Chapitre 1. Dualité en dimension finie

Dans ce chapitre, nous allons examiner les liens qui existent entre un espace vectoriel E de dimension finie sur un corps commutatif quelconque K, et **son espace vectoriel dual** E^*, qui est par définition l'espace vectoriel de toutes les formes linéaires sur E. Autrement dit, $E^* = \mathcal{L}(E, K)$.

Nous savons déjà que
$$\dim E^* = \dim E.$$
En effet, ceci découle du fait que pour deux espaces vectoriels E, F de dimension finie, la dimension de l'espace des applications linéaires de E vers F est donnée par la formule $\dim \mathcal{L}(E, F) = \dim E \cdot \dim F$. En plus, le corps K, vu comme espace vectoriel sur K, est de dimension 1.

Nous verrons dans ce chapitre qu'une base B de E détermine de manière naturelle une base de E^*. Cette base, notée B^*, sera appelée **la base duale de** B.

Commençons par la définition des formes linéaires coordonnées associées à une base B.

Définition.
Soit E un espace vectoriel de dimension finie n sur un corps commutatif K.
Soit $B = (e_1, \ldots, e_n)$ une base quelconque de E.
Soit $i \in [[1, n]]$. **La i-ème forme linéaire coordonnée** (*associée à B*) *est la forme linéaire*
$$\varphi_i : \left\{ \begin{array}{ccc} E & \to & K \\ x = \displaystyle\sum_{k=1}^{n} \lambda_k e_k & \mapsto & \lambda_i \end{array} \right.$$

Quelques commentaires sont nécessaires.
La définition de φ_i a un sens, puisque la décomposition $x = \displaystyle\sum_{k=1}^{n} \lambda_k e_k$ est unique. En effet B est une base. Ensuite, il s'agit bien d'une forme linéaire ; c'est une conséquence des axiomes de définition des espaces vectoriels.

Il faut remarquer que les n formes linéaires coordonnées dépendent de la base B.

Finalement, pour tout $i, j \in [[1, n]]$, on a
$$\varphi_i(e_j) = \delta_{ij} = \left\{ \begin{array}{ll} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{array} \right.$$

Théorème et définition.
Soit E un espace vectoriel de dimension finie n sur un corps commutatif K.
Soit $B = (e_1, \ldots, e_n)$ une base quelconque de E.
Alors la famille $(\varphi_1, \ldots, \varphi_n)$ des formes linéaires coordonnées est une base de E^*.

On appelle $(\varphi_1, \ldots, \varphi_n)$ **la base duale de** *B. On la note B^*.*

Preuve.
Comme nous savons déjà que $\dim E^* = n$, il est suffisant de montrer que $(\varphi_1, \ldots, \varphi_n)$ est une famille libre de E^*.
Supposons que pour un certain n-uplet de scalaires $(\alpha_1, \ldots, \alpha_n)$ on a
$$\sum_{k=1}^{n} \alpha_k \varphi_k = 0.$$
Fixons $i \in [[1, n]]$. Alors on a évidemment
$$\left(\sum_{k=1}^{n} \alpha_k \varphi_k \right) (e_i) = 0.$$

Le membre de gauche se simplifie en

$$\alpha_i = 0.$$

Puisque cela est vrai pour tout i, nous avons montré que $(\varphi_1, \ldots, \varphi_n)$ est bien une famille libre. Donc c'est une base. □

Avec les notations ci-dessus, on peut écrire

$$\forall x \in E, \quad x = \sum_{i=1}^{n} \varphi_i(x) e_i.$$

Corollaire.
Pour tout vecteur e non nul de E, il existe une forme linéaire φ sur E avec $\varphi(e) = 1$.

Preuve.
Nous savons qu'on peut compléter e en une base $(e, e_2, \ldots, e_n)$ de E. Alors la première forme linéaire coordonnée φ_1 associée à cette base obéit à la relation

$$\varphi_1(e) = 1.$$

□

Corollaire.
Soit v un vecteur de E tel que $\forall \varphi \in E^*$, $\varphi(v) = 0$. Alors $v = 0$.

Preuve.
Évidente si on considère la contraposée. On est ramené au corollaire précédent. □

Pour l'instant, nous avons associé à toute base B de E une base B^* de E^*. On peut aussi se demander si on peut inverser la construction. Etant donnée une base L de E^*, est-ce qu'on peut trouver une base B de E dont la base duale est L, c'est-à-dire $L = B^*$? Y a-t-il unicité de cette base B ?

La réponse à ces questions est positive.

Théorème de la base anté-duale.
Soit E un espace vectoriel de dimension n sur K, et E^* son espace vectoriel dual.
On considère une base $L = (\omega_1, \ldots, \omega_n)$ de E^*.
Alors il existe une base unique B de E telle que $B^* = L$.
Cette unique base B est appelée **la base anté-duale de L.**

Preuve.
A l'aide de la base L, on peut définir une application linéaire

$$\omega : \left\{ \begin{array}{ccc} E & \to & K^n \\ x & \mapsto & (\omega_1(x), \ldots, \omega_n(x)) \end{array} \right.$$

La linéarité de ω est une conséquence directe de la linéarité de $\omega_1, \ldots, \omega_n$.
Nous allons montrer que ω est un isomorphisme linéaire. Comme E et K^n ont la même dimension, il suffit de montrer que le noyau de ω est réduit au vecteur nul.
Si x est dans le noyau de ω, alors

$$\forall i \in [[1, n]], \omega_i(x) = 0.$$

Comme, par ailleurs, la famille des ω_i est une famille génératrice de E^*, on en déduit que

$$\forall \phi \in E^*, \phi(x) = 0.$$

Par le corollaire précédent, on peut dire que $x = 0$. Bref, $\ker \omega = \{0\}$.

L'application ω est donc un isomorphisme linéaire.
Nous cherchons des vecteurs $(e_1, \ldots, e_n)$ tels que

$$\forall i, j \in [[1, n]], \omega_i(e_j) = \left\{ \begin{array}{ll} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{array} \right.$$

C'est équivalent à

$$\forall j \in [[1, n]], \omega(e_j) = (0, \ldots, 0, 1, 0, \ldots, 0),$$

avec 1 en position j.
Puisque ω est une bijection, cela s'écrit alors

$$\forall j \in [[1, n]], e_j = \omega^{-1}(0, \ldots, 0, 1, 0, \ldots, 0).$$

Vu de cette manière, l'existence et l'unicité de e_j est claire. On a donc trouvé une famille de vecteurs $(e_1, \ldots, e_n)$ de E qui convient. Cette famille est une base, car c'est l'image de la base canonique de K^n par l'isomorphisme ω^{-1}.

$\square$

Théorème.
Soit F un sous-espace vectoriel de E, avec $\dim F = p$ et $\dim E = n$.
L'ensemble $\{\varphi \in E^* | \forall x \in F, \varphi(x) = 0\}$ des formes linéaires de E s'annulant sur F est un sous-espace vectoriel de E^* de dimension $n - p$.

Preuve.
On choisit une base $(e_1, \ldots, e_p)$ de F, et on considère l'application linéaire

$$\Phi : \left\{ \begin{array}{ccc} E^* & \to & K^p \\ \varphi & \mapsto & (\varphi(e_1), \ldots, \varphi(e_p)) \end{array} \right.$$

D'abord, prouvons que $\{\varphi \in E^* | \forall x \in F, \varphi(x) = 0\}$ est le noyau de Φ.
Il est évident que $\{\varphi \in E^* | \forall x \in F, \varphi(x) = 0\} \subset \ker \Phi$.
Réciproquement, si $\Phi(\varphi) = 0$, alors φ s'annule sur tous les vecteurs de la base $(e_1, \ldots, e_p)$ de F. Par linéarité de φ, on en déduit de suite que φ s'annule sur tout vecteur de F. Donc $\ker \Phi \subset \{\varphi \in E^* | \forall x \in F, \varphi(x) = 0\}$.

Nous avons prouvé que $\{\varphi \in E^* | \forall x \in F, \varphi(x) = 0\} = \ker \Phi$.
C'est donc un sous-espace vectoriel de E^*. Nous allons calculer sa dimension à l'aide du théorème du rang : $\dim \ker \Phi = \dim E^* - \text{rang}(\Phi) = n - \text{rang}(\Phi)$.
Il suffit de montrer que le rang de Φ vaut p, autrement dit que Φ est surjective.

Complétons la base $(e_1, \ldots, e_p)$ de F en une base $B = (e_1, \ldots, e_n)$ de E, et appelons $(\varphi_1, \ldots, \varphi_n)$ la base duale de B. Prenons un élément quelconque $(\lambda_1, \ldots, \lambda_p) \in K^p$. On a alors

$$\begin{array}{rcl} \Phi(\sum_{j=1}^{p} \lambda_j \varphi_j) & = & (\sum_{j=1}^{p} \lambda_j \varphi_j(e_1), \ldots, \sum_{j=1}^{p} \lambda_j \varphi_j(e_p)) \\ & = & (\lambda_1, \ldots, \lambda_p). \end{array}$$

Cela montre que $(\lambda_1, \ldots, \lambda_p) \in K^p$ possède un antécédent par Φ. La surjectivité de Φ est établie et la preuve est achevée.

$\square$

Examinons un cas particulier de ce théorème, à savoir $p = n - 1$. Cela signifie que F est un hyperplan de E. Ce théorème dit alors que le sous-espace vectoriel des formes linéaires de E qui s'annulent sur l'hyperplan est de dimension 1. Autrement dit, deux formes s'annulant sur un même hyperplan sont nécessairement colinéaires.

Notre prochain résultat est une sorte de symétrique du théorème précédent.

Théorème.
Soit E un espace vectoriel de dimension n sur K.
Soit $(\varphi_1, \dots, \varphi_q)$ une famille **libre** *de q formes linéaires sur E.*
Alors $\displaystyle\bigcap_{i=1}^{q} \ker \varphi_i$ est un sous-espace vectoriel de E de dimension $n-q$. En outre, toute forme linéaire qui s'annule sur $\displaystyle\bigcap_{i=1}^{q} \ker \varphi_i$ est une combinaison linéaire de $\varphi_1, \dots, \varphi_q$.

Preuve.
Définissons une application linéaire

$$\Psi : \left\{ \begin{array}{ccl} E & \to & K^q \\ x & \mapsto & (\varphi_1(x), \dots, \varphi_q(x)). \end{array} \right.$$

L'intersection des noyaux des φ_i est évidemment le noyau de Ψ. Donc il s'agit d'un sous-espace vectoriel de E dont la dimension (par le théorème du rang) vaut $n - \mathrm{rang}(\Psi)$. Il nous faut montrer que le rang de Ψ est égal à q, ou encore que Ψ est surjective.
Puisque la famille $(\varphi_1, \dots, \varphi_q)$ de formes linéaires est libre, on peut (par le théorème de la base incomplète) la compléter en une base $(\varphi_1, \dots, \varphi_n)$ de l'espace dual E^*. Cette base de E^* possède alors une base anté-duale

$$L = (e_1, \dots, e_n).$$

Si on se donne un élément quelconque $(\lambda_1, \dots, \lambda_q) \in K^q$, on observe alors ceci :

$$\begin{aligned} \Psi(\sum_{j=1}^{q} \lambda_j e_j) &= (\varphi_1(\sum_{j=1}^{q} \lambda_j e_j), \dots, \varphi_q(\sum_{j=1}^{q} \lambda_j e_j)) \\ &= (\lambda_1, \dots, \lambda_q). \end{aligned}$$

C'est la preuve de la surjectivité de Ψ. Nous avons montré la première partie du théorème.

Il reste à montrer que toute forme linéaire s'annulant sur $F = \displaystyle\bigcap_{i=1}^{q} \ker \varphi_i$ est forcément une combinaison linéaire des $\varphi_1, \dots, \varphi_q$. Nous savons par le théorème précédent que l'ensemble G des formes linéaires qui s'annulent sur F est un sous-espace vectoriel de E^* de dimension

$$\dim G = n - \dim F = n - (n - q) = q.$$

Or les q formes linéaires $\varphi_1, \dots, \varphi_q$ appartiennent manifestement toutes à G, puisque $F = \displaystyle\bigcap_{i=1}^{q} \ker \varphi_i$ est contenu dans tous les noyaux. Donc $\mathrm{Vect}(\varphi_1, \dots, \varphi_q) \subset G$. Or G est de dimension q, et comme la famille $(\varphi_1, \dots, \varphi_q)$ est libre, on a aussi que $\mathrm{Vect}(\varphi_1, \dots, \varphi_q)$ est de dimension q. Par conséquent $G = \mathrm{Vect}(\varphi_1, \dots, \varphi_q)$. Cela veut dire que toute forme linéaire s'annulant sur F est dans $\mathrm{Vect}(\varphi_1, \dots, \varphi_q)$, et s'écrit donc comme combinaison linéaire de ces q formes. $\square$

On utilise assez souvent ce théorème dans le cas $q = 1$: le noyau d'une forme linéaire φ non nulle sur E est un hyperplan de E, et toute forme linéaire ω s'annulant sur cet hyperplan est colinéaire à φ. Cela signifie qu'il existe un scalaire $\lambda \in K$ avec $\omega = \lambda \cdot \varphi$.

Si E est un espace vectoriel sur K de dimension n, et $B = (e_1, \ldots, e_n)$ une base de E, on note souvent $(e_1^*, \ldots, e_n^*)$ les vecteurs de sa base duale B^*. Cette notation est un peu dangereuse, parce qu'on peut penser que e_1^* dépend seulement de e_1. Cela n'est pas le cas. La forme linéaire e_1^* est correctement définie seulement si on se donne tous les vecteurs de la base B.

Nous avons vu lors de notre étude des espaces euclidiens (où le corps K est par définition le corps des réels $\mathbb{R}$) qu'un produit scalaire $\langle\ ,\ \rangle$ sur E sert à définir un isomorphisme $\alpha : E \to E^*$. On rappelle sa construction :

$$\forall x, y \in E, (\alpha(x))(y) = \langle x, y \rangle.$$

Si $B = (e_1, \ldots, e_n)$ est une base de E, alors $(\alpha(e_1), \ldots, \alpha(e_n))$ est une base de E^*, car α est un isomorphisme. Mais cette construction ne doit pas être confondue avec celle de la base duale. En général $(\alpha(e_1), \ldots, \alpha(e_n)) \neq B^*$. Il est néanmoins intéressant d'étudier dans quel cas il y a égalité entre ces deux bases de E^*. L'égalité signifie que $\forall i \in [[1, n]], \alpha(e_i) = \varphi_i$. Deux formes linéaires sur E sont égales si et seulement si elles prennent des valeurs égales sur tous les vecteurs d'une base de E. Autrement dit

$$(\alpha(e_1), \ldots, \alpha(e_n)) = B^* \iff \forall i \in [[1, n]], \forall j \in [[1, n]], (\alpha(e_i))(e_j) = \varphi_i(e_j).$$

La dernière égalité s'écrit également $\langle e_i, e_j \rangle = \delta_{ij}$, ce qui signifie précisément que B est une base **orthonormée** pour le produit scalaire $\langle\ ,\ \rangle$. Bref, l'égalité $(\alpha(e_1), \ldots, \alpha(e_n)) = B^*$ a lieu si et seulement si B est une base orthonormée pour $\langle\ ,\ \rangle$.

Chapitre 2. Sous-espaces vectoriels stables par un endomorphisme

Définition.
Soit E un espace vectoriel sur un corps commutatif K, et u un endomorphisme de E.
On dit qu'un sous-espace vectoriel F de E est **stable par** *u si $u(F) \subset F$, c'est-à-dire*

$$\forall x \in F, u(x) \in F.$$

Lorsqu'un sous-espace vectoriel F est stable par u, on peut définir **l'endomorphisme de F induit par** u. On le note u_F. C'est l'application linéaire suivante :

$$u_F : F \to F : x \mapsto u(x).$$

La stabilité de F par u permet de remplacer l'espace d'arrivée E par son sous-espace vectoriel F. L'endomorphisme de F induit par u ne doit pas être confondu avec la restriction $u|_F$ de u à F, car cette restriction est une application linéaire de F vers E.

À l'aide des vecteurs propres d'un endomorphisme, il est facile de construire certains sous-espaces stables.

Théorème.
Soit u un endomorphisme de E, et soient $v_1, \ldots, v_k$ des vecteurs propres de u (pas forcément de même valeur propre). Alors $F = Vect(v_1, \ldots, v_k)$ est stable par u.

Preuve.
Soit x un élément quelconque de F. Alors on peut écrire $x = \sum_{j=1}^{k} \alpha_j v_j$. Appelons λ_j la valeur propre associée à v_j. On a

$$u(x) = u(\sum_{j=1}^{k} \alpha_j v_j) = \sum_{j=1}^{k} \alpha_j u(v_j) = \sum_{j=1}^{k} (\alpha_j \lambda_j) v_j \in F.$$

□

Voici un autre moyen simple d'obtenir des sous-espaces stables :

Théorème.
Soient u, v deux endomorphismes de E **qui commutent :** $u \circ v = v \circ u$.
Alors l'image et le noyau de u sont stables par v, et l'image et le noyau de v sont stables par u.

Preuve.
Il suffit par symétrie de montrer que l'image et le noyau de u sont stables par v.
Montrons que l'image de u est stable par v :
Soit $x \in \text{Im } u$. Alors il existe $y \in E$ avec $x = u(y)$.
Dès lors

$$v(x) = v(u(y)) = u(v(y)) \in \text{Im } u.$$

La stabilité de Im u par v est prouvée.

Montrons que le noyau de u est stable par v :
Soit $x \in \ker u$. Alors

$$u(v(x)) = v(u(x)) = v(0) = 0.$$

Donc $v(x) \in \ker u$, et la stabilité de $\ker u$ par v est prouvée. □

En particulier, si u et v commutent, alors tout sous-espace propre de u est stable par v. En effet, comme u, v commutent, on a aussi, pour tout $\lambda \in K$, la relation

$$(u - \lambda \mathrm{id}) \circ v = u \circ v - \lambda v = v \circ u - \lambda v = v \circ (u - \lambda \mathrm{id}).$$

Donc le noyau de $u - \lambda\mathrm{id}$ (à savoir l'espace propre de u correspondant à la valeur propre λ) est stable par v, grâce au théorème que nous venons de prouver.

Si un sous-espace vectoriel est stable par u, on peut le voir sur certaines représentations matricielles de u :

Théorème.
Soit F un sous-espace vectoriel d'un espace vectoriel E de dimension finie n. Soit u un endomorphisme de E, et $e = (e_1, \ldots, e_n)$ une base de E telle que $F = \mathrm{Vect}(e_1, \ldots, e_p)$ pour un certain $p \in [[1, n]]$.
Le sous-espace vectoriel F est stable par u si et seulement si la matrice $A = \mathcal{M}_e(u)$ de u par rapport à la base e possède la propriété :

$$\forall i \in [[p + 1, n]], \forall j \in [[1, p]], A_{ij} = 0.$$

Preuve.
F est stable par u
$\iff \forall j \in [[1, p]], u(e_j) \in F$
$\iff \forall j \in [[1, p]], u(e_j) \in \mathrm{Vect}(e_1, \ldots, e_p)$
$\iff \forall j \in [[1, p]], \sum_{i=1}^{n} A_{ij} e_i \in \mathrm{Vect}(e_1, \ldots, e_p)$
$\iff \forall j \in [[1, p]], \forall i \in [[p + 1, n]], A_{ij} = 0.$ □

En d'autres termes, la matrice A est de la forme

$$\begin{pmatrix} * & \cdots & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & * & \cdots & * \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix}$$

avec un bloc de zéros de $n - p$ lignes et p colonnes en bas à gauche.

Notre prochain théorème nous indique comment calculer le déterminant d'une telle matrice.

Théorème.
Soit A une matrice carrée $n \times n$ possédant un bloc de zéros en bas à gauche. On l'écrit sous la forme

$$A = \begin{pmatrix} C & D \\ 0 & E \end{pmatrix}$$

où C est une matrice $p \times p$, D une matrice $p \times (n - p)$, 0 la matrice nulle $(n - p) \times p$, et E une matrice $(n - p) \times (n - p)$.
Le déterminant de la matrice A est égal au produit des déterminants de C et E :

$$\det A = (\det C) \cdot (\det E).$$

Preuve.
Utilisons la formule exprimant le déterminant de A comme somme de $n!$ termes :

$$\det A = \sum_{\sigma \in \mathrm{Sym}(n)} \epsilon(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n}.$$

Or, si $j \leqslant p$ et $i \geqslant p+1$, on a $A_{ij} = 0$. On peut donc enlever de la somme toutes les permutations σ pour lesquelles il existe $j \in [[1,p]]$ avec $\sigma(j) \notin [[1,p]]$.
Donc il reste seulement les permutations σ telles que

$$\sigma([[1,p]]) \subset [[1,p]].$$

Comme σ est une permutation de $[[1,n]]$, on a forcément

$$\sigma([[1,p]]) = [[1,p]] \qquad \text{et } \sigma([[p+1,n]]) = [[p+1,n]].$$

Cela signifie que les σ qui restent dans la somme se décomposent en composée (commutative) d'une permutation de $[[1,p]]$ et d'une permutation de $[[p+1,n]]$. Donc

$$\det A = \sum_{\alpha \in \mathrm{Sym}([[1,p]]), \beta \in \mathrm{Sym}([[p+1,n]])} \epsilon(\alpha \circ \beta) A_{\alpha(1)1} \cdots A_{\alpha(p)p} \cdot A_{\beta(p+1),p+1} \cdots A_{\beta(n),n}.$$

Puisque $\epsilon(\alpha \circ \beta) = \epsilon(\alpha)\epsilon(\beta)$, on obtient ensuite

$$\det A = \left(\sum_{\alpha \in \mathrm{Sym}([[1,p]])} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(p)p} \right) \cdot \left(\sum_{\beta \in \mathrm{Sym}([[p+1,n]])} \epsilon(\beta) A_{\beta(p+1),p+1} \cdots A_{\beta(n),n} \right)$$

Le premier facteur est évidemment égal à $\displaystyle\sum_{\alpha \in \mathrm{Sym}(p)} \epsilon(\alpha) C_{\alpha(1)1} \cdots C_{\alpha(p)p} = \det C$.
Pour le second facteur, si on associe à $\beta \in \mathrm{Sym}([[p+1,n]])$ la permutation β' de $[[1,n-p]]$ définie par

$$\beta'(k) = \beta(k+p) - p,$$

on a $\epsilon(\beta) = \epsilon(\beta')$, et le second facteur s'écrit alors

$$\sum_{\beta \in \mathrm{Sym}(n-p)} \epsilon(\beta) E_{\beta'(1),1} \cdots E_{\beta'(n-p),n-p} = \det E.$$

D'où le résultat

$$\det A = (\det C)(\det E).$$

$\square$

Attention : Ne pas remplacer $\det C \cdot \det E$ par $\det(C \cdot E)$, car en général $C \cdot E$ n'existe pas.

Par itération de ce théorème, on peut calculer facilement les déterminants "triangulaires par blocs". Ainsi, par exemple

$$\det \begin{bmatrix} 1 & -1 & 4 & 3 & 7 & -2 \\ 1 & 2 & -2 & -3 & 5 & -6 \\ 0 & 0 & 4 & 2 & -3 & 7 \\ 0 & 0 & -1 & -3 & 5 & 2 \\ 0 & 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 0 & 3 & -2 \end{bmatrix} = \det \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \cdot \det \begin{bmatrix} 4 & 2 \\ -1 & -3 \end{bmatrix} \cdot \det \begin{bmatrix} -2 & 3 \\ 3 & -2 \end{bmatrix}.$$

Comme le déterminant est invariant par transposition, on peut aussi calculer des déterminants "triangulaires inférieurs par blocs" :

$$\det \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} = \det \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \cdot \det \begin{bmatrix} 3 & 4 \\ 7 & 8 \end{bmatrix}.$$

Nous allons voir maintenant comment décrire les endomorphismes qui stabilisent chaque sous-espace vectoriel d'une somme directe.

Théorème.
Soit E un espace vectoriel de dimension finie sur un corps commutatif K.
On suppose que E est une somme directe de sous-espaces vectoriels

$$E = \bigoplus_{i=1}^{k} E_i.$$

Soit e une base de E adaptée à cette décomposition, c'est-à-dire : pour tout $i \in [[1,k]]$, la sous-famille

$$(e_{\dim E_1 + \cdots + \dim E_{i-1} + 1}, \ldots, e_{\dim E_1 + \cdots + \dim E_{i-1} + \dim E_i})$$

est une base du sous-espace vectoriel E_i.
Un endomorphisme u de E possède la propriété que chaque E_i est stable par u si et seulement si la matrice de u par rapport à e est de la forme "diagonale par blocs" :

$$\mathcal{M}_e(u) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_k \end{pmatrix},$$

où, pour tout $i \in [[1,k]]$, A_i est une matrice carrée de taille $(\dim E_i) \times (\dim E_i)$, et les zéros de la matrice représentent des blocs de zéros.

Preuve.
Il suffit de l'écrire. □

Corollaire évident.
Soit $e = (e_1, \ldots, e_n)$ une base d'un espace vectoriel E de dimension n. Un endomorphisme u de E possède une matrice diagonale par rapport à la base e si et seulement si, pour tout $i \in [[1,n]]$, le sous-espace vectoriel de dimension 1 engendré par e_i, à savoir $\mathrm{Vect}(e_i)$, est stable par u.

Le déterminant d'une matrice "diagonale par blocs" est facile à calculer : c'est le produit des déterminants de tous les blocs.

Théorème.

1. Soit A une matrice diagonale par blocs : $A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_k \end{pmatrix}$.

Alors $\det A = \displaystyle\prod_{i=1}^{k} \det A_i$.

2. Soit u un endomorphisme de $E = \displaystyle\bigoplus_{i=1}^{k} E_i$ tel que chaque E_i soit stable par u.
On appelle $u_i : E_i \to E_i$ l'endomorphisme induit par u sur le sous-espace E_i.
Alors $\det u = \displaystyle\prod_{i=1}^{k} \det u_i$.

Preuve.
La seconde partie est seulement la traduction de la première partie au niveau des endomorphismes.
La première partie se prouve en se ramenant à la formule $\det \begin{pmatrix} C & D \\ 0 & E \end{pmatrix} = (\det C)\cdot(\det E)$.
□

Donnons une interprétation des matrices triangulaires supérieures en termes d'endomorphismes et de sous-espaces stables.

Théorème.
Soit $e = (e_1, \ldots, e_n)$ une base d'un espace vectoriel E de dimension finie. Soit u un endomorphisme de E.
La matrice $\mathcal{M}_e(u)$ est triangulaire supérieure si et seulement si les n sous-espaces vectoriels

$$\mathrm{Vect}(e_1), \mathrm{Vect}(e_1, e_2), \ldots, \mathrm{Vect}(e_1, \ldots, e_p), \ldots, \mathrm{Vect}(e_1, \ldots, e_n)$$

sont tous stables par u.

Preuve.
Il suffit d'observer qu'une matrice triangulaire possède en bas à gauche un bloc de zéros de taille $(n-p) \times p$ pour tout $p \in [[1, n]]$, et d'appliquer le théorème qui décrit les matrices des endomorphismes pour lesquels $\mathrm{Vect}(e_1, \ldots, e_p)$ est stable. □

Enfin, nous rappelons que nous avons déjà rencontré la notion de stabilité à travers l'adjoint d'un endomorphisme dans un espace euclidien. Nous avons vu en particulier : Si F est un sous-espace de E (euclidien) stable par u, alors son sous-espace orthogonal $F^{\perp}$ est stable par l'adjoint u^*.

Chapitre 3. Polynômes d'un endomorphisme, d'une matrice. Polynôme minimal.

Définition.
Soit u un endomorphisme d'un espace vectoriel E sur un corps commutatif K.
Soit $P \in K[X]$ un polynôme. On peut écrire

$$P = \sum_{k=0}^{d} a_k X^k,$$

où d est un entier (dépendant de P), et $a_0, \dots, a_d$ des scalaires.
On note $P(u)$ l'endomorphisme de E défini par

$$P(u) := \sum_{k=0}^{d} a_k u^k$$

où $u^k = \underbrace{u \circ u \circ \cdots \circ u}_{k}$ est la composée k-uple de u, et $u^0 = \mathrm{id}_E$.

Exemples.
Si u est un projecteur, et $P = X^2 - X$, alors $P(u) = u \circ u - u = 0$.
Si $u = \lambda \cdot id_E$, où $\lambda \in K$, et P un polynôme quelconque, on trouve $P(u) = P(\lambda) \cdot id_E$.

Théorème.
Soit u un endomorphisme de E. L'application $\Phi : K[X] \to \mathcal{L}(E)$ définie par

$$\Phi(P) = P(u)$$

possède les propriétés suivantes :
1. $\Phi(P+Q) = \Phi(P) + \Phi(Q)$
2. $\Phi(\lambda P) = \lambda \Phi(P)$
3. $\Phi(P \cdot Q) = \Phi(P) \circ \Phi(Q)$
où P, Q sont deux polynômes quelconques, et $\lambda \in K$ un scalaire quelconque.
On peut résumer ces trois propriétés en disant que Φ est à la fois une application linéaire de l'espace vectoriel $K[X]$ vers l'espace vectoriel $\mathcal{L}(E)$ et un morphisme de l'anneau $K[X]$ vers l'anneau $\mathcal{L}(E)$.

Preuve.
Écrivons $P = \sum_{k=0}^{\infty} a_k X^k$ et $Q = \sum_{k=0}^{\infty} b_k X^k$. Ces sommes sont en fait des sommes finies, puisque P, Q sont des polynômes.
1. $\Phi(P+Q) = (P+Q)(u) = \left(\sum_{k=0}^{\infty}(a_k + b_k)X^k\right)(u) = \sum_{k=0}^{\infty}(a_k+b_k)u^k = \sum_{k=0}^{\infty} a_k u^k + \sum_{k=0}^{\infty} b_k u^k$
$= P(u) + Q(u) = \Phi(P) + \Phi(Q)$.
2. $\Phi(\lambda P) = (\lambda P)(u) = \left(\sum_{k=0}^{\infty}(\lambda a_k)X^k\right)(u) = \sum_{k=0}^{\infty} \lambda a_k u^k = \lambda \sum_{k=0}^{\infty} a_k u^k = \lambda P(u) = \lambda \Phi(P)$.
3. Montrons d'abord que la relation est vraie dans le cas particulier où P, Q sont deux monômes : $P = X^p, Q = X^q$. Clairement $\Phi(P \cdot Q) = \Phi(X^{p+q}) = u^{p+q} = u^p \circ u^q = \Phi(P) \circ \Phi(Q)$.

Avec cette règle, et celles prouvées auparavant, on peut faire le cas général : En effet, on peut écrire tout polynôme comme une combinaison linéaire de monômes : $P = \sum_{i=0}^{\infty} a_i X^i$ et $Q = \sum_{j=0}^{\infty} b_j X^j$. Dès lors

$$\Phi(P \cdot Q) = \Phi\left(\sum_{(i,j)\in\mathbb{N}^2} a_i b_j X^i \cdot X^j\right) = \sum_{(i,j)\in\mathbb{N}^2} a_i b_j \Phi(X^i \cdot X^j) = \sum_{(i,j)\in\mathbb{N}^2} a_i b_j \Phi(X^i)\circ\Phi(X^j) = \left(\sum_{i\in\mathbb{N}} a_i \Phi(X^i)\right) \circ \left(\sum_{j\in\mathbb{N}} b_j \Phi(X^j)\right) = \Phi\left(\sum_{i\in\mathbb{N}} a_i X^i\right) \circ \Phi\left(\sum_{j\in\mathbb{N}} b_j X^j\right) = \Phi(P)\circ\Phi(Q).$$

□

Théorème.
Soit u un endomorphisme d'un espace vectoriel E sur K.
1. Pour P, Q deux polynômes quelconques de $K[X]$, on a $P(u)\circ Q(u) = Q(u)\circ P(u)$.
2. Pour tout $P \in K[X]$, les sous-espaces $\operatorname{im} P(u)$ et $\ker P(u)$ sont stables par u.
3. Pour tout $\varphi \in GL(E)$ et tout $P \in K[X]$, on a $P(\varphi^{-1}\circ u\circ\varphi) = \varphi^{-1}\circ P(u)\circ\varphi$.

Preuve.
1. Il suffit d'observer que $P(u)\circ Q(u) = (P\cdot Q)(u) = (Q\cdot P)(u) = Q(u)\circ P(u)$.
2. D'après 1., on a $P(u)\circ u = u\circ P(u)$ (c'est le cas particulier où Q est le polynôme X). Comme $P(u)$ et u commutent, on sait alors que les sous-espaces vectoriels $\operatorname{im} P(u)$ et $\ker P(u)$ sont stables par u.
3. Il suffit de montrer la relation pour les monômes $P = X^k$ (la propriété générale en découlera par linéarité). Or

$$\begin{aligned}
P(\varphi^{-1}\circ u\circ\varphi) &= (\varphi^{-1}\circ u\circ\varphi)^k \\
&= (\varphi^{-1}\circ u\circ\varphi)\circ(\varphi^{-1}\circ u\circ\varphi)\circ\cdots\circ(\varphi^{-1}\circ u\circ\varphi) \\
&= \varphi^{-1}\circ u\circ(\varphi\circ\varphi^{-1})\circ u\circ(\varphi\circ\varphi^{-1})\circ\cdots\circ u\circ\varphi \\
&= \varphi^{-1}\circ(u\circ u\circ\cdots\circ u)\circ\varphi \\
&= \varphi^{-1}\circ u^k\circ\varphi \\
&= \varphi^{-1}\circ P(u)\circ\varphi.
\end{aligned}$$

□

Définition.
Soit u un endomorphisme de E.
On appelle **polynôme annulateur de** *u tout polynôme $P \in K[X]$ tel que $P(u) = 0$.*

Exemples.
Le polynôme nul est un polynôme annulateur pour tout endomorphisme.
Le polynôme constant 1 n'est un polynôme annulateur pour aucun endomorphisme.
Le polynôme $X - 1$ est un polynôme annulateur de $u = id_E$.
Le polynôme $X^2 - 1$ est un polynôme annulateur de toute symétrie s de E.
Le polynôme $X^2 - X$ est un polynôme annulateur de tout projecteur p de E.

L'ensemble des polynômes annulateurs de u possède une structure intéressante.

Théorème.
Soit u un endomorphisme de E.
L'ensemble des polynômes annulateurs de u est un sous-espace vectoriel de $K[X]$.
C'est aussi un idéal de l'anneau $K[X]$.

Preuve.
Reprenons l'application $\Phi : K[X] \to \mathcal{L}(E)$ défini par

$$\Phi(P) = P(u)$$

Nous savons que Φ est une application linéaire. Son noyau est évidemment l'ensemble des polynômes annulateurs de u. Or nous savons que le noyau d'une application linéaire est un sous-espace vectoriel de l'espace de départ.

Il reste à prouver que $\ker \Phi$ est un idéal de l'anneau de $K[X]$, c'est-à-dire :
1. Si P et Q sont deux polynômes annulateurs de u, alors $P+Q$ est aussi annulateur. C'est évident !
2. Si P est annulateur de u, et $Q \in K[X]$ quelconque, alors $P \cdot Q$ est aussi annulateur. C'est facile à comprendre :

$$\begin{aligned}(PQ)(u) &= P(u) \circ Q(u) \\ &= 0 \circ Q(u) \\ &= 0\end{aligned}$$

□

Profitons de cette preuve pour rappeler qu'un **idéal** I de $K[X]$ est un sous-groupe du groupe $(K[X], +)$ qui possède aussi la propriété : $\forall (P, Q) \in I \times K[X], P \cdot Q \in I$. Un idéal ne doit pas être confondu avec un sous-anneau.

Si on connaît un polynôme annulateur non nul de u, on a déjà une information précieuse sur le spectre de u :

Théorème.
Soit u un endomorphisme de E, et P un polynôme annulateur de u.
Alors toute valeur propre de u est une racine de P.
Autrement dit, le spectre de u est contenu dans l'ensemble des racines de P.

Preuve.
Soit λ une valeur propre de u. Choisissons un vecteur propre non nul $\vec{v} \in E$ associé à la valeur propre λ. Comme $P(u) = 0$, on a alors

$$P(u)(\vec{v}) = \vec{0}.$$

Comme $u(\vec{v}) = \lambda \vec{v}$, on voit que

$$P(u)(\vec{v}) = \sum_{k=0}^{\infty} a_k (u^k(\vec{v})) = \sum_{k=0}^{\infty} a_k \lambda^k \vec{v} = P(\lambda) \cdot \vec{v}.$$

Donc $P(\lambda) \cdot \vec{v} = \vec{0}$. Mais $\vec{v}$ est non nul. D'où $P(\lambda) = 0$. □

On peut en déduire par exemple que les seules valeurs propres possibles pour un projecteur p d'un espace vectoriel sont 0 et 1, car ce sont les racines du polynôme $X^2 - X$, qui est toujours annulateur pour p.
Mais il ne faut pas en déduire que le spectre d'un projecteur est forcément **égal** à $\{0, 1\}$, comme le montre le cas $p = \mathrm{id}$, qui est un projecteur, mais dont le spectre est plus petit : $\mathrm{Spec}(\mathrm{id}) = \{1\}$.

Nous sommes maintenant en mesure de définir le polynôme minimal d'un endomorphisme en dimension finie.

Théorème.
Soit E un espace vectoriel de dimension finie, et u un endomorphisme de E.
Alors il existe un unique polynôme $P \in K[X]$ possédant les propriétés suivantes :
1. P est unitaire,
2. P est un polynôme annulateur de u,
3. Si Q est un polynôme annulateur de u, alors P divise Q.

Preuve.
Montrons d'abord l'unicité. Si P_1 et P_2 ont ces propriétés, alors en vertu de la propriété 3, on peut dire que

$$P_1 | P_2 \qquad \text{et} \qquad P_2 | P_1$$

Mais cela signifie que P_2 et P_1 sont proportionnels, et comme ils sont unitaires tous deux, ils sont égaux.

Il reste à montrer l'existence de P. Tout d'abord, observons qu'il existe au moins un polynôme annulateur non nul. Pour cela, regardons l'application linéaire $\Phi : K[X] \to \mathcal{L}(E)$. L'espace de départ est de dimension infinie, mais celui d'arrivée est de dimension finie. Il en découle que Φ n'est pas injective, et donc que son noyau contient des polynômes non nuls. Or le noyau de Φ est exactement l'espace des polynômes annulateurs de u.
On peut maintenant considérer un polynôme annulateur non nul de degré minimal. Évidemment, si on le divise par son coefficient dominant, il reste annulateur. On appelle P ce polynôme unitaire annulateur de degré minimal. Il vérifie 1. et 2. Reste à montrer qu'il vérifie aussi 3.

Soit Q un polynôme annulateur de u. Effectuons la division euclidienne de Q par P. On écrit donc

$$Q = PA + R$$

avec R de degré strictement inférieur au degré de P. Comme $Q(u) = 0$, on trouve

$$(PA + R)(u) = 0$$

ou encore, d'après les propriétés prouvées ci-dessus

$$P(u) \circ A(u) + R(u) = 0$$

et comme $P(u) = 0$, on aboutit à

$$R(u) = 0$$

Donc R est un polynôme annulateur de u. Mais si R était non nul, on aurait alors trouvé un polynôme annulateur non nul de u d'un degré strictement plus petit que le degré de P. C'est en contradiction avec la minimalité du degré de P. Il faut que $R = 0$. Mais alors on a montré que P divise Q. $\square$

Définition.
L'unique polynôme défini ci-dessus s'appelle **le polynôme minimal de u.**

Exemples.
1. Le polynôme minimal de l'endomorphisme nul est le polynôme X.

2. Le polynôme minimal de l'endomorphisme identité est le polynôme $X - 1$.

3. Si p est un projecteur (différent de l'identité et de l'endomorphisme nul), alors le polynôme minimal de p est le polynôme $X^2 - X$. En effet, $X^2 - X$ est annulateur, donc le polynôme minimal est un diviseur unitaire de $X^2 - X$. Comme un polynôme minimal n'est jamais de degré 0, il faut que le polynôme minimal soit $X^2 - X$ ou X ou $X - 1$. Les deux derniers cas sont à exclure, car ils signifieraient que $p = 0$ ou $p = \mathrm{id}$. Donc le polynôme minimal de p est $X^2 - X$.

4. De la même façon, si s est une symétrie (différente de id et de $-$id), alors le polynôme minimal de s est $X^2 - 1$.

Théorème.
Deux endomorphismes semblables ont le même polynôme minimal.

Preuve.
Soient u_1, u_2 deux endomorphismes semblables d'un espace vectoriel de dimension finie E. Il existe par conséquent un automorphisme φ de E avec

$$u_2 = \varphi^{-1} \circ u_1 \circ \varphi.$$

Appelons P_1, P_2 les polynômes minimaux de u_1, u_2. Nous allons montrer que P_1 divise P_2.

En effet, P_2 est par définition un polynôme annulateur de u_2. Donc

$$P_2(\varphi^{-1} \circ u_1 \circ \varphi) = 0$$

ou encore

$$\varphi^{-1} \circ P_2(u_1) \circ \varphi = 0.$$

En composant à gauche par φ, à droite par φ^{-1}, on trouve alors

$$P_2(u_1) = 0.$$

Cela signifie que P_2 est un polynôme annulateur pour u_1. Or le polynôme minimal de u_1 divise tout polynôme annulateur de u_1. Donc P_1 divise P_2.

En échangeant les rôles de u_1, u_2, on prouve de la même façon que P_2 divise P_1. Il vient alors que P_1 et P_2 sont proportionnels. Enfin, comme ils sont tous deux unitaires, ils sont égaux. $\square$

Le théorème ci-dessous est fondamental pour la suite. C'est le théorème de décomposition des noyaux, parfois aussi appelé "lemme des noyaux".

Théorème de décomposition des noyaux
Soit u un endomorphisme d'un espace vectoriel E sur K, et soient P, Q deux polynômes de $K[X]$ **premiers entre eux.** *Alors on a la décomposition en somme directe*

$$\ker(P \cdot Q)(u) = \ker P(u) \oplus \ker Q(u).$$

Preuve.
Tout d'abord, observons que $\ker P(u) \subset \ker(PQ(u))$. En effet, si $P(u)(x) = 0$, alors

$$(PQ(u))(x) = (QP(u))(x) = (Q(u) \circ P(u))(x) = Q(u)(P(u)(x)) = Q(u)(0) = 0.$$

De la même façon on obtient que $\ker Q(u) \subset \ker(PQ(u))$.

Nous allons prouver que $\ker(PQ(u)) = \ker P(u) + \ker Q(u)$ et ensuite que

$$0 = w_1 + w_2 \in \ker P(u) + \ker Q(u) \Rightarrow w_1 = w_2 = 0.$$

Soit x un élément de $\ker(PQ(u))$. On a donc $PQ(u)(x) = 0$.
Puisque P, Q sont des polynômes premiers entre eux, il existe, grâce au théorème de Bezout pour les polynômes, deux polynômes A, B tels que

$$AP + BQ = P \wedge Q = 1$$

et donc

$$(AP + BQ)(u) = \mathrm{id}.$$

Mais alors

$$\begin{aligned} x &= \mathrm{id}(x) \\ &= ((AP + BQ)(u))(x) \\ &= ((BQ + AP)(u))(x) \\ &= (B(u) \circ Q(u))(x) + (A(u) \circ P(u))(x). \end{aligned}$$

Or le premier terme de cette décomposition de x appartient à $\ker P(u)$, puisque

$$\begin{aligned} P(u)\left((B(u) \circ Q(u))(x)\right) &= (P(u) \circ B(u) \circ Q(u))(x) \\ &= ((PBQ)(u))(x) \\ &= ((BPQ)(u))(x) \\ &= (B(u) \circ (PQ)(u))(x) \\ &= B(u)(0) \\ &= 0. \end{aligned}$$

De la même façon, le second terme $(A(u) \circ P(u))(x)$ est dans $\ker Q(u)$.
Nous avons montré que $\ker(PQ(u)) = \ker P(u) + \ker Q(u)$.

Il reste à montrer que si $w_1 \in \ker P(u), w_2 \in \ker Q(u)$ et $w_1 + w_2 = 0$, alors ces deux vecteurs sont forcément nuls.

On reprend la relation $AP + BQ = 1$. On a alors

$$(AP + BQ)(u)(w_1) = w_1$$

et donc

$$(BQ)(u)(w_1) = w_1.$$

Mais $w_1 = -w_2$. Donc

$$-(BQ)(u)(w_2) = w_1$$

et comme $w_2 \in \ker Q(u)$, on conclut que $0 = w_1$. Ensuite $0 = w_2$, et le théorème est démontré. $\square$

Généralisons maintenant le lemme des noyaux à une famille finie de polynômes deux à deux premiers entre eux.

Théorème général de décomposition des noyaux
Soit u un endomorphisme d'un espace vectoriel E sur K, et soient $P_1, \cdots, P_k$ des polynômes de $K[X]$ **deux à deux premiers entre eux.** *Alors on a la décomposition en somme directe*

$$\ker(P_1 P_2 \cdots P_k)(u) = \ker P_1(u) \oplus \ker P_2(u) \oplus \cdots \oplus \ker P_k(u)$$

Preuve.
Il suffit d'appliquer la proposition précédente aux deux polynômes $P = P_1 P_2 \cdots P_{k-1}$ et $Q = P_k$ (dont on montre qu'ils sont bien premiers entre eux) et de travailler par récurrence sur k. □

Le lemme des noyaux a des corollaires très importants.

Théorème.
Soit u un endomorphisme de E, avec E de dimension finie. Si le polynôme minimal de u est scindé et à racines simples, alors :
1. u est diagonalisable,
2. le spectre de u est égal à l'ensemble des racines du polynôme minimal.

Preuve.
Soit P le polynôme minimal de u. Par hypothèse $P = \prod_{j=1}^{k}(X - r_j)$, avec $r_1, \ldots, r_k$ des scalaires deux à deux distincts (car les racines sont simples). On applique le lemme des noyaux aux k polynômes de degré 1 définis par $P_j = X - r_j$. Évidemment, ces polynômes P_j sont deux à deux premiers entre eux. On a bien le droit d'appliquer le lemme des noyaux. Observons que $\ker(P_1 P_2 \cdots P_k)(u) = \ker P(u) = \ker 0 = E$. Donc

$$E = \bigoplus_{j=1}^{k} \ker P_j(u) = \bigoplus_{j=1}^{k} \ker(u - r_j \cdot \mathrm{id}_E).$$

Notons $E_j = \ker(u - r_j \cdot \mathrm{id}_E)$. Aucun E_j ne peut être réduit au vecteur nul. En effet, si on suppose par l'absurde qu'il existe i avec $E_i = \{0\}$, on aurait $E = \bigoplus_{j \neq i} E_j$, et donc

$$E = \ker(P_1 P_2 \cdots P_{i-1} P_{i+1} \cdots P_k)(u).$$

Mais alors le polynôme $P_1 P_2 \cdots P_{i-1} P_{i+1} \cdots P_k$ serait annulateur de u, ce qui contredit le fait que P est le polynôme minimal.
De la propriété

$$E_j = \ker(u - r_j \cdot \mathrm{id}_E) \neq \{0\}$$

on déduit que r_j est une valeur propre de u, et que E_j est l'espace propre associé à r_j. Comme on a

$$E = \bigoplus_{j=1}^{k} E_j$$

on en déduit que u est diagonalisable, et que le spectre de u est égal à $\{r_1, \ldots, r_k\}$. □

Exemple. Si p est un projecteur (autre que l'identité et l'endomorphisme nul), nous avons vu que le polynôme minimal de p est $X^2 - X$. Comme $X^2 - X = X(X - 1)$ est scindé à racines simples, on en déduit que p est toujours diagonalisable, et que son spectre est $\{0, 1\}$. Nous connaissions déjà ce théorème, et nous l'avons simplement redécouvert grâce au théorème précédent.

Nous allons maintenant étudier le polynôme minimal d'un endomorphisme diagonalisable.

Théorème.
Soit u un endomorphisme diagonalisable de E (supposé de dimension finie).
On suppose que le spectre de u est $\{\lambda_1, \ldots, \lambda_k\}$ (où les λ_i deux à deux distincts).
Alors le polynôme minimal de u est

$$(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k) = \prod_{i=1}^{k} (X - \lambda_i).$$

Preuve.
Le polynôme $(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$ est un polynôme annulateur de u, car pour tout vecteur $x \in E$, on a

$$(u - \lambda_1 \cdot \mathrm{id}) \circ (u - \lambda_2 \cdot \mathrm{id}) \circ \cdots \circ (u - \lambda_k \cdot \mathrm{id})(x) = 0$$

(cela se voit en décomposant x en somme de vecteurs propres). Donc le polynôme minimal est un diviseur du polynôme $(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$.
D'un autre côté, nous savons que chaque valeur propre de u est nécessairement racine de tout polynôme annulateur de u. En particulier, chaque λ_i est une racine du polynôme minimal. Donc il faut que le polynôme minimal soit aussi un multiple du polynôme $(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$. Il en découle que le polynôme minimal est proportionnel à $(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$. Enfin, comme le polynôme minimal est unitaire par définition, on voit que ces deux polynômes sont égaux. $\square$

Nous pouvons maintenant donner une nouvelle caractérisation des endomorphismes diagonalisables.

Théorème.
Soit u un endomorphisme d'un espace vectoriel de dimension finie.
u est diagonalisable $\iff$ le polynôme minimal de u est scindé à racines simples.

Preuve.
La partie "seulement si" est exactement le théorème précédent.
La partie "si" a été démontrée plus haut. $\square$

Pour terminer ce chapitre, nous allons définir par analogie les notions de "polynôme d'une matrice carrée" et "polynôme minimal d'une matrice".

Définition.
Soit A une matrice carrée $n \times n$ à coefficients dans un corps commutatif K, et soit P un polynôme à coefficients dans K. Si $P = \sum_{k=0}^{\infty} c_k X^k$, on note $P(A)$ la matrice carrée $n \times n$ ci-dessous :

$$P(A) := \sum_{k=0}^{\infty} c_k A^k.$$

Cette construction jouit de propriétés très similaires à celles des polynômes d'endomorphismes :

$$\forall P, Q \in K[X], (P + Q)(A) = P(A) + Q(A),$$

$$\forall \lambda \in K, \forall P \in K[X], (\lambda P)(A) = \lambda \cdot P(A),$$

$$\forall P, Q \in K[X], (P \cdot Q)(A) = P(A) \cdot Q(A).$$

On a aussi, pour toute matrice inversible $B \in GL(n, K)$, la propriété

$$P(B^{-1}AB) = B^{-1}P(A)B.$$

Observons que pour toute base e d'un espace vectoriel de dimension finie E, tout endomorphisme u de E, et tout polynôme P, on a l'égalité

$$\mathcal{M}_e(P(u)) = P(\mathcal{M}_e(u)).$$

Etant donnée une matrice carrée A, on appellera **polynôme annulateur de** A tout polynôme P tel que $P(A) = 0$.

Evidemment, pour un endomorphisme u, le polynôme P est annulateur de u si et seulement si le polynôme P est annulateur de toute matrice représentant u.
On observera aussi que deux matrices semblables ont forcément les mêmes polynômes annulateurs.

L'ensemble des polynômes annulateurs d'une matrice A est un idéal de l'anneau $K[X]$, et un sous-espace de l'espace vectoriel $K[X]$.

Comme pour les endomorphismes, les valeurs propres d'une matrice A sont nécessairement des racines de tout polynôme annulateur de A.

On appelle **polynôme minimal de** A l'unique polynôme unitaire P annulateur de A qui vérifie :
Si Q est annulateur de A, alors $P|Q$.

On observera que l'ensemble des polynômes annulateurs de A est exactement l'ensemble des multiples du polynôme minimal de A.

Deux matrices semblables ont toujours le même polynôme minimal.

Le lemme des noyaux se traduit comme suit en termes de matrices :

Théorème général de décomposition des noyaux
Soit A une matrice carrée à coefficients dans K, et soient $P_1, \cdots, P_k$ des polynômes de $K[X]$ **deux à deux premiers entre eux.** *Alors on a la décomposition en somme directe*

$$\ker(P_1P_2\cdots P_k)(A) = \ker P_1(A) \oplus \ker P_2(A) \oplus \cdots \oplus \ker P_k(A).$$

Enfin, nous avons aussi la caractérisation des matrices diagonalisables à l'aide du polynôme minimal :
Une matrice A est diagonalisable si et seulement si son polynôme minimal est scindé sur K, et que toutes ses racines sont simples.

Chapitre 4.
Polynôme caractéristique. Théorème de Cayley-Hamilton.

Commençons ce chapitre par quelques rappels.

Si u est un endomorphisme d'un espace vectoriel de dimension finie n sur un corps commutatif K, on lui associe **son polynôme caractéristique**. Il est dorénavant noté C_u, et il est défini par

$$C_u = \det(X \cdot \mathrm{id} - u) \in K[X].$$

Le polynôme caractéristique est un polynôme unitaire de degré n.

De manière équivalente, on peut aussi définir le polynôme caractéristique d'une matrice carrée A de taille $n \times n$, noté C_A, par

$$C_A = \det(X \cdot I_n - A) \in K[X].$$

Nous avons vu en particulier que deux matrices semblables ont toujours le même polynôme caractéristique. Du côté des endomorphismes, l'équivalent de cette phrase est : si u, v sont deux endomorphismes, et qu'il existe un automorphisme φ avec $v = \varphi^{-1} \circ u \circ \varphi$, alors u et v ont le même polynôme caractéristique.

L'intérêt du polynôme caractéristique de u (respectivement de A) est le fait que ses racines sont les valeurs propres de l'endomorphisme (respectivement de la matrice).

Au chapitre précédent, nous avons aussi rencontré le polynôme minimal d'un endomorphisme (respectivement d'une matrice). Si u est un endomorphisme, le polynôme minimal de u est le polynôme unitaire P de degré minimal tel que $P(u) = 0$. C'est le "plus petit" (au sens du degré) polynôme annulateur de u.

Nous savons déjà que toute valeur propre de A est nécessairement racine du polynôme minimal de A. Autrement dit, toute racine du polynôme caractéristique est nécessairement une racine du polynôme minimal.

Grâce au théorème de Cayley-Hamilton, nous pourrons montrer que la réciproque est vraie : toute racine du polynôme minimal est nécessairement une racine du polynôme caractéristique.

L'énoncé du théorème de Cayley-Hamilton est très simple :
Si u est un endomorphisme en dimension finie, alors $C_u(u) = 0$.

Cela signifie que le polynôme d'endomorphisme $C_u(u)$ est l'endomorphisme nul. Une autre manière de dire les choses : Si on écrit le polynôme caractéristique de u et qu'on remplace la variable X par l'endomorphisme u, alors on obtient l'endomorphisme nul. On peut aussi dire : C_u est un polynôme annulateur de u.

La même chose vaudra pour les matrices : *Si A est une matrice carrée, alors $C_A(A) = 0$.*

Tout ce chapitre est consacré à la preuve du théorème de Cayley-Hamilton. Comme cette preuve est un peu longue, nous allons commencer par une série de lemmes de préparation.

Notation.
Soit E un espace vectoriel sur K, x un vecteur de E, et u un endomorphisme de E.
On note $K[u](x)$ l'ensemble

$$K[u](x) := \{P(u)(x) | P \in K[X]\}.$$

Observons que $K[u](x) \subset E$, et que c'est un sous-espace vectoriel de E.

Lemme 1.
Soit E un espace vectoriel de dimension finie sur K (corps commutatif), x un vecteur de E, et u un endomorphisme de E. Alors il existe un unique polynôme P unitaire de degré minimal dans $K[X]$ tel que $P(u)(x) = 0$.

Preuve.
Il existe évidemment des polynômes unitaires P dans $K[X]$ avec $P(u)(x) = 0$. Il suffit par exemple de prendre pour P le polynôme minimal de u. Celui-ci possède la propriété

$$\forall x \in E, P(u)(x) = 0.$$

Il reste à montrer l'unicité : si P_1, P_2 sont deux polynômes unitaires, avec $P_1(u)(x) = P_2(u)(x) = 0$, et tels que P_1, P_2 sont de degré **minimal** parmi tous les polynômes unitaires P vérifiant $P(u)(x) = 0$, alors $P_1 = P_2$. Pour cela il suffit de montrer ceci :
Si P_1 est unitaire de degré minimal, et si $P_1(u)(x) = P(u)(x)$, alors P_1 divise P.

L'argument est classique. Faisons la division euclidienne de P par P_1 :

$$P = P_1 Q + R$$

avec $\deg R < \deg P_1$. Alors

$$P(u) = P_1(u) \circ Q(u) + R(u) = Q(u) \circ P_1(u) + R(u).$$

En évaluant cet endomorphisme sur le vecteur x, on trouve grâce à $P(u)(x) = P_1(u)(x) = 0$:

$$0 = Q(u)(0) + R(u)(x).$$

Donc $R(u)(x) = 0$ avec $\deg R_1 < \deg P_1$. Par minimalité du degré de P_1, il faut alors que R soit le polynôme nul. Donc P_1 est bien un diviseur de P. □

Lemme 2.
L'espace vectoriel $K[u](x)$ est stable par u.

Preuve.
Soit $y \in K[u](x)$. Alors il existe un polynôme $P \in K[X]$ avec $y = P(u)(x)$.
Dès lors

$$u(y) = u(P(u)(x)) = (u \circ P(u))(x) = Q(u)(x).$$

où Q est le polynôme $Q = XP \in K[X]$. Il est maintenant évident que $u(y) \in K[u](x)$. □

Lemme 3.
On note $u' : K[u](x) \to K[u](x)$ l'endomorphisme induit par u sur $K[u](x)$. On note

$$P = X^d + a_{d-1}X^{d-1} + a_{d-2}X^{d-2} + \cdots + a_1 X + a_0 \in K[X]$$

l'unique polynôme unitaire du lemme 1. Alors il existe une base e de $K[u](x)$ telle que

$$\mathcal{M}_e(u') = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}$$

Preuve.
On considère la famille $e = (x, u(x), u^2(x), \ldots, u^{d-1}(x))$. Ces vecteurs sont tous dans $K[u](x)$. Nous allons montrer qu'il s'agit d'une famille libre et génératrice.

e est une famille libre : Si on suppose que

$$\sum_{i=0}^{d-1} \alpha_i u^i(x) = 0$$

alors on peut écrire

$$\left(\sum_{i=0}^{d-1} \alpha_i u^i\right)(x) = 0.$$

Mais, par minimalité du degré du polynôme P, il faut alors que le polynôme $\sum_{i=0}^{d-1} \alpha_i X^i$, qui est de degré $\leqslant d-1$, soit nul (sinon on a une contradiction avec $\deg P = d$). Il vient alors que $\forall i \in [[0, d-1]], \alpha_i = 0$. La famille est déjà libre.

e est aussi une famille génératrice : Montrons que tout élément de $K[u](x)$ est combinaison linéaire des vecteurs de e. Soit $y \in K[u](x)$. Alors il existe $Q \in K[X]$ tel que $y = Q(u)(x)$. Faisons maintenant la division euclidienne de Q par P :

$$Q = PA + R,$$

avec R de degré au plus $d-1$. Mais alors

$$y = Q(u)(x) = R(u)(x),$$

et comme R est de degré au plus $d-1$, le vecteur y est combinaison linéaire de

$$x, u(x), u^2(x), \ldots, u^{d-1}(x),$$

les coefficients étant les coefficients du polynôme R. Cela montre que la famille e est génératrice. Nous pouvons désormais dire que e est une base de $K[u](x)$.

Observons que

$$\forall k \in [[0, d-2]], \quad u(u^k(x)) = u^{k+1}(x)$$

et

$$u(u^{d-1}(x)) = u^d(x) = -a_0 x - a_1 u(x) - a_2 u^2(x) \cdots - a_{d-1} u^{d-1}(x)$$

D'où la matrice de u' par rapport à la base $e = (x, u(x), \ldots, u^{d-1}(x))$:

$$\mathcal{M}_e(u') = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}.$$

$\square$

Lemme 4.
Le polynôme caractéristique de la matrice

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}$$

est égal à $X^d + a_{d-1}X^{d-1} + a_{d-2}X^{d-2} + \cdots + a_1 X + a_0$.

Preuve.
Calculons

$$\det \begin{pmatrix} X & 0 & 0 & \dots & 0 & a_0 \\ -1 & X & 0 & \cdots & 0 & a_1 \\ 0 & -1 & X & \cdots & 0 & a_2 \\ 0 & 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & X & a_{d-2} \\ 0 & 0 & 0 & \cdots & -1 & X + a_{d-1} \end{pmatrix}$$

en développant suivant la dernière colonne (numéro d). On trouve que le polynôme caractéristique de la matrice vaut

$$\left(\sum_{k=0}^{d-2} (-1)^{d+k+1} a_k \Gamma_k\right) + (X + a_{d-1})X^{d-1} = \left(\sum_{k=0}^{d-2} a_k (-1)^{d+k+1} \Gamma_k\right) + a_{d-1}X^{d-1} + X^d,$$

où, pour tout $k \in [[0, d-2]]$, Γ_k désigne le mineur associé au coefficient a_k. Pour conclure, il suffit de montrer que $\Gamma_k = (-1)^{d+k+1} X^k$. Or

$$\Gamma_k = \det \begin{pmatrix} X & & & & & & & \\ -1 & X & & & & & & \\ & \ddots & \ddots & & & & & \\ & & -1 & X & & & & \\ & & & & -1 & X & & \\ & & & & & -1 & \ddots & \\ & & & & & & \ddots & X \\ & & & & & & & -1 \end{pmatrix},$$

où, sur la diagonale, le nombre de coefficients X est égal à k, et le nombre de coefficients -1 est égal à $d-1-k$.
Observons qu'il s'agit d'une matrice **diagonale par blocs** à deux blocs. Le déterminant du premier bloc diagonal (de taille $k \times k$) vaut X^k, et le déterminant du second bloc diagonal (de taille $(d-1-k) \times (d-1-k)$ vaut $(-1)^{d-1-k}$. D'où

$$\Gamma_k = X^k (-1)^{d-1-k} = (-1)^{d+k+1} X^k,$$

étant donné que $k + 1 - (-1 - k) = 2k + 2$ est pair. $\square$

Nous pouvons maintenant passer à la preuve théorème de Cayley-Hamilton proprement dit.

Théorème de Cayley-Hamilton.
Soit E un espace vectoriel de dimension finie, et u un endomorphisme de E. On note C_u le polynôme caractéristique de u. Alors $C_u(u) = 0$.

Preuve.
Nous devons montrer $C_u(u) = 0$ ou encore :

$$\forall x \in E, C_u(u)(x) = 0.$$

Fixons un vecteur x quelconque dans E. On note $F = K[u](x)$ et $d = \dim F$. Nous savons que F est stable par u (lemme 2). Appelons P le polynôme (dépendant de x) du lemme 1. Si on note u' l'endomorphisme induit par u sur F, alors les lemmes 3 et 4 nous disent que

$$C_{u'} = P.$$

D'un autre côté, le polynôme $C_{u'}$ est un diviseur du polynôme C_u. En effet, on peut choisir une base $e = (e_1, \ldots, e_n)$ de E telle que $(e_1, \ldots, e_d)$ soit une base de F, alors la matrice de u dans cette base e prend la forme "par blocs"

$$\begin{pmatrix} A & M \\ 0 & B \end{pmatrix},$$

où A est une matrice carrée $d \times d$ et B une matrice carrée $(n-d) \times (n-d)$. Clairement

$$C_u = \det \begin{pmatrix} XI_d - A & -M \\ 0 & XI_{n-d} - B \end{pmatrix} = \det(XI_d - A) \cdot \det(XI_{n-d} - B)$$

Or A est une matrice représentant l'endomorphisme u' de F. Donc $\det(XI_d - A) = C_{u'}$. Il vient alors

$$C_u = C_{u'} \cdot \det(XI_{n-d} - B)$$

Le second facteur du membre de droite est évidemment un polynôme. Notons-le Q. Nous avons prouvé que $C_{u'} | C_u$. Nous pouvons conclure :

$$\begin{aligned} C_u(u)(x) &= (Q \cdot C_{u'})(u)(x) \\ &= (Q \cdot P)(u)(x) \\ &= (Q(u) \circ P(u))(x) \\ &= Q(u)(P(u)(x)) \\ &= Q(u)(0) \\ &= 0. \end{aligned}$$

□

Le théorème de Cayley-Hamilton possède aussi une version matricielle, qui découle évidemment de sa version pour les endomorphismes.

Théorème de Cayley-Hamilton.
Soit A une matrice carrée, et C_A son polynôme caractéristique. Alors $C_A(A) = 0$.

Tous les corollaires ci-dessous sont des conséquences immédiates du théorème de Cayley-Hamilton.

Corollaires.
1. Le polynôme caractéristique d'un endomorphisme u est un polynôme annulateur de u.
2. Le polynôme caractéristique d'une matrice A est un polynôme annulateur de A.
3. Le polynôme minimal de u divise le polynôme caractéristique de u.
4. Le polynôme minimal de A divise le polynôme caractéristique de A.
5. Le spectre de u est égal à l'ensemble des racines du polynôme minimal de u.
6. Le spectre de A est égal à l'ensemble des racines du polynôme minimal de A.
7. Le polynôme minimal de u est de degré inférieur ou égal à la dimension de l'espace vectoriel sur lequel u est défini.
8. Le polynôme minimal de A est de degré inférieur ou égal à la taille de la matrice A.

Illustrons le théorème de Cayley-Hamilton sur **le cas des matrices diagonalisables.** Soit A une matrice carrée diagonalisable. Soit $\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_k\} \subset K$. Alors le polynôme caractéristique de A est égal à

$$C_A = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \cdots (X - \lambda_k)^{m_k},$$

où $m_1, m_2, \ldots, m_k$ sont les multiplicités des k valeurs propres. Ce sont aussi les dimensions des espaces propres de chaque valeur propre (parce que A est diagonalisable).
D'un autre côté, le polynôme minimal de A, comme nous l'avons vu au chapitre précédent, est le polynôme

$$(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k).$$

Dans ce cas, on voit directement que le polynôme minimal est un diviseur du polynôme caractéristique, et que l'ensemble des racines du polynôme minimal est le spectre. C'est confirmé par le théorème de Cayley-Hamilton.

Au cours de la preuve du théorème de Cayley-Hamilton, nous avons établi un résultat intéressant concernant le polynôme caractéristique de l'endomorphisme induit sur un sous-espace stable. La preuve ayant été donnée ci-dessus, nous nous contentons de mettre en évidence l'énoncé.

Théorème.
Soit u un endomorphisme d'un espace vectoriel E de dimension finie, et F un sous-espace vectoriel de E stable par u. On note u_F l'endomorphisme induit par u sur F. Alors $C_{u_F} | C_u$.

Nous avons vu qu'une matrice à coefficients dans K est diagonalisable si et seulement si son polynôme caractéristique est scindé sur K et que la dimension de chaque sous-espace propre est égale à la multiplicité de la valeur propre dans le polynôme caractéristique.
Lorsque le polynôme caractéristique est scindé, ce n'est pas suffisant pour affirmer que la matrice est diagonalisable. On a néanmoins un joli résultat concernant la trace et le déterminant de la matrice.

Théorème.
Soit A une matrice carrée à coefficients dans K. On suppose que C_A est scindé sur K :

$$C_A = \prod_{\lambda \in \mathrm{Spec}(A)} (X - \lambda)^{m(\lambda)}.$$

où $m(\lambda)$ désigne la multiplicité de λ. Alors on a les égalités

$$\begin{aligned} \mathrm{tr}\, A &= \sum_{\lambda \in \mathrm{Spec}(A)} m(\lambda) \cdot \lambda, \\ \det A &= \prod_{\lambda \in \mathrm{Spec}(A)} \lambda^{m(\lambda)}. \end{aligned}$$

Preuve.
On suppose A de taille $n \times n$ et on pose $C_A = \det(XI_n - A)$.
Quand on développe ce déterminant par la somme à $n!$ termes, on voit que le seul terme qui contribue des monômes X^{n-1} est le terme qui correspond à la permutation identique $\sigma = \mathrm{id}_{[[1,n]]}$. Ce terme est

$$(X - A_{11})(X - A_{22}) \cdots (X - A_{nn}).$$

Donc le coefficient de X^{n-1} dans le polynôme caractéristique est $-\mathrm{tr}\, A$. Et comme le polynôme caractéristique s'écrit aussi

$$\underbrace{(X-\lambda_1)\cdots(X-\lambda_1)}_{m_1}\cdots\underbrace{(X-\lambda_k)\cdots(X-\lambda_k)}_{m_k},$$

son coefficient devant X^{n-1} vaut $-m_1\lambda_1 - \cdots - m_k\lambda_k$. D'où l'égalité

$$\sum_{\lambda \in \mathrm{Spec}(A)} m(\lambda) \cdot \lambda = \mathrm{tr}\, A.$$

Montrons maintenant la formule pour le déterminant. On trouve

$$C_A(0) = \det(0 \cdot I_n - A) = \det(-A) = (-1)^n \det A.$$

D'autre part

$$C_A(0) = \prod_{\lambda \in \mathrm{Spec}(A)} (0 - \lambda)^{m(\lambda)} = (-1)^{\sum_\lambda m(\lambda)} \prod_{\lambda \in \mathrm{Spec}(A)} \lambda^{m(\lambda)}.$$

D'où

$$(-1)^n \det A = (-1)^{\sum_\lambda m(\lambda)} \prod_{\lambda \in \mathrm{Spec}(A)} \lambda^{m(\lambda)},$$

et comme $n = \sum_\lambda m(\lambda)$, on peut conclure. $\square$

Terminons enfin par l'inégalité reliant la dimension d'un espace propre et la multiplicité de la valeur propre dans le polynôme caractéristique.

Théorème.
Soit u un endomorphisme d'un espace vectoriel de dimension finie, λ une valeur propre de u, et $m(\lambda)$ la multiplicité de λ dans le polynôme caractéristique C_u. On rappelle que E_λ est le sous-espace propre associé à la valeur propre λ. Alors on a les inégalités

$$1 \leqslant \dim E_\lambda \leqslant m(\lambda).$$

Preuve.
Le fait que $1 \leqslant \dim E_\lambda$ vient de la définition d'une valeur propre : son espace propre n'est jamais réduit au vecteur nul.
Montrons l'autre inégalité. Le sous-espace E_λ de E est évidemment stable par u. Donc le polynôme caractéristique de u_{E_λ} divise C_u. Or u_{E_λ} est évidemment l'homothétie $\lambda \cdot id$, agissant sur un espace vectoriel de dimension égale à $\dim E_\lambda$.
Donc le polynôme caractéristique de u_{E_λ} vaut

$$(X - \lambda)^{\dim E_\lambda}$$

et nous savons qu'il divise C_u. Donc

$$\dim E_\lambda \leqslant m(\lambda)$$

$\square$

Chapitre 5. Théorie de la réduction en dimension finie.

La théorie de la réduction étudie les endomorphismes u d'un espace vectoriel. On y cherche les conditions sous lesquelles il est possible de trouver une base dans laquelle l'endomorphisme u s'exprime par une matrice très simple. Les matrices très simples qui nous intéressent ici sont surtout les matrices diagonales, mais aussi les matrices triangulaires supérieures.

1 Endomorphismes diagonalisables

Nous connaissons déjà les endomorphismes diagonalisables. Ce sont des endomorphismes u d'un espace vectoriel pour lesquels il existe une base de vecteurs propres. En dimension finie, "u est diagonalisable" signifie qu'il existe une base e telle que $\mathcal{M}_e(u)$ soit une matrice diagonale.

Dans cette section, nous allons décrire plusieurs conditions nécessaires et suffisantes pour qu'un endomorphisme en dimension finie soit diagonalisable.

Commençons par rappeler les résultats que nous connaissons déjà à ce propos.

Théorème.
Soit u un endomorphisme d'un espace vectoriel de dimension finie sur le corps K.
u est diagonalisable si et seulement si :
1. Son polynôme caractéristique C_u est scindé sur K,
2. La multiplicité de chaque valeur propre dans le polynôme caractéristique est égale à la dimension de l'espace propre correspondant.

On peut reformuler ce théorème d'une manière plus courte :

Version équivalente.
Soit u un endomorphisme d'un espace vectoriel E de dimension finie.
u est diagonalisable si et seulement si la somme des dimensions des espaces propres de u est égale à la dimension de E, c'est-à-dire

$$\sum_{\lambda \in \mathrm{Spec}(u)} \dim E_\lambda = \dim E.$$

Preuve.
Soient 1. et 2. les conditions dans l'énoncé du théorème précédent. Il s'agit de prouver que (1. et 2.) est équivalent à l'affirmation

$$\sum_{\lambda \in \mathrm{Spec}(u)} \dim E_\lambda = \dim E.$$

Supposons 1. et 2. vraies. Alors

$$\sum_{\lambda \in \mathrm{Spec}(u)} \dim E_\lambda = \sum_{\lambda \in \mathrm{Spec}(u)} m(\lambda) = \deg C_u = \dim E,$$

car le polynôme caractéristique est scindé.

Réciproquement, supposons que

$$\sum_{\lambda \in \mathrm{Spec}(u)} \dim E_\lambda = \dim E.$$

Nous savons que $\dim E_\lambda \leqslant m(\lambda)$. D'un autre côté, par la théorie des polynômes, nous savons aussi que

$$\sum_{\lambda \in \mathrm{Spec}(u)} m(\lambda) \leqslant \deg C_u = \dim E.$$

Il est donc nécessaire d'avoir pour toute valeur propre : $\dim E_\lambda = m(\lambda)$. La condition 2. est vraie, et comme la somme des multiplicités des racines de C_u est égale à son degré, on peut dire également que la condition 1. est vraie.

□

En pratique, on applique souvent ce théorème lorsque le polynôme caractéristique est scindé et à racines simples :

Théorème.
Soit u un endomorphisme de E (de dimension finie sur K).
Si le polynôme caractéristique de u est scindé, et que toutes ses racines sont simples, alors u est diagonalisable et tous ses espaces propres sont de dimension 1.

Preuve. Il suffit de montrer que les conditions 1. et 2. sont vérifiées. La première condition fait partie de l'hypothèse. Pour la seconde, il suffit de montrer que chaque espace propre est de dimension 1. Or nous savons qu'un espace propre, par définition, est au moins de dimension 1. D'un autre côté, par l'inégalité

$$\dim E_\lambda \leqslant m(\lambda)$$

et l'hypothèse $m(\lambda) = 1$, on déduit que $\dim E_\lambda = 1$. L'endomorphisme u est bien diagonalisable, et ses espaces propres sont tous de dimension 1. □

Nous avons également rencontré une classe très spéciale d'endomorphismes diagonalisables, à savoir les endomorphismes auto-adjoints (ou symétriques). Ces endomorphismes n'existent que dans le contexte des espaces euclidiens et des espaces hermitiens. Il faut donc en particulier que $K = \mathbb{R}$ (ou $K = \mathbb{C}$).

Théorème spectral réel.
Soit u un endomorphisme auto-adjoint d'un espace euclidien E (donc sur le corps $\mathbb{R}$).
Alors u est diagonalisable sur, et il existe une base orthonormée de vecteurs propres.

Théorème spectral complexe.
Soit u un endomorphisme auto-adjoint d'un espace hermitien E (donc sur le corps $\mathbb{C}$).
Alors u est diagonalisable, et il existe une base orthonormée de vecteurs propres.
En outre, toutes les valeurs propres de u sont réelles.

Revenons au cas général.

En rapport avec le polynôme minimal, nous avons également vu cette autre caractérisation des endomorphismes diagonalisables.

Théorème.
Un endomorphisme u en dimension finie est diagonalisable si et seulement si le polynôme minimal de u est scindé et à racines simples.

Nous pouvons en déduire cet autre énoncé, très proche.

Théorème.
Un endomorphisme u en dimension finie est diagonalisable si et seulement s'il existe un polynôme annulateur de u scindé et à racines simples (et non nul) .

Preuve.
La partie "seulement si" est évidente : il suffit de prendre pour polynôme annulateur le polynôme minimal !
Supposons qu'il existe un polynôme annulateur P de u scindé et à racines simples. Le polynôme minimal de u est par définition un diviseur de P. Or tout diviseur d'un polynôme scindé non nul à racines simples est encore scindé à racines simples. Donc le polynôme minimal de u est un polynôme scindé à racines simples. Ce qui entraîne que u est diagonalisable.
□

Comme application immédiate, on a : tout projecteur p est diagonalisable. En effet, p possède le polynôme annulateur $X^2 - X$, et $X^2 - X = X(X-1)$ est scindé à racines simples.
De la même façon, toute symétrie s d'un espace vectoriel est diagonalisable, car le polynôme $X^2 - 1 = (X+1)(X-1)$ est annulateur et scindé et à racines simple.

Ci-dessous encore une version très proche de ce théorème.

Théorème.
Un endomorphisme u en dimension finie est diagonalisable si et seulement si le polynôme $\prod_{\lambda \in \mathrm{Spec}(u)} (X-\lambda)$ est annulateur de u.

Preuve.
Commençons par "si". Par hypothèse, u possède un polynôme annulateur scindé à racines simples. Donc u est diagonalisable.
Poursuivons par "seulement si". u est supposé diagonalisable. On sait alors que le polynôme minimal de u est

$$\prod_{\lambda \in \mathrm{Spec}(u)} (X-\lambda)$$

Or le polynôme minimal de u est par définition un polynôme annulateur de u. □

Nous avons vu que si F est un sous-espace vectoriel de E stable par un endomorphisme u, alors le polynôme caractéristique C_{u_F} de l'endomorphisme induit sur F est un diviseur du polynôme caractéristique C_u. La même chose vaut pour le polynôme minimal :

Théorème.
Soit u un endomorphisme de E, et F un sous-espace stable par u. Alors le polynôme minimal de u_F divise le polynôme minimal de u.

Preuve.
Le polynôme minimal de u, noté M_u est un polynôme annulateur de u, donc c'est aussi un polynôme annulateur de u_F. Il en découle que M_{u_F} divise M_u. □

Comme corollaire immédiat, nous avons que la propriété "être diagonalisable" passe aux sous-espaces stables :

Théorème.
Soit u un endomorphisme diagonalisable, et F un sous-espace vectoriel stable par u. Alors u_F est également diagonalisable.

Preuve.
Il suffit de montrer que le polynôme minimal de u_F est scindé à racines simples. Or par hypothèse, le polynôme minimal de u est scindé à racines simples. Et comme M_{u_F} est un diviseur de M_u, on peut dire que M_{u_F} est également scindé à racines simples. □

2 Endomorphismes trigonalisables

Définition.
Soit u un endomorphisme d'un espace vectoriel E de dimension finie. On dit que u est **trigonalisable** *s'il existe une base e de E telle que $\mathcal{M}_e(u)$ soit une matrice triangulaire supérieure.*

Cela revient à dire que la base e possède la propriété suivante : Pour tout $j \in [[1, n]]$, il existe j scalaires $(t_{ij})_{i \in [[1,j]]}$ tels que

$$u(e_j) = \sum_{i=1}^{j} t_{ij} e_i.$$

En particulier, le premier vecteur e_1 est un vecteur propre de u, de valeur propre t_{11}.

L'endomorphisme u (en dimension n) est trigonalisable si et seulement s'il existe une chaîne de sous-espaces vectoriels

$$\{0\} = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_{n-1} \subset F_n = E$$

avec $\forall k, \dim F_k = k$, telle que chaque F_k soit stable par u. En effet, si u est trigonalisable, prenons une base e dans laquelle la matrice de u est triangulaire supérieure. Il suffit alors de définir les sous-espaces F_k par

$$F_k = \text{Vect}(e_1, \ldots, e_k)$$

Réciproquement, s'il existe une chaîne de sous-espaces vectoriels comme ci-dessus, on peut construire une base e de E de la manière suivante : On choisit e_1 comme étant une base de F_1, puis on complète la famille libre (e_1) de F_2 en une base (e_1, e_2) de F_2, puis on complète la famille libre (e_1, e_2) de F_3 en une base de F_3 et ainsi de suite. Dans cette base e, la matrice de u sera triangulaire supérieure.

Comme pour les endomorphismes diagonalisables, nous cherchons des conditions nécessaires ou suffisantes pour qu' un endomorphisme soit trigonalisable.

Observons que les conditions suffisantes de diagonalisabilité sont aussi des conditions suffisantes de trigonalisabilité, car : si u est diagonalisable, alors u est trigonalisable.

Notre premier résultat donne une condition nécessaire et suffisante pour que u soit trigonalisable :

Théorème.
1. Un endomorphisme u d'un espace vectoriel de dimension finie est trigonalisable si et seulement s'il possède un polynôme annulateur scindé non nul.
2. Si e est une base telle que $\mathcal{M}_e(u)$ soit triangulaire, alors la diagonale de cette matrice est formée par les valeurs propres de u (le nombre d'apparitions de chaque valeur propre sur la diagonale étant égal à sa multiplicité dans le polynôme caractéristique).

Preuve.
Nous nous occupons d'abord de la partie 1.
Prouvons "seulement si". On suppose u trigonalisable. Alors il existe une base e, de cardinal n, telle que la matrice $A := \mathcal{M}_e(u)$ soit triangulaire supérieure. Par les propriétés du déterminant d'une matrice triangulaire, on voit alors que le polynôme caractéristique de u (qui est celui de A) est égal à

$$C_u = \prod_{i=1}^{n}(X - A_{ii}).$$

Ce polynôme est scindé. Or, par le théorème de Cayley-Hamilton, C_u est un polynôme annulateur de u. Donc u possède un polynôme annulateur scindé.

Prouvons "si". On va montrer la proposition par récurrence sur la dimension de E. Si E est de dimension 0 ou 1, il n'y a rien à montrer, car un endomorphisme en dimension $\leqslant 1$ est toujours trigonalisable, quel que soit le corps. On suppose le résultat vrai pour toutes les dimensions $\leqslant n-1$, et on va montrer qu'il reste vrai pour la dimension n.
Soit P un polynôme annulateur scindé de u, qui existe par hypothèse. Choisissons une valeur propre λ de u. Il y en a au moins une, car le spectre de u est égal à l'ensemble des racines du polynôme minimal M_u, et le polynôme minimal est un diviseur de P, qui est scindé. Donc M_u est scindé (et non constant par définition). Il admet donc une racine λ. Nous considérons maintenant le sous-espace vectoriel

$$F = \mathrm{Im}\ (u - \lambda \mathrm{id}).$$

Par le théorème du rang, la dimension de F est égale à $n - \dim E_\lambda \leqslant n-1$. La dernière inégalité est vraie, car λ est bien une valeur propre de u. Le sous-espace vectoriel F est stable par u, donc on peut former l'endomorphisme induit u_F, et on sait que le polynôme minimal M_{u_F} divise M_u. Mais M_u est scindé, ce qui entraîne que M_{u_F} est scinde. Ce dernier étant un annulateur de u_F, on a établi l'existence d'un polynôme annulateur scindé de u_F. D'où la possibilité de recourir à l'hypothèse de récurrence.
Notons $p = \dim F$. On sait qu'il existe une base $(e_1, \ldots, e_p)$ de F telle que pour tout $j \in [[1,p]]$, il existe des scalaires $(t_{ij})_{1 \leqslant i \leqslant j}$ avec $u(e_j) = \sum_{i=1}^{j} t_{ij} e_i$. Complétons maintenant en une base de E de manière quelconque. Soit $j > p$. Alors on a

$$u(e_j) = (u - \lambda \mathrm{id})(e_j) + \lambda e_j.$$

Par définition de F, on a $(u - \lambda \mathrm{id})(e_j) \in F = \mathrm{Vect}(e_1, \ldots, e_p)$. Il existe alors des scalaires $(a_{ij})_{i \in [[1,p]]}$ tels que

$$u(e_j) = \sum_{i=1}^{p} a_{ij} e_i + \lambda e_j = \sum_{i=1}^{j} t_{ij} e_j,$$

où les (t_{ij}) sont définis par :

$$\forall i \in [[1,p]], t_{ij} = a_{ij}, \qquad \forall i \in [[p+1, j-1]], t_{ij} = 0, \qquad t_{jj} = \lambda.$$

La base e de E possède bien la propriété que $\mathcal{M}_e(u)$ est triangulaire supérieure, ce qui achève la preuve de la partie 1.
Il reste la partie 2. à démontrer. Soit T une matrice triangulaire supérieure qui représente u dans une certaine base de E. Alors le polynôme caractéristique de u est égal à C_T, qui se calcule facilement. D'où

$$C_u = \prod_{i=1}^{\dim E} (X - T_{ii}).$$

Les listes des racines (multiplicités comprises) des deux membres sont identiques. Donc les racines de C_u coïncident avec les coefficients diagonaux de la matrice T. $\square$

Nous en déduisons immédiatement les deux corollaires ci-dessous.

Corollaire 1.
Un endomorphisme u est trigonalisable $\iff$ le polynôme caractéristique de u est scindé.

Preuve.
En effet, le polynôme caractéristique est annulateur de l'endomorphisme (Cayley-Hamilton). Ceci établit la partie $\Leftarrow$. La réciproque est la partie 2 du théorème précédent. □

Observons que l'on peut remplacer le mot "caractéristique" par "minimal" dans le corollaire 1. Le lecteur fournira facilement la preuve.

Corollaire 2.
Tout endomorphisme d'un espace vectoriel sur le corps $\mathbb{C}$ est trigonalisable.

Preuve.
En effet, le polynôme caractéristique ayant ses coefficients dans $\mathbb{C}$, il est scindé (théorème de Gauss-d'Alembert). □

3 Un algorithme de trigonalisation sur $\mathbb{C}$

Comme toute matrice à coefficients dans $\mathbb{C}$ est trigonalisable, montrons comment faire en pratique pour réduire une matrice complexe quelconque à une matrice triangulaire. Cela signifie : étant donnée une matrice complexe A, chercher deux matrices complexes P, T avec P inversible et T triangulaire supérieure, telles que
$$T = P^{-1}AP$$
ou encore $A = PTP^{-1}$.

Soit A_n une matrice complexe de taille $n \times n$ (avec $n \geqslant 1$). On sait qu'il existe au moins une valeur propre λ_n de A_n (grâce à Gauss-d'Alembert). Choisissons une matrice-colonne v non nulle telle que $A_n \cdot v = \lambda_n v$, puis choisissons un entier $j \in [[1, n]]$ tel que le coefficient de ligne j dans v soit non nul. Il est clair que j existe. Puis complétons le vecteur v en une matrice carrée $n \times n$ en lui adjoignant dans l'ordre tous les vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbb{C})$ **sauf** le vecteur numéro j. On obtient alors une matrice P_n de la forme

$$P_n = \begin{bmatrix} v_1 & 1 & 0 & \cdots & & & \cdots & 0 \\ v_2 & 0 & 1 & \cdots & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & & & & \\ v_{j-1} & & & & 1 & & & \\ v_j & & & & 0 & 0 & & \\ v_{j+1} & & & & & 1 & & \\ \vdots & \vdots & & & & & \ddots & \vdots \\ v_n & 0 & \cdots & & & & \cdots & 1 \end{bmatrix}.$$

En développant son déterminant suivant la première colonne, on trouve qu'il vaut $(-1)^{j+1}v_j$. La matrice P_n est alors inversible. Comme $A_n v = \lambda_n v$, on trouve que

$$A_n \cdot \begin{bmatrix} v_1 & 1 & 0 & \cdots & & & \cdots & 0 \\ v_2 & 0 & 1 & \cdots & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & & & & \\ v_{j-1} & & & & 1 & & & \\ v_j & & & & 0 & 0 & & \\ v_{j+1} & & & & & 1 & & \\ \vdots & \vdots & & & & & \ddots & \vdots \\ v_n & 0 & \cdots & & & & \cdots & 1 \end{bmatrix} = \begin{bmatrix} v_1 & 1 & 0 & \cdots & & & \cdots & 0 \\ v_2 & 0 & 1 & \cdots & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & & & & \\ v_{j-1} & & & & 1 & & & \\ v_j & & & & 0 & 0 & & \\ v_{j+1} & & & & & 1 & & \\ \vdots & \vdots & & & & & \ddots & \vdots \\ v_n & 0 & \cdots & & & & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_n & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix},$$

où les symboles $*$ désignent des nombres complexes quelconques. Notons B_n la matrice de droite. On a la relation

$$A_n = P_n B_n P_n^{-1}.$$

La matrice B_n n'est en général pas triangulaire supérieure. Écrivons-la sous la forme de blocs

$$B_n = \begin{bmatrix} \lambda_n & L_{n-1} \\ 0 & A_{n-1} \end{bmatrix},$$

où L_{n-1} est une matrice-ligne $1 \times (n-1)$, et A_{n-1} une matrice carrée $(n-1) \times (n-1)$.

Résumons. En partant de la matrice A_n, nous avons défini successivement :
un nombre complexe λ_n,
une matrice inversible P_n de taille $n \times n$,
une matrice carrée B_n de taille $n \times n$,
une matrice-ligne L_{n-1} de taille $1 \times (n-1)$,
une matrice carrée A_{n-1} de taille $(n-1) \times (n-1)$.

L'idée de notre algorithme est de refaire les opérations ci-dessus en partant de la matrice A_{n-1} pour obtenir les données $\lambda_{n-1}, P_{n-1}, B_{n-1}, L_{n-2}, A_{n-2}$, recommencer avec A_{n-2} et ainsi de suite jusqu'à obtenir A_1.

En particulier on peut écrire les égalités matricielles

$$\begin{aligned} A_n &= P_n B_n P_n^{-1} \\ A_{n-1} &= P_{n-1} B_{n-1} P_{n-1}^{-1} \\ \vdots &= \vdots \\ A_2 &= P_2 B_2 P_2^{-1}. \end{aligned}$$

Observons que la matrice $B_2 = \begin{bmatrix} \lambda_2 & L_1 \\ 0 & A_1 \end{bmatrix}$ est triangulaire supérieure.

On définit maintenant les matrices inversibles $Q_n, Q_{n-1}, \ldots, Q_2$ (toutes de taille $n \times n$) par $Q_n = P_n$ et, pour tout $k \in [[2, n-1]]$,

$$Q_k = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & P_k \end{bmatrix} \in \mathrm{GL}(n, \mathbb{C}),$$

le nombre de coefficients 1 étant égal à $n - k$. Il est clair que Q_k et P_k ont le même déterminant, ce qui explique que Q_k est toujours inversible.

Définissons encore une fois une suite de matrices par récurrence descendante. On pose cette fois $C_n = A_n = A$ et, pour tout $k \in [[1, n-1]]$, $C_{k-1} := Q_k^{-1} C_k Q_k$. Les matrices $C_n, C_{n-1}, \ldots, C_1$ sont toutes de taille $n \times n$.

Prouvons par récurrence sur k que la matrice C_k est de la forme par blocs

$$C_k = \begin{bmatrix} T_{n-k} & * \\ 0 & A_k \end{bmatrix},$$

où T_{n-k} est une matrice triangulaire supérieure de taille $(n-k) \times (n-k)$. Pour $k = n$, il n'y a rien à prouver. Ensuite, avec l'hypothèse de récurrence, on obtient

$$C_{k-1} = Q_k^{-1} \begin{bmatrix} T_{n-k} & * \\ 0 & A_k \end{bmatrix} Q_k = \begin{bmatrix} T_{n-k} & *P_k \\ 0 & P_k^{-1} A_k P_k \end{bmatrix} = \begin{bmatrix} T_{n-k} & *P_k \\ 0 & B_k \end{bmatrix}.$$

Et comme $B_k = \begin{bmatrix} \lambda_k & L_{k-1} \\ 0 & A_{k-1} \end{bmatrix}$, le résultat suit. CQFD

En particulier, il découle de ce résultat (pour $k = 1$) que la matrice C_1 est triangulaire supérieure.

Nous avons alors

$$C_1 = Q_2^{-1} C_2 Q_2 = Q_2^{-1} Q_3^{-1} C_2 Q_3 Q_2 = \cdots = Q_2^{-1} \cdots Q_n^{-1} A Q_n \cdots Q_2.$$

Il suffit maintenant de poser $T = C_1$, qui est triangulaire supérieure, et $P = Q_n \cdots Q_2$, de sorte que

$$T = P^{-1} A P,$$

comme souhaité.

Exemple.

Illustrons notre algorithme sur la matrice complexe 4×4

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & -3 \\ -1 & 0 & 1 & 2 \\ 0 & -1 & -1 & 1 \end{bmatrix} = A_4.$$

Calculons le polynôme caractéristique de A. Des calculs de déterminant nous donnent

$$\det \begin{bmatrix} X-1 & -2 & -1 & 0 \\ -1 & X-1 & 0 & 3 \\ 1 & 0 & X-1 & -2 \\ 0 & 1 & 1 & X-1 \end{bmatrix} = X^2 (X-2)^2.$$

Commençons par choisir $\lambda_4 = 0$. On choisit un vecteur propre correspondant à cette valeur propre, par exemple $\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$. On peut alors prendre (après élimination du vecteur numéro 1 de la base canonique)

$$P_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ceci donne ensuite

$$A_4 P_4 = P_4 \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 3 & 1 & -3 \\ 0 & -2 & 0 & 2 \\ 0 & -1 & -1 & 1 \end{bmatrix} = P_4 B_4.$$

On recommence avec $A_3 = \begin{bmatrix} 3 & 1 & -3 \\ -2 & 0 & 2 \\ -1 & -1 & 1 \end{bmatrix}$. Son polynôme caractéristique vaut $X(X-2)^2$. Prenons $\lambda_3 = 0$ et, comme vecteur propre associé, la colonne $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. On peut choisir

$$P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Ceci nous donne la matrice B_3 :

$$A_3P_3 = P_3 \begin{bmatrix} 0 & 1 & -3 \\ 0 & 0 & 2 \\ 0 & -2 & 4 \end{bmatrix} = P_3B_3.$$

On recommence avec $A_2 = \begin{bmatrix} 0 & 2 \\ -2 & 4 \end{bmatrix}$. Son polynôme caractéristique vaut $(X-2)^2$. Prenons $\lambda_2 = 2$ et, comme vecteur propre associé, la colonne $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. On peut choisir

$$P_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Ceci donne enfin la matrice B_2 :

$$A_2P_2 = P_2 \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} = P_2B_2.$$

Maintenant on pose

$$Q_4 = P_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad Q_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \qquad Q_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Il reste encore à calculer C_3, C_2, C_1 (cette dernière sera une matrice triangulaire semblable à A) : On a $C_4 = A$ par définition, puis, grâce au résultat sur la forme des C_k :

$$C_3 = B_4 = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 3 & 1 & -3 \\ 0 & -2 & 0 & 2 \\ 0 & -1 & -1 & 1 \end{bmatrix}.$$

Ici $* = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$ et donc

$$*P_3 = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}.$$

Il en découle que

$$C_2 = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 4 \end{bmatrix}.$$

Ici $* = \begin{bmatrix} 1 & 0 \\ 1 & -3 \end{bmatrix}$ et donc

$$*P_2 = \begin{bmatrix} 1 & 0 \\ 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & -3 \end{bmatrix}.$$

D'où finalement

$$C_1 = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -2 & -3 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

On a donc $T = P^{-1}AP$ avec

$$T = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -2 & -3 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad P = Q_4Q_3Q_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

On pourra vérifier qu'on a bien $AP = PT$, c'est-à-dire

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & -3 \\ -1 & 0 & 1 & 2 \\ 0 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -2 & -3 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Sommaire du cours d'Algèbre 4