calculer $M_{\mathcal{B}\mathcal{B}}(f)^3$, $M_{\mathcal{B}\mathcal{B}}(f)^4$, $M_{\mathcal{B}\mathcal{B}}(f)^5$ etc. à partir de la formule (1).

5. Le théorème de Cayley-Hamilton (donné sans preuve)

Théorème (Théorème de Cayley-Hamilton)

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, et soit $f \in End(E)$ un endomorphisme de E. Soit P le polynôme caractéristique de f. Pour toute base $\mathcal{B}$ de E,

$$P\big(M_{\mathcal{B}\mathcal{B}}(f)\big) = 0 \ .$$

Autrement dit, le polynôme caractéristique annule les matrices de représentations $M_{\mathcal{B}\mathcal{B}}(f)$ de f.

Le terme a_0 de P est ici à comprendre comme $a_0 \mathrm{Id}_n$. Si par exemple $n = 3$, si $P(X) = -X^3 + 2X^2 + X - 4$, et si $A = M_{\mathcal{B}\mathcal{B}}(f)$, alors ce que dit Cayley-Hamilton dans ce cas particulier est que

$$-A^3 + 2A^2 + A - 4Id_3 = 0$$

où 0 est la matrice nulle 3×3.

Définition

Soit A une matrice $n \times n$, $n \in \mathbb{N}$. On dit que A est une matrice diagonalisable s'il existe M une matrice inversible $n \times n$ avec la propriété que $M^{-1}AM$ est une matrice diagonale.

On se donne E un $\mathbb{R}$-espace vectoriel de dimension n, et $\mathcal{B}$ une base de E. Par exemple $E = \mathbb{R}^n$ et $\mathcal{B}$ la base canonique de $\mathbb{R}^n$ constituée des vecteurs $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$,..., et $(0, \ldots, 0, 1)$. Il existe (on l'a déjà vu) un unique endomorphisme f de E qui est tel que

$$M_{\mathcal{B}\mathcal{B}}(f) = A \ .$$

Si $(e_1, \ldots, e_n)$ sont les vecteurs de $\mathcal{B}$, f est caractérisé par le fait que les coordonnées de $f(e_i)$ dans la base $\mathcal{B}$ sont les composantes de la $i$ème colonne de A. On suppose que A est diagonalisable. On note $\tilde{\mathcal{B}}$ la famille de vecteurs de E qui est telle que

$$M_{\mathcal{B}\to\tilde{\mathcal{B}}} = M \ .$$

Si φ est l'endomorphisme de E défini par le fait que

$$M_{\mathcal{B}\mathcal{B}}(\varphi) = M \ ,$$

les vecteurs $\tilde{e}_i$ de $\tilde{\mathcal{B}}$ sont données par les relations $\varphi(e_i) = \tilde{e}_i$, et φ est un isomorphisme puisque M est inversible, de sorte que $\tilde{\mathcal{B}}$ est une base. On a alors que

$$\begin{aligned} M^{-1}AM &= M^{-1}_{\mathcal{B}\to\tilde{\mathcal{B}}} M_{\mathcal{B}\mathcal{B}}(f) M_{\mathcal{B}\to\tilde{\mathcal{B}}} \\ &= M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f) \ . \end{aligned}$$

Par suite, dire que A est diagonalisable entraîne que f est diagonalisable. La réciproque est vraie, et A est diagonalisable si et seulement si f l'est. On a $M^{-1}AM = D$ avec $A = M_{\mathcal{B}\mathcal{B}}(f)$ qui équivaut alors à $M = M_{\mathcal{B}\to\tilde{\mathcal{B}}}$ et $D = M_{\tilde{\mathcal{B}}\tilde{\mathcal{B}}}(f)$. Dans la pratique, nul n'est besoin de déterminer f. On calcule le polynôme caractéristique $P(X) = det(A - X\mathrm{Id}_n)$, où Id_n est la matrice identité $n \times n$. On calcule les racines réelles de P, et on récupère ce qui a été dit dans la section précédente.

Exercice : Soit $\alpha, a, b, c \in \mathbb{R}$ des nombres réels. On considère la matrice

$$A = \begin{pmatrix} \alpha & a & b \\ 0 & \alpha & c \\ 0 & 0 & \alpha \end{pmatrix} .$$

Montrer que A est diagonalisable si et seulement si $a = b = c = 0$.

Solution : Si $a = b = c = 0$ alors $A = \alpha \mathrm{Id}_3$ et A est clairement diagonalisable (puisque diagonale). A l'inverse, supposons que A est diagonalisable. Son polynôme caractéristique est

$$P(X) = \det \begin{pmatrix} \alpha - X & a & b \\ 0 & \alpha - X & c \\ 0 & 0 & \alpha - X \end{pmatrix} = -(X - \alpha)^3$$

(le déterminant d'une matrice triangulaire est égal au produit de ses termes diagonaux). Donc A a une seule valeur propre qui est α. On a supposé que A était diagonalisable et donc il existe P une matrice inversible 3×3 telle que

$$P^{-1}AP = \alpha \mathrm{Id}_3 .$$

Soit encore, en multipliant à -gauche par P et à droite par P^{-1},

$$A = \alpha P \mathrm{Id}_3 P^{-1} = \alpha P P^{-1} = \alpha \mathrm{Id}_3$$

ce qui n'est possible que si $a = b = c = 0$. $\square$

Fin du chapitre 3