

7.1 Proposition Soit V et W des $\mathbb{K}$ -esp. vectoriels de dimension m et n , et $f: V \rightarrow W$ $\mathbb{K}$ -linéaire. Alors les conditions suivantes sont équivalentes :

(a) f est de rang r ;

(b) Il existe des bases $B = \{e_1, \dots, e_m\}$ de V et $C = \{f_1, \dots, f_n\}$ de W avec

$$f(e_i) = \begin{cases} f_i & \text{si } 1 \leq i \leq r, \\ 0 & \text{si } i > r. \end{cases}$$

(en d'autres termes :

$$\{f\}_B^C = \left(\begin{array}{cc} \underbrace{\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}}_r & \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}}_{m-r} \\ \underbrace{0}_{r} & \underbrace{0}_{n-r} \end{array} \right) \begin{matrix} \left. \vphantom{\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}} \right\} r \\ \left. \vphantom{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}} \right\} n-r \end{matrix}$$

preuve : (b) $\Rightarrow$ (a) : On a

$\text{Im}(f) = \text{Vect}(\{f_1, \dots, f_r\})$, et comme $\{f_1, \dots, f_r\}$ fait partie d'une base, il est libre ; donc $\dim \text{Im}(f) = r$.

(a) $\Rightarrow$ (b). Par le Théorème du rang 6.5, nous avons $\dim \text{Ker}(f) = m - \dim \text{Im}(f) = m - r$.

On choisit une base $\{e_{r+1}, e_{r+2}, \dots, e_m\}$ de $\text{Ker}(f)$, grâce à 5.32, on peut la compléter en une base $B = \{e_1, \dots, e_m\}$ de V . On pose $f_1 = f(e_1), \dots, f_r = f(e_r)$.

Alors $\{f_1, \dots, f_r\}$ engendre $\text{Im}(f)$ (de dimension r), donc est libre ; on peut le compléter en une base $C = \{f_1, \dots, f_n\}$. Alors B et C satisfont à (b). #

7.2 Remarque : Donc, si on peut choisir librement une base B de V et une base C de W , on peut s'arranger pour que $\{f\}_B^C$ ait une forme très simple, déterminée par le rang de f . On va s'intéresser maintenant aux endomorphismes $f: V \rightarrow V$, et à la question : peut-on trouver une base $B \subset V$ telle que $\{f\}_B^B$ ait une

forme aussi simple que possible, qui nous permette de "comprendre" f de façon plus "géométrique". Nous allons voir que cela revient à classer les matrices de $M_n(\mathbb{K})$ à similitude près. Dans ce chapitre, nous fixons un $n \in \mathbb{N}$ et un espace vectoriel V de dimension n (typiquement $\mathbb{K}^n$).

7.3 Proposition : Soit $A, B \in M_n(\mathbb{K})$. Alors les conditions suivantes sont équivalentes :

- (a) Il existe un endomorphisme $f: V \rightarrow V$, et des bases A et B de V , avec $A = \{f\}_A^A$ et $B = \{f\}_B^B$
 (b) Il existe une matrice $P \in GL_n(\mathbb{K})$ avec $B = P^{-1} A P$.

Preuve : (a) $\Rightarrow$ (b) Il suffit de prendre $P = \{id\}_B^A$, la matrice de passage de A à B . Alors

$$B = \{f\}_B^B = \{id\}_A^B \{f\}_A^A \{id\}_B^A = P^{-1} A P \text{ par 6.22.e.}$$

(b) $\Rightarrow$ (a) Soit $A = \{e_1, \dots, e_n\}$ une base quelconque de V . Soit $f: V \rightarrow V$ défini par $f(e_i) = a_{1i}e_1 + \dots + a_{ni}e_n$, avec $A = (a_{ij})$. Donc on obtient $\{f\}_A^A = A$.

Similairement, si $P = (p_{ij})$, soit $p: V \rightarrow V$ défini par $p(e_i) = p_{1i}e_1 + \dots + p_{ni}e_n$. Comme P est inversible, $B = \{f_1 = p(e_1), \dots, f_n = p(e_n)\}$ est une base de V , et l'on a $P = \{id\}_B^A$.

$$\text{On en déduit } B = P^{-1} A P = \{id\}_A^B \{f\}_A^A \{id\}_B^A = \{f\}_B^B, \#$$

7.4 Définition : Soient $A, B \in M_n(\mathbb{K})$. On dit que A et B sont semblables si l'une des conditions équivalentes (a) ou (b) de 7.3 est vérifiée. On le note $A \approx B$.

7.5 Exemple : les matrices $A = \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix}$ et

(83)

$B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ sont semblables : en effet, si l'on prend $P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, on trouve $B = P^{-1} A P$ dans $M_2(\mathbb{R})$.

On remarque que B a une forme plus simple que A ,

7.6 Définition: (i) Une matrice $A = (a_{ij}) \in M_n(\mathbb{K})$ est dite diagonale si $a_{ij} = 0$ pour tous $1 \leq i, j \leq n$ avec $i \neq j$:

$$A = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$

On la note alors $A = \text{Diag}(a_1, \dots, a_n)$ avec $a_i := a_{ii}$.

(ii) Une matrice $A \in M_n(\mathbb{K})$ est dite diagonalisable s'il existe une matrice diagonale D avec $A \approx D$.

(iii) Soit $f: V \rightarrow V$ $\mathbb{K}$ -linéaire, $\dim V = n$.

On dit que f est diagonalisable s'il existe une base $\mathcal{B}$ de V avec $\{f\}_{\mathcal{B}}^{\mathcal{B}} \in M_n(\mathbb{K})$ diagonale.

7.7 Exemple : Soit $A = \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix} \in M_2(\mathbb{R})$ comme dans 7.5. Alors A est diagonalisable, puisque $A \approx B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ et B est diagonale.

Pour le reste du chapitre, nous supposons $f: V \rightarrow V$ $\mathbb{K}$ -linéaire donnée (avec $\dim V = n$).

7.8 Définition Soit $\lambda \in \mathbb{K}$. On dit que λ est une valeur propre de f s'il existe $x \in V$ avec $x \neq 0$ et $f(x) = \lambda x$. Dans ce cas, on dit qu'un tel x est un vecteur propre de f (pour la valeur propre λ).

Remarquons que si $x \in V$ est un vecteur propre de f , il l'est pour une seule valeur propre: $f(x) = \lambda x = \mu x \Rightarrow \lambda = \mu$, puisque $x \neq 0$.

7.9 Exemples: (a) Considérons $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avec

$g(x, y) = (5x + 3y, -6x - 4y)$ (Donc si B est la base canonique de $\mathbb{R}^2$, $\{g\}_B^B = A = \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix}$ de 7.5).

Alors

$$g(1, -1) = (2, -2) = 2 \cdot (1, -1)$$

$$g(-1, 2) = (1, -2) = (-1) \cdot (-1, 2)$$

Donc $(1, -1) \in \mathbb{R}^2$ est un vecteur propre de g pour la valeur propre 2, et $(-1, 2)$ est un vecteur propre de g pour la valeur propre -1. Nous savons par 7.7 que g est diagonalisable.

(b) Si $\text{Ker}(f) \neq 0$, alors 0 est valeur propre de f , et tout $x \in \text{Ker}(f) \setminus \{0\}$ est un vecteur propre pour la valeur propre 0. En particulier, f est injectifssi 0 n'est pas valeur propre de f .

(c) Si $\lambda \in \mathbb{K}$, et si $h: V \rightarrow V$, $x \mapsto \lambda x$, alors, si $V \neq 0$, λ est valeur propre de h , et tout $x \in V \setminus \{0\}$ est un vecteur propre de h pour λ .

7.10 Proposition: Les conditions suivantes sont équivalentes:

(a) $f: V \rightarrow V$ est diagonalisable

(b) Pour toute base B de V , $\{f\}_B^B$ est diagonalisable

(c) V admet une base dont tous les éléments sont des vecteurs propres de f .

preuve: exercice facile. #

On en déduit: l'étude de la diagonalisation se ramène à la recherche de vecteurs propres.

7.11 Proposition: Soit $A \in M_n(\mathbb{K})$. Alors la fonction

$C_A: \mathbb{K} \rightarrow \mathbb{K}$, $t \mapsto \det(A - t I_n)$ est une

fonction polynomiale de deg. n en t . On appelle le polynôme correspondant $C_A(x) \in \overset{\mathbb{K}}{P}(x)$ le polynôme caractéristique de A .

preuve : il suffit d'utiliser la définition du déterminant ; nous omettons la preuve. #.

7.12 Proposition : Si $A \approx B$ dans $M_n(K)$, alors

$$C_A(x) = C_B(x) \text{ dans } P^K[x].$$

En particulier, si A et B sont des bases de V , alors

$$C_{\{f\}_A}(x) = C_{\{f\}_B}(x).$$

On note $C_f(x)$ ce polynôme, et on l'appelle le polynôme caractéristique de l'endomorphisme f .

preuve : Si $A \approx B$, il existe $P \in GL_n(K)$ avec
 $B = P^{-1} A P$. Si $t \in K$, alors $P^{-1} t I_n P = t I_n$, donc
 $B - t I_n = P^{-1} A P^{-1} - P^{-1} t I_n P = P^{-1} (A - t I_n) P$.
 Or $\det(P^{-1} n P) = \det(P^{-1}) \cdot \det(n) \cdot \det(P) = \det(n)$.
 pour tout $n \in M_n(K)$. Donc $C_B(t) = C_A(t) \quad \forall t \in K$,
 et par suite $C_A(x) = C_B(x)$. #.

7.13 Proposition : Les conditions suivantes sur $\lambda \in K$ sont équivalentes :

(a) λ est valeur propre de f .

(b) λ est racine de $C_f(x)$: $C_f(\lambda) = 0$.

preuve : (a) $\Leftrightarrow \exists x \neq 0$ avec $f(x) = \lambda x \Leftrightarrow$
 $\exists x \neq 0$ avec $(f - \lambda \text{Id}_V)(x) = 0 \Leftrightarrow f - \lambda \text{Id}_V$ n'est pas injectif $\Leftrightarrow f - \lambda \text{Id}_V$ n'est pas un isomorphisme
 $\Leftrightarrow$ Si B est une base de V , $\{f - \lambda \text{Id}_V\}_B^B \in M_n(K)$ n'est pas inversible $\Leftrightarrow \det \{f - \lambda \text{Id}_V\}_B^B = 0$
 $\Leftrightarrow C_f(\lambda) = 0$. #

7.14 Exemple : Soit $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ comme en 7.3. Alors

$$C_g(x) = \det(A - x I_2) = \det\left(\begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}\right) = \det\begin{pmatrix} 5-x & 3 \\ -6 & -4-x \end{pmatrix} = x^2 - x - 2 = (x+1)(x-2).$$

Nous avons vu en 7.9 que 2 et -1 sont des valeurs propres de g , et ce sont effectivement des racines de $C_g(x)$.

7.15 Définition: Soit $f: V \rightarrow V$ comme ci-dessus, et $P(x) \in P^{\mathbb{K}}[x]$

avec $P(x) = a_0 + a_1 x + \dots + a_m x^m$. On définit

$$(a) \quad f^0 := \text{id}_V, \quad f^1 = f \quad \text{et} \quad f^m = f \circ f^{m-1} = \overbrace{f \circ \dots \circ f}^m,$$

donc $f^m \in \mathcal{L}(V, V)$.

$$(b) \quad P(f) := a_0 f^0 + a_1 f^1 + \dots + a_m f^m \in \mathcal{L}(V, V).$$

Similairement, si $A \in M_n(\mathbb{K})$, on définit

$$(c) \quad A^0 = I_n, \quad A^1 = A, \quad \text{et} \quad A^m = A \cdot A^{m-1} = \overbrace{A \cdot \dots \cdot A}^m,$$

donc $A^m \in M_n(\mathbb{K})$.

$$(d) \quad P(A) = a_0 I_n + a_1 A + \dots + a_m A^m \in M_n(\mathbb{K}).$$

7.16 Remarques: (a) Si $A = \{f\}_B^B$, $B \subset V$ une base, alors on a aussi $P(A) = \{P(f)\}_B^B$: cela suit de 6.17 et 6.18 a.

(b) Si x est un vecteur propre de f pour la valeur propre 2, alors x est un vecteur propre de $P(f)$ pour la v.p. $P(2) \in \mathbb{K}$: On a en effet $f^k(x) = f^{k-1}(2x) = f^{k-2}(2^2x) = \dots = 2^k x$, donc $P(f)(x) = a_0 x + a_1 2x + a_2 2^2 x + \dots + a_m 2^m x = P(2) \cdot x$.

(c) Si $A = \text{Diag}(a_1, \dots, a_n) \in M_n(\mathbb{K})$, alors $P(A) = \text{Diag}(P(a_1), \dots, P(a_n))$: suit de $\text{Diag}(a_1, \dots, a_n)^k = \text{Diag}(a_1^k, \dots, a_n^k)$.

7.17 Théorème: Si $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ sont deux-à-deux distincts, et si $x_1, \dots, x_k \in V$ sont des vecteurs propres de f avec $f(x_i) = \lambda_i x_i$, $1 \leq i \leq k$, alors $\{x_1, \dots, x_k\}$ est libre.

Preuve: Supposons $a_1, \dots, a_k \in \mathbb{K}$ avec $a_1 x_1 + \dots + a_k x_k = 0$.

Pour $1 \leq i \leq k$, posons $P_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^k \frac{x - \lambda_j}{\lambda_i - \lambda_j}$

On a donc $P_i(\lambda_j) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

On dit alors de 7.16. (b) : $P_i(f)(x_j) = P_i(\lambda_j) x_j = \begin{cases} 0 & j \neq i \\ x_i & j = i \end{cases}$

Mais alors $0 = P_i(f)(0) = P_i(f)(a_1 x_1 + \dots + a_k x_k) = a_1 P_i(f)(x_1) + \dots + a_k P_i(f)(x_k) = a_i x_i$. Comme $x_i \neq 0$ (c'est une valeur propre!), on a donc $a_i = 0$, pour $1 \leq i \leq k$. #

7.18 Definition : Si $\lambda \in \mathbb{K}$ est une valeur propre de f , alors on appelle $\text{Ker}(f - \lambda \text{id}_V)$ l'espace propre de f pour la valeur propre λ . On le note $E(\lambda)$ (c'est un sous-espace de f).

7.19 Exemple : Si on reprend $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de 7.9, avec $g(x, y) = (5x + 3y, -6x - 4y)$, on sait que g admet 2 et -1 comme valeurs propres. On a donc

$$E(2) = \text{Ker}(g - 2 \text{id}_{\mathbb{R}^2}) = \left\{ (x, y) \mid \begin{cases} 3x + 3y = 0 \\ -6x - 6y = 0 \end{cases} \right\} = \text{Vect}\{(1, -1)\}, \text{ et}$$

$$E(-1) = \text{Ker}(g + \text{id}_{\mathbb{R}^2}) = \left\{ (x, y) \mid \begin{cases} 6x + 3y = 0 \\ -6x - 3y = 0 \end{cases} \right\} = \text{Vect}\{(-1, 2)\}.$$

7.20 Théorème : Si $\lambda_1, \dots, \lambda_k$ sont des valeurs propres distinctes de f , alors $E(\lambda_1) \oplus \dots \oplus E(\lambda_k)$

preuve : Il suffit de montrer que si $x_i \in E(\lambda_i)$ et $x_1 + \dots + x_k = 0$, alors $x_1 = \dots = x_k = 0$. Supposons donc $x_1 + \dots + x_k = 0$, et supposons par l'absurde que pas tous les x_i sont nuls : disons $x_{i_1} \neq 0, \dots, x_{i_\ell} \neq 0$, les autres x_i étant nuls. Alors $x_{i_1} + \dots + x_{i_\ell} = x_1 + \dots + x_k = 0$; or $\{x_{i_1}, \dots, x_{i_\ell}\}$ est linéaire et non vide, une contradiction !

Donc tous les x_i sont nuls, et $E(\lambda_1) + \dots + E(\lambda_k) = E(\lambda_1) \oplus \dots \oplus E(\lambda_k)$ #

7.21 Exemples: (a) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (-y, x)$. (94)

Dans la base canonique $\mathcal{B}$ de $\mathbb{R}^2$, on a $\{f\}_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$

Donc $C_f(x) = \begin{vmatrix} -x & -1 \\ 1 & -x \end{vmatrix} = x^2 + 1$. On observe

que $C_f(x)$ n'admet aucune racine dans $\mathbb{R}$, donc f n'admet aucun vecteur propre (c'est clair "géométriquement": f est une rotation d'angle $\pi/2$).

(b) $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$, $(x, y) \mapsto (-y, x)$. Dans la base canonique $\mathcal{B}$ de $\mathbb{C}^2$, on a $\{f\}_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, et comme ci-dessus $C_f(x) = x^2 + 1 = (x - i)(x + i)$.

On calcule facilement

$$E(i) = \text{Ker}(f - i \text{Id}) = \text{Vect}\{(i, 1)\}$$

$$E(-i) = \text{Ker}(f + i \text{Id}) = \text{Vect}\{(1, i)\}.$$

Or $\mathcal{C} = \{(i, 1), (1, i)\}$ est une base de $\mathbb{C}^2$, donc $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ est diagonalisable, avec $\{f\}_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

7.22 Proposition: Rappelons que l'on a supposé $\dim V = n$.

Alors si f admet n valeurs propres distinctes, f est diagonalisable.

Preuve: Si $\lambda_1, \dots, \lambda_n$ sont des v.p. distinctes de f , et x_i un vecteur propre de f pour λ_i , alors $\{x_1, \dots, x_n\}$ est libre (7.17), donc est une base de V . Donc V admet une base de vecteurs propres, et f est diagonalisable. #

7.23 Exemples: Si f n'admet pas n valeurs propres distinctes, la situation est plus compliquée.

(a) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f = \text{id}$ admet comme seule valeur propre 1, mais est clairement diagonalisable ($E(1) = \mathbb{R}^2$).

(b) Considérons $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $g(x, y) = (x + y, y)$.

Si $\mathcal{B}$ = base can. de $\mathbb{R}^2$, $\{g\}_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

On en déduit $C_g(x) = \begin{vmatrix} 1-x & 1 \\ 0 & 1-x \end{vmatrix} = (1-x)^2$,

qui admet 1 comme racine double. On a

$$E(1) = \text{Ker}(g - \text{id}_M) = \{ (x, y) \mid y = 0 \} =$$

$\text{Vect}\{(1, 0)\}$, donc $E(1)$ est de dimension 1.

Ainsi, g n'est pas diagonalisable. En effet:

Lemme 7.24 voir page 87.

7.25 Proposition: Supposons que f admette une valeur propre, et soient $\{\lambda_1, \dots, \lambda_r\}$ les valeurs propres de f .

Alors f est diagonalisable ssi $V = E(\lambda_1) \oplus \dots \oplus E(\lambda_r)$.

preuve: Nous savons que les $E(\lambda_i)$ sont en somme directe par 7.20. Si $V = E(\lambda_1) \oplus \dots \oplus E(\lambda_r)$, alors

en faisant l'union des bases des $E(\lambda_i)$, on obtient une base de V formée de vecteurs propres. Donc f est diagonalisable. Inversement, si f est diagonalisable, on choisit une base de V formée de vecteurs propres: $\{e_1, \dots, e_n\}$.

Or chaque e_i appartient à l'un des $E(\lambda_j)$. Donc

$$V = E(\lambda_1) + \dots + E(\lambda_r). \quad \#$$

7.26 Remarque: Si f admet λ comme valeur propre, alors $C_f(\lambda) = 0$, et on sait que $(x-\lambda)$ divise $C_f(x)$.

Par 7.22, si f admet n valeurs propres distinctes $\lambda_1, \dots, \lambda_n$, alors on en déduit $C_f(x) = (-1)^n (x-\lambda_1) \cdot \dots \cdot (x-\lambda_n)$.

Si chacun des $E(\lambda_i)$ est de dimension 1, $V = E(\lambda_1) \oplus \dots \oplus E(\lambda_n)$ et f est diagonalisable. Dans l'exemple 7.23, on a:

$C_g(x) = (x-1)^2$	$C_g(x) = (x-1)^2$	
dim $E(1) = 2$	dim $E(1) = 1$	Ceci motive:
$\Rightarrow f$ diagonalisable	g non-diagonalisable	

7.27 Définition: Soit λ une valeur propre de f . On appelle

(a) multiplicité algébrique de λ le nombre

$$m_\lambda(\lambda) = \max \{ m \mid (x-\lambda)^m \text{ divise } C_f(x) \} \in \mathbb{N}$$

(b) multiplicité géométrique de λ le nombre
 $m_g(\lambda) = \dim E(\lambda) \in \mathbb{N}$.

7.28 Proposition : Si λ est valeur propre de f , alors

$$1 \leq m_g(\lambda) \leq m_a(\lambda).$$

Preuve : Si λ est valeur propre, il existe un $x \in E(\lambda)$ avec $x \neq 0$, donc $m_g(\lambda) = \dim E(\lambda) \geq 1$.

Si $m_g(\lambda) = n = \dim V$, alors $E(\lambda) = V$, donc

$f: V \rightarrow V, x \mapsto \lambda x$. Alors $\{f\}_B^B = \text{diag}(\lambda, \dots, \lambda)$ dans n'importe quelle base B , et donc $C_f(x) = (\lambda - x)^n$, donc $m_a(\lambda) = n = m_g(\lambda)$.

Supposons $m_g(\lambda) = m < n$, et soit $\{x_1, \dots, x_m\}$ une base de $E(\lambda)$, que l'on complète en une base $B = \{x_1, \dots, x_m, x_{m+1}, \dots, x_n\}$ de V .

Alors $\{f\}_B^B = A = \begin{pmatrix} \lambda I_m & B \\ 0 & C \end{pmatrix}$ pour certaines notations

$$B \in M_{m, n-m}(\mathbb{K}), \quad C \in M_{n-m, n-m}(\mathbb{K}).$$

$$\text{Ainsi } A - xI_n = \begin{pmatrix} (\lambda - x)I_m & B \\ 0 & C - xI_{n-m} \end{pmatrix}.$$

En développant successivement le déterminant par rapport à la $i^{\text{ème}}$ colonne m fois (chacune de ces colonnes contient alors qu'un terme non-nul égal à $(\lambda - x)$), on trouve

$$C_f(x) = \det(A - xI_n) = (\lambda - x)^m \cdot \det(C - xI_{n-m}).$$

Donc $m_a(\lambda) \geq m = m_g(\lambda)$. #

7.29 Théorème : Soit $f: V \rightarrow V$, V de dim n . Alors les

conditions suivantes sont équivalentes: (a) f est diagonalisable.

(b) $C_f(x)$ admet toutes ses racines dans $\mathbb{K}$, donc

$$C_f(x) = (-1)^n (x - \lambda_1)^{m_a(\lambda_1)} \cdots (x - \lambda_k)^{m_a(\lambda_k)}, \quad \text{et de plus,}$$

$$m_g(\lambda_i) = m_a(\lambda_i) \quad \text{pour tous } 1 \leq i \leq k.$$

7.24 Lemme: Soient $\lambda_1, \dots, \lambda_r$ des valeurs propres distinctes de f . Si $x \in E(\lambda_1) \oplus \dots \oplus E(\lambda_r) \subset V$ est un vecteur propre, alors il existe $1 \leq i \leq r$ unique avec $x \in E(\lambda_i)$.

preuve: posons $x = x_1 + \dots + x_r$ avec $x_i \in E(\lambda_i)$, et supposons que $f(x) = \mu x$, $\mu \in \mathbb{R}$, et $x \neq 0$.

Alors $\mu x_1 + \dots + \mu x_r = \mu x = f(x) = f(x_1 + \dots + x_r) = f(x_1) + \dots + f(x_r) = \lambda_1 x_1 + \dots + \lambda_r x_r$, donc

$$(\mu - \lambda_1)x_1 + \dots + (\mu - \lambda_r)x_r = 0, \quad \stackrel{\oplus}{\Rightarrow} (\mu - \lambda_j)x_j = 0 \quad \forall 1 \leq j \leq r.$$

Comme $x \neq 0$, il existe un $1 \leq i \leq r$ avec $x_i \neq 0$,

donc $(\mu - \lambda_i)x_i = 0 \Rightarrow \mu = \lambda_i$. Un tel i est donc unique,

et $(\mu - \lambda_j) \neq 0$ si $j \neq i \Rightarrow x_j = 0$ si $j \neq i \Rightarrow$

$$x = x_i \in E(\lambda_i). \quad \#$$

preuve de 7.29. (a) $\Rightarrow$ (b). Supposons donc f diagonalisable, et soient $\lambda_1, \dots, \lambda_r$ les valeurs propres de f . On a donc par 7.25

$$V = E(\lambda_1) \oplus \dots \oplus E(\lambda_r), \text{ donc } n = m_g(\lambda_1) + \dots + m_g(\lambda_r).$$

D'autre part, il existe une base B de V avec $\{f\}_B^B =$

$$\text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{m_g(\lambda_1)}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{m_g(\lambda_r)}). \text{ On en déduit}$$

$$C_f(X) = (-1)^n (x - \lambda_1)^{m_g(\lambda_1)} \dots (x - \lambda_r)^{m_g(\lambda_r)}, \text{ donc}$$

$$m_g(\lambda_i) = m_e(\lambda_i).$$

$$(b) \Rightarrow (a) \text{ Supposons } C_f(X) = (-1)^n (x - \lambda_1)^{m_1} \dots (x - \lambda_r)^{m_r},$$

et de plus $m_g(\lambda_i) = m_i$. On a par 7.20:

$$E(\lambda_1) \oplus \dots \oplus E(\lambda_r) \subset V. \text{ Or } \dim(E(\lambda_1) \oplus \dots \oplus E(\lambda_r)) =$$

$$\dim(E(\lambda_1)) + \dots + \dim(E(\lambda_r)) = m_g(\lambda_1) + \dots + m_g(\lambda_r) = m_1 + \dots + m_r = n.$$

Donc $E(\lambda_1) \oplus \dots \oplus E(\lambda_r) = V$, et f est diagonalisable par 7.25. $\#$

7.30 Remarque: Rappelons que toutes les notions ci-dessus s'appliquent aussi à une matrice $A \in M_n(\mathbb{K})$ au lieu de $f: V \rightarrow V$:

On prend $V = \mathbb{R}^n$ avec la base canonique B , et on utilise

l'isomorphisme $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \rightarrow M_n(\mathbb{K})$, $f \mapsto \{f\}_B^B$.

7.31. Exemple: Soit $A \in M_n(\mathbb{K})$ et supposons A diagonalisable: $\exists P \in GL_n(\mathbb{K})$ et $D = \text{diag}(a_1, \dots, a_n)$ avec $A = P^{-1} D P$. Alors on démontre facilement par récurrence sur m que $A^m = P^{-1} D^m P = P^{-1} \text{diag}(a_1^m, \dots, a_n^m) P$.

7.32 Exemple (suites récurrentes linéaires). Supposons

une suite (u_n) de $\mathbb{K}$ définie par u_0, u_1 donné

et $u_{n+2} = c u_n + d u_{n+1}$, $c, d \in \mathbb{K}$. Alors

$$\begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix} \begin{pmatrix} u_{n-1} \\ u_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}^2 \begin{pmatrix} u_{n-2} \\ u_{n-1} \end{pmatrix} = \dots = \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}^n \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

Si $\begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}$ est diagonalisable, alors on peut calculer $\begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}^n$ et donner une valeur explicite à u_n pour tout n .

Un exemple célèbre: la suite de Fibonacci, donnée par $u_0 = u_1 = 1$, $u_{n+2} = u_n + u_{n+1}$.

Donc, si $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in M_2(\mathbb{R})$, on a

$$\begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Essayons de calculer $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n$ par la méthode 7.31.

$$\begin{aligned} C_A(x) &= |A - x I_2| = \begin{vmatrix} -x & 1 \\ 1 & 1-x \end{vmatrix} = x^2 - x - 1 = \\ &= (x - 2)(x - \mu) \quad \text{avec} \quad \lambda = \frac{1+\sqrt{5}}{2}, \quad \mu = \frac{1-\sqrt{5}}{2} \end{aligned}$$

Remarquons: $\bullet \lambda + \mu = 1 \quad \bullet 1 + \lambda = \lambda^2$
 $\bullet \lambda - \mu = \sqrt{5} \quad \bullet 1 + \mu = \mu^2 \quad \bullet \lambda \mu = -1$

A a deux valeurs propres distinctes, donc est diagonalisable!

Cherchons des vecteurs propres: posons $v_1 = (1, \lambda)$ et $v_2 = (1, \mu)$.

Alors v_1 est vecteur propre pour λ , et v_2 pour μ . En effet $A \begin{pmatrix} 1 \\ \lambda \end{pmatrix} =$

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \end{pmatrix} = \begin{pmatrix} \lambda \\ 1 + \lambda \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda^2 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \lambda \end{pmatrix}, \quad \text{et de même} \quad A \begin{pmatrix} 1 \\ \mu \end{pmatrix} = \mu \begin{pmatrix} 1 \\ \mu \end{pmatrix}.$$

On a donc la base $B = \{v_1, v_2\}$ de $\mathbb{R}^2$ formée de vecteurs propres. Si C est la base canonique, posons

$$P = \{id\}_B^C = \begin{pmatrix} 1 & 1 \\ \lambda & \mu \end{pmatrix}$$

Donc $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = P \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} P^{-1}$

(Rappel : si $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = (y, x+y)$,
alors $A = \{f\}_C^C = \{id\}_B^C \{f\}_B^B \{id\}_C^B = P \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} P^{-1}$.)

On calcule $P^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\mu & 1 \\ \lambda & -1 \end{pmatrix}$, donc

$$\begin{aligned} \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \lambda & \mu \end{pmatrix} \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} \begin{pmatrix} -\mu & 1 \\ \lambda & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda^n & \mu^n \\ \lambda^{n+1} & \mu^{n+1} \end{pmatrix} \begin{pmatrix} 1-\mu \\ \lambda-1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda^n & \mu^n \\ \lambda^{n+1} & \mu^{n+1} \end{pmatrix} \begin{pmatrix} \lambda \\ -\mu \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda^{n+1} - \mu^{n+1} \\ \lambda^{n+2} - \mu^{n+2} \end{pmatrix}, \text{ donc} \end{aligned}$$

$$u_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right).$$

7.33 Définition : Une matrice $T \in M_n(\mathbb{K})$ est dite triangulaire supérieure si $T = (t_{ij})$ et $t_{ij} = 0$ si $i > j$:

$$T = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ & \ddots & \\ 0 & & t_{nn} \end{pmatrix}.$$

Une matrice $A \in M_n(\mathbb{K})$ est dite triagonalisable s'il existe une matrice triangulaire supérieure T avec $A \approx T$.

Si V est de dimension finie, on dit qu'un endomorphisme $f: V \rightarrow V$ est triagonalisable si $\{f\}_B^B$ est triangulaire, où B est une base quelconque (ou, de façon équivalente, s'il existe une base C de V avec $\{f\}_C^C$ triangulaire supérieure).

7.34 Théorème: Un endomorphisme $f: V \rightarrow V$ (où V est un $\mathbb{K}$ esp. vectoriel de dim. finie) est triséparable si et seulement si $C_f(x)$ a toutes ses racines dans $\mathbb{K}$. (100)

De même, $A \in M_n(\mathbb{K})$ est triséparable ssi $C_A(x)$ a toutes ses racines dans $\mathbb{K}$.

Preuve: Si f est triséparable, alors $\exists$ une base B avec $\{f\}_B^B = T = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ 0 & \ddots & 0 \\ \vdots & & t_{nn} \end{pmatrix}$, et $C_T(x) = (t_{11}-x) \dots (t_{nn}-x)$

donc $C_f(x)$ a toutes ses racines dans $\mathbb{K}$. Inversement, si f a toutes ses racines dans $\mathbb{K}$, on diminue par récurrence sur $n = \dim V$ que f est triséparable.

Si $n=1$: clair!

Supposons $n > 1$ et le théorème diminue si $\dim V < n$.

Par hypothèse, f a au moins une valeur propre λ ;

soit $v \in V$ un vecteur propre pour λ , $v \neq 0$.

On choisit une base $B = \{v, e_2, \dots, e_n\}$ de V .

Alors $\{f\}_B^B = \begin{pmatrix} \lambda & * & * \\ 0 & & \\ \vdots & & n \end{pmatrix} = A$, $n \in M_{n-1}(\mathbb{K})$

En développant par la première ligne, on trouve

$$C_f(x) = C_A(x) = (\lambda - x) \cdot \det(n - xI_{n-1}),$$

et $\det(n - xI_{n-1})$ a aussi toutes ses racines dans $\mathbb{K}$.

Posons $W = \text{Vect}(e_2, \dots, e_n) \subset V$ (un sous-esp. de dim. $n-1$)

et considérons $P: V \rightarrow V$ la projection sur W donnée

par $P(a_1 v + a_2 e_2 + \dots + a_n e_n) = a_2 e_2 + \dots + a_n e_n$.

Alors $g: W \rightarrow W$, $x \mapsto P \circ f(x)$ est un endomorphisme

Si $C = \{e_2, \dots, e_n\}$ (base de W), alors on obtient

$\{g\}_C^C$. Il existe donc, par hyp. de réc., une base

$D = \{f_2, \dots, f_n\}$ de W avec $\{g\}_D^D = T = \begin{pmatrix} t_{22} & * \\ 0 & \ddots & t_{nn} \end{pmatrix}$.

Dans la base $\mathcal{E} = \{v_1, p_1, \dots, p_n\}$ de V , on a alors (101)

$$\{f\}_{\mathcal{E}}^{\mathcal{E}} = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & T \end{array} \right) \text{ Triangulaire supérieure. } \#$$

7.35 Théorème (Classification des $A \in M_2(\mathbb{C})$). Soit $A \in M_2(\mathbb{C})$, et $C_A(x) = (x-\lambda)(x-\mu)$. On a exactement une des options suivantes :

(a) A est diagonalisable (donc $A \approx \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$)

(b) A n'est pas diagonalisable : alors $\lambda = \mu$ et $A \approx \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$

preuve : Supposons que A ne soit pas diagonalisable : Alors 7.22

implique $\lambda = \mu$. Par 7.34, nous savons que A est

triagonalisable, donc est semblable à $\begin{pmatrix} \lambda & a \\ 0 & b \end{pmatrix}$

pour certains $a, b \in \mathbb{C}$. On a que b est une racine

de $C_A(x)$, donc $b = \lambda$: Ainsi, A est semblable à

$$\begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}, \text{ et on calcule } \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

On en déduit $A \approx \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$

(Remarque : on vérifie facilement que $\approx$ est une relation d'équivalence ; en particulier, $A \approx B$ et $B \approx C \Rightarrow A \approx C$). #

7.36 Théorème (Classification des $A \in M_2(\mathbb{R})$)

Supposons $A \in M_2(\mathbb{R})$. On a exactement une des options suivantes.

(a) A est diagonalisable

(b) A n'est pas diagonalisable, mais $C_A(x)$ admet une racine réelle λ . Alors $A \approx \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

(c) $C_A(x)$ n'admet aucune racine réelle.

Soit $a+bi \in \mathbb{C} \setminus \mathbb{R}$ une racine complexe de $C_A(x)$.

$$\text{Alors } A \approx \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad (\text{avec } b \neq 0)$$

preuve : Supposons A non-diagonalisable. Si $C_A(x)$ admet une racine réelle λ , alors on a $C_A(x) = (x-\lambda)^2$ (si $C_A(x)$ a deux racines distinctes, alors A est diagonalisable) et on procède exactement comme pour le cas 7.35(b).

Supposons donc $C_A(x)$ irréductible dans $\mathbb{R}$.

Soit $\lambda = a+bi$ une racine de $C_A(x)$ dans $\mathbb{C}$, et $v = (x_1+iy_1, x_2+iy_2) \in \mathbb{C}^2$ un vecteur propre de f (avec $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$, $\{p\}_B^B = A$ pour $B \subset \mathbb{C}^2$ base canon).

Soit $x = \text{Re}(v) = (x_1, x_2) \in \mathbb{R}^2$, $y = \text{Im}(v) = (y_1, y_2) \in \mathbb{R}^2$, et $X = \{x\}^{\mathbb{C}}$, $Y = \{y\}^{\mathbb{C}}$ ($\mathbb{C}$ base canon de $\mathbb{R}^2$), $V = \{v\}^{\mathbb{B}}$. On a $AV = \lambda V$, et $A\bar{v} = \overline{AV} = \overline{\lambda v} = \bar{\lambda} \bar{v}$.

On en déduit que $\bar{\lambda}$ est aussi une valeur propre de A , donc une racine de $C_A(x)$, et $\bar{v}$ est un vecteur propre pour $\bar{\lambda}$. On a $\lambda \neq \bar{\lambda}$ (sinon $\lambda \in \mathbb{R}$!), et donc $\{v, \bar{v}\}$ est une base de $\mathbb{C}^2$.

On calcule :

$$AX = \text{Re}(AV) = \text{Re}(\lambda v) = \text{Re}(ax + ay + biX - by) = ax - by$$

$$AY = \text{Im}(AV) = \text{Im}(\lambda v) = ay + bX, \text{ donc si } P = (X \ Y) = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}, \text{ alors}$$

$$AP = (AX \ AY) = (ax - by \ ay + bX) = (X \ Y) \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = P \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Or P est inversible : en effet, ses colonnes sont $X = \{x\}^{\mathbb{C}}$ et $Y = \{y\}^{\mathbb{C}}$, donc il faut vérifier que $\{x, y\} \subset \mathbb{R}^2$ est liée. Si $\alpha, \beta \in \mathbb{R}$ avec $\alpha x + \beta y = 0$, alors

$$\frac{\alpha}{2}(v + \bar{v}) + \frac{\beta}{2}(v - \bar{v}) = 0, \text{ donc } (\alpha + \beta)v + (\alpha - \beta)\bar{v} = 0$$

donc $\begin{cases} \alpha + \beta = 0 \\ \alpha - \beta = 0 \end{cases}$ car $\{v, \bar{v}\}$ est une base de $\mathbb{C}^2$, et donc $\alpha = \beta = 0$.

$$\Rightarrow A = P \begin{pmatrix} a & b \\ -b & a \end{pmatrix} P^{-1} \text{ et } A \sim \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \quad \#$$

7.37 Remarque : dans le Théorème 7.36, il est clair

que $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \not\sim \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$: cela se voit de plusieurs façons :

par exemple : $P \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} P^{-1} = P \cdot 2 \text{Id}_2 P^{-1} = 2 P P^{-1} = 2 \text{Id}_2$,

donc $A \sim \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ si et seulement si $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

En dimension supérieure, pour étudier la question : $A \sim B$?

on introduit un outil très utile, le polynôme minimal.

En 7.15, nous avons défini $P(A)$, si $A \in M_n(K)$ et $P(x) \in P^K[x]$, ainsi que $P(f) \in \mathcal{L}(V, V)$ si $f \in \mathcal{L}(V, V)$.

7.38 Lemme : Soit $A \in M_n(K)$ et $Q \in GL_n(K)$ et

$P(x) \in P^K[x]$. Alors $P(QAQ^{-1}) = QP(A)Q^{-1}$.

En particulier, si $A \sim B$ alors $P(A) \sim P(B)$.

preuve : Soit de $(QAQ^{-1})^m = QA^mQ^{-1}$ et

$$Q(2M + \mu N)Q^{-1} = 2QMQ^{-1} + \mu QNQ^{-1}, \text{ pour tout } m \in \mathbb{N}, 2, \mu \in K, M, N \in M_n(K).$$

7.39 Lemme : Soit $A \in M_n(K)$. Il existe un polynôme $Q(x) \in$

$P^K[x]$ avec : $Q \neq 0$ et $Q(A) = 0$.

preuve : On considère $\{I_n, A, \dots, A^{n^2}\} \subset M_n(K)$;

cet ensemble a $n^2 + 1$ éléments, et $\dim M_n(K) = n^2$, donc

il est lié : $\exists a_0, a_1, \dots, a_{n^2} \in K$ avec

$$a_0 I_n + a_1 A + \dots + a_{n^2} A^{n^2} = 0$$

donc $Q(A) = 0$ si $Q(x) = a_0 + a_1 x + \dots + a_{n^2} x^{n^2}$. $\#$

7.40 Théorème + Définition : Soit $A \in M_n(K)$. Il existe un

unique polynôme unitaire $m_A(x) \in P^K[x]$ vérifiant

(i) $m_A(A) = 0$

(ii) Si $Q(x) \in P^K[x]$ et $Q(A) = 0$, alors

$m_A(x)$ divise $Q(x)$. En particulier, $m_A(x)$ est le polynôme unitaire de degré minimal avec $m_A(A) = 0$

On appelle $m_A(x)$ le polynôme minimal de A .

(104)

Preuve : On pose $\text{Ann}(A) = \{Q(x) \in P^{\mathbb{K}}[x] \mid Q(A) = 0\}$.

Par 7.39, $\text{Ann}(A) \neq \{0\}$. On choisit un polynôme $m'_A(x) \in \text{Ann}(A)$ non nul et de degré minimal.

$m'_A(x) = a_0 + \dots + a_k x^k$ avec $a_k \neq 0$. On pose $m_A(x) = \frac{1}{a_k} m'_A(x)$, et on a : $m_A(x)$ est unitaire, et $m_A(A) = 0$. En utilisant la division euclidienne sur $P^{\mathbb{K}}(x)$, on peut montrer : $Q(x) \in \text{Ann}(A) \iff m_A(x) \text{ divise } Q(x)$. #

7.41 Remarque : Soit $f: V \rightarrow V$ comme plus haut, A et B deux bases de V et $A = \{p\}_A^A$, $B = \{p\}_B^B$.

On a alors $m_A(x) = m_B(x)$: en effet, on sait qu'il existe $P \in GL_n(\mathbb{K})$, $B = P A P^{-1}$. Par 7.38 on a donc $m_A(B) = m_A(P A P^{-1}) = P m_A(A) P^{-1} = P 0 P^{-1} = 0$ donc $m_B(x) \mid m_A(x)$. De même,

$m_B(A) = m_B(P^{-1} B P) = P^{-1} m_B(B) P = 0$, donc $m_A(x) \mid m_B(x)$. Puisque tous deux sont unitaires, on en déduit $m_A(x) = m_B(x)$. Donc la définition suivante ne dépend pas du choix de la base B :

7.4.2 Définition : On appelle polynôme minimal de f le polynôme $m_f(x) = m_{\{f\}_B^B}(x)$, où B est une base quelconque de V .

les deux lemmes suivants montrent que l'on peut assez facilement calculer le polynôme minimal si l'on connaît le polynôme caractéristique.

7.43 Théorème (Cayley-Hamilton) : On a $C_f(f) = 0$, donc $m_f(x)$ divise $C_f(x)$. Ainsi : $m_A(x)$ divise $C_A(x)$, $\forall A \in M_n(\mathbb{K})$.

preuve: nous omettons la preuve du Théorème. #

7.44 Théorème (Frobenius): $m_f(x)$ et $c_f(x)$ ont les mêmes facteurs irréductibles (de même pour $m_A(x)$ et $c_A(x)$).

preuve: admis. #

7.45. Exemples: (a) Soit $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ $2 \neq 0$.

Alors $c_A(x) = -(x-2)^3$, et

$$m_A(x) = x-2 : m_A(A) = A - 2I_3 = 0.$$

(b) Reprenons notre vel exemple 7.3, $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avec

$$g(x,y) = (5x+3y, -6x-4y), \quad B \text{ base cano. Alors}$$

$$A = \{g\}_B^B = \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix}.$$

Nous avons calculé $c_g(x) = x^2 - x - 2 = (x+1)(x-2)$.

D'après 7.43 et 7.44 on a $c_g(x) = m_g(x)$.

Vérifions :

$$m_A(A) = A^2 - A - 2I_2 = \begin{pmatrix} 7 & 3 \\ -6 & -2 \end{pmatrix} - \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

(c) $B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ $c_B(x) = -(x-2)^3$. On sait par 7.43+44

que $m_B(x) \in \{ (x-2), (x-2)^2, (x-2)^3 \}$.

On calcule :

$$B - 2I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq 0, \quad (B - 2I_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow m_B(x) = (x-2)^3. \quad \text{Si } C = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad m_C(x) = (x-2)^2.$$

(d) Calculer le polynôme minimal de $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$:

On calcule d'abord le polynôme caractéristique :

$$c_A(x) = -(x-2)^3. \quad \text{Donc } m_A(x) \in \{ (x-2), (x-2)^2, (x-2)^3 \}.$$

$$\text{on a } (A-2I) = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0, \quad (A-2I)^2 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 = 0,$$

donc $m_A(x) = (x-2)^2$.

7.45 Definition: Un sous-espace $W \subset V$ est dit stable par $f: V \rightarrow V$ si $f(W) \subset W$.

7.46 Exemples: (a) 0 et V sont stables par $f: V \rightarrow V$.

(b) Soit $x \in V$, $x \neq 0$, et $W = \text{Vect}\{x\}$. Alors W est stable par f si et seulement si x est un vecteur propre de f :
 $f(W) \subset W \Leftrightarrow f(x) \in \text{Vect}\{x\} \Leftrightarrow \exists \lambda, f(x) = \lambda x$.

(c) Si $W \subset V$ est stable par f , et si $P \in P^k[x]$ est un polynôme, alors W est stable par $P(f)$:
c'est clair, car si $x \in W$, $f(x) \in W$, donc $f^2(x) = f(f(x)) \in W$, et par récurrence, $f^k(x) \in W \forall k \geq 0$.
Finalement, $P(f)(x)$ est comb. linéaire des $f^k(x)$, tous dans W , donc $P(f)(x) \in W$.

7.47 Proposition: Si g est un autre endomorphisme de V avec $f \circ g = g \circ f$, alors $\text{Ker}(g)$ et $\text{Im}(g)$ sont stables par f .

preuve: Soit $x \in \text{Ker}(g)$; alors $g(f(x)) = f(g(x)) = f(0) = 0$, donc $f(x) \in \text{Ker}(g)$. Donc $f(\text{Ker}(g)) \subset \text{Ker}(g)$.

De même, si $x \in \text{Im}(g)$, soit $y \in V$ avec $x = g(y)$.

Alors $f(x) = f(g(y)) = g \circ f(y) \in \text{Im}(g)$, donc $f(\text{Im}(g)) \subset \text{Im}(g)$. #

7.48 Corollaire: Si $P(x) \in P^k[x]$, alors $\text{Ker } P(f)$ est stable par f .

preuve: Nous avons $P(f) \circ f = f \circ P(f)$, donc cela suit de 7.47.

7.49 Théorème: Soient $P(x), Q(x) \in P^k[x]$ premiers entre eux. Alors $\text{Ker } P(f) \oplus \text{Ker } Q(f) = \text{Ker } (P(f) \circ Q(f))$

preuve : Par le lemme de Bézout, il existe $R(x)$ et $S(x) \in P^k[x]$ avec $R(x)P(x) + S(x)Q(x) = 1$, donc $R(f) \circ P(f) + S(f) \circ Q(f) = \text{id}_V$. (*)

(i) La somme est directe : Supposons $x \in \text{Ker } P(f) \cap \text{Ker } Q(f)$.
alors en évaluant (*) en x on trouve

$$x = R(f)(P(f)(x)) + S(f)(Q(f)(x)) = R(f)(0) + S(f)(0) = 0.$$

(ii) $\text{Ker } P(f) + \text{Ker } Q(f) \subset \text{Ker } (P(f) \circ Q(f))$:

clair, car si $y \in \text{Ker } Q(f)$, alors $y \in \text{Ker } (P(f) \circ Q(f))$.

De même, $P(f) \circ Q(f) = P(x) \cdot Q(x)(f) = Q(f) \circ P(f)$,

donc si $x \in \text{Ker } P(f)$, alors $x \in \text{Ker } (Q(f) \circ P(f)) = \text{Ker } (P(f) \circ Q(f))$.

(iii) $\text{Ker } (P(f) \circ Q(f)) \subset \text{Ker } P(f) + \text{Ker } Q(f)$:

Soit $x \in \text{Ker } (P(f) \circ Q(f))$; d'après (*) on a

$$x = \underbrace{R(f) \circ P(f)(x)}_y + \underbrace{S(f) \circ Q(f)(x)}_z$$

$$\text{Or } Q(f)(y) = (Q \cdot R \cdot P)(f)(x) = (RPQ)(f)(x) =$$

$$R(f)P(f)Q(f)(x) = P(f)(0) = 0, \text{ donc } y \in \text{Ker } (Q(f)).$$

Similairement, $z \in \text{Ker } (P(f))$. Donc $x \in \text{Ker } P(f) + \text{Ker } Q(f)$. #

7.50 Corollaire : Soient $P_1(x), \dots, P_r(x) \in P^k[x]$ deux-à-deux premiers entre eux. Alors

$$\text{Ker } P_1(f) \oplus \dots \oplus \text{Ker } P_r(f) = \text{Ker } (P_1 \dots P_r(f)).$$

preuve : par récurrence sur r , en utilisant 7.49 : $P_r(x)$ et $P_1 \dots P_{r-1}(x)$ sont premiers entre eux, donc

$$\text{Ker } (P_1 \dots P_{r-1}(f)) \oplus \text{Ker } (P_r(f)) = \text{Ker } (P_1 \dots P_r(f)). \quad \#$$

7.51 Corollaire : Soit $m_f(x) = P_1^{n_1} \dots P_r^{n_r}(x)$ avec $P_1, \dots, P_r$ deux-à-deux premiers et irréductibles

$$\text{Ker}(P_1(f)^{n_1}) \oplus \dots \oplus \text{Ker}(P_r(f)^{n_r}) = V$$

(108)

En particulier, si $m_f(x)$ a toutes ses racines dans $\mathbb{K}$,

$$m_f(x) = (x - \lambda_1)^{n_1} \dots (x - \lambda_r)^{n_r}, \quad \text{alors}$$

$$\text{Ker}(f - \lambda_1 I_V)^{n_1} \oplus \dots \oplus \text{Ker}(f - \lambda_r I_V)^{n_r} = V,$$

(avec chacun des $\text{Ker}(f - \lambda_i I_V)^{n_i}$ invariant par f).

C'est la disposition primaire de f .

preuve : Soit de 7.50 : Comme $m_f(f) = 0$,

$$\text{Ker}(P_1(f)^{n_1}) \oplus \dots \oplus \text{Ker}(P_r(f)^{n_r}) = \text{Ker}(m_f(f)) = V. \quad \#$$

7.52 Définition : Soit $\lambda \in \mathbb{K}$, $m \in \mathbb{N}$, $m \geq 1$. On appelle

bloc de Jordan de taille m pour λ la matrice

$$J(\lambda, m) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in M_m(\mathbb{K})$$

Nous allons montrer que si $A \in M_n(\mathbb{K})$ est diagonalisable,

alors

$$A \approx \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & J_k \end{pmatrix} \quad \text{où } J_k \text{ sont des blocs de Jordan.}$$

7.53 Définition : Soit $f: V \rightarrow V$. Alors f est dit nilpotent s'il existe $m > 0$ avec $f^m = 0$.

7.54 Lemme : Soit $f: V \rightarrow V$ nilpotent, $x \in V$ et $m \in \mathbb{N}$, $m \geq 1$, avec $f^{m-1}(x) \neq 0$ et $f^m(x) = 0$. Alors

(a) $B = \{f^{m-1}(x), \dots, f(x), x\}$ est une base.

(b) Si $W = \text{Vect } B \subset V$, alors W est stable par f et $\{f|_W\}_B^B = J(0, m)$, où $f|_W: W \rightarrow W$ est la restriction de f à W .

preuve : (a) Supposons $y = a_1 f^{m-1}(x) + \dots + a_{m-1} f(x) + a_m x = 0$

$$\text{Alors } 0 = f^{m-1}(y) = f^{m-1}(0) = a_m f^{m-1}(x) \Rightarrow a_m = 0.$$

$$\Rightarrow 0 = f^{m-2}(y) = f^{m-2}(0) = a_{m-1} f^{m-1}(x) \Rightarrow a_{m-1} = 0 \text{ etc.}$$

(b) est clair : $f(f^i(x)) = f^{i+1}(x)$. #

7.55 Notation: dans 7.54, on note $Z(x, f)$

le sous-espace $W = \text{Vect} \{ f^{m-1}(x), \dots, f(x), x \}$.

On l'appelle le sous-espace cyclique de V engendré par x : c'est le plus petit sous-espace invariant contenant x .

7.56 Théorème (Décomposition cyclique): Supposons $g: V \rightarrow V$ nilpotent. Alors il existe $x_1, \dots, x_r \in V$ tel que

$$V = Z(x_1, g) \oplus \dots \oplus Z(x_r, g).$$

preuve: Recurrence sur $n = \dim V$.

Si $\dim V = 0$: c'est clair.

Supposons $\dim V \geq 1$ et le théorème prouvé dans les cas de dimension inférieure.

Comme g est nilpotent, g n'est pas injectif (g injectif $\Rightarrow g^m$ injectif $\forall m \geq 1$). Donc g n'est pas surjectif.

Donc $\dim(\text{Im}(g)) < n$. De plus, $\text{Im}(g)$ est stable par g . Par hypothèse de récurrence, il existe

$y_1, \dots, y_r \in \text{Im}(g)$ avec

$$\text{Im}(g) = Z(y_1, g) \oplus \dots \oplus Z(y_r, g).$$

Comme $y_1, \dots, y_r$ sont dans $\text{Im}(g)$, on peut choisir $x_1, \dots, x_r \in V$ avec $y_i = g(x_i)$, $i=1, \dots, r$.

Posons $W = Z(x_1, g) + \dots + Z(x_r, g)$. La somme est directe! : si $z_1 + \dots + z_r = 0$, alors

$$0 = g(0) = g(z_1) + \dots + g(z_r) \in Z(y_1, g) \oplus \dots \oplus Z(y_r, g)$$

donc $g(z_1) = \dots = g(z_r) = 0$ (cette somme est directe),

car $g(z_i) \in Z(y_i, g)$.

Donc $z_i \in Z(x_i, g) \cap \text{Ker}(g) \subset Z(y_i, g)$, et

donc $z_1 = \dots = z_r = 0$.

Ainsi $\dim W = \sum_{i=1}^r \dim Z(x_i, g) = \sum_{i=1}^r (\dim Z(y_i, g) + 1)$
 $= r + \text{rang}(g)$.

$$\text{D'autre part } \dim (Ker(g) \cap W) = \dim \left(\bigoplus_{i=1}^r Ker(g) \cap Z(x_i, g) \right) \quad (110)$$

$$= \sum_{i=1}^r \dim Ker(g) \cap Z(x_i, g) = \sum_{i=1}^r 1 = r.$$

Soit U un supplémentaire de $W \cap Ker(g)$ dans $Ker(g)$:

Alors $W + U = W \oplus U$: car $W \cap U = (W \cap Ker(g)) \cap U = 0$.

On calcule $\dim U = \dim Ker(g) - \dim (W \cap Ker(g)) = \dim V - \text{rang}(g) - r = \dim V - \dim W$. Donc

$V = W \oplus U$. On choisit une base $x_{r+1}, \dots, x_{r+s}$

de U . Comme $x_{r+i} \in U \subset Ker(g)$ on a

$$Z(x_{r+i}, g) = \text{Vect} \{x_{r+i}\}, \text{ donc}$$

$$U = Z(x_{r+1}, g) \oplus \dots \oplus Z(x_{r+s}, g), \text{ et finalement}$$

$$V = W \oplus U = Z(x_1, g) \oplus \dots \oplus Z(x_{r+s}, g). \quad \#$$

7.56 Corollaire: Soit $f: V \rightarrow V$ et supposons $m_f(x) = (x-\lambda)^m$,
 et donc $C_f(x) = (-1)^n (x-\lambda)^n$, $n = \dim V$. Soit $r = \dim E(\lambda)$.

Alors il existe $m_1, \dots, m_r \in \mathbb{N}$ avec $m_1 + \dots + m_r = n$,

$m = m_1 \geq m_2 \geq \dots \geq m_r \geq 1$ et une base B de V avec

$$\{f\}_B^B = \underbrace{J(\lambda, m_1) \oplus \dots \oplus J(\lambda, m_r)}_{B(\lambda)} := \begin{pmatrix} J(\lambda, m_1) & & 0 \\ & \ddots & \\ 0 & & J(\lambda, m_r) \end{pmatrix}.$$

preuve: Soit $g = f - \lambda \text{id}_V: V \rightarrow V$.

Comme $V = Ker(g^m)$, g est nilpotente. Donc

il existe $x_1, \dots, x_r \in V$ avec

$$V = Z(x_1, g) \oplus \dots \oplus Z(x_r, g).$$

On pose $m_i = \dim Z(x_i, g)$, et on suppose les x_i choisis dans un ordre tel que $m_1 \geq \dots \geq m_r \geq 1$.

On a alors $m = \max \{m_1, \dots, m_r\}$, donc $m = m_1$.

Alors $B = \{g^{m_1-1}(x_1), \dots, x_1, g^{m_2-1}(x_2), \dots, x_2, \dots, g^{m_r-1}(x_r), \dots, x_r\}$

est une base de V , et dans cette base on a

(111)

$$\{g\}_B^B = J(0, m_1) \oplus \dots \oplus J(0, m_r) = \begin{pmatrix} J(0, m_1) & & 0 \\ & \ddots & \\ 0 & & J(0, m_r) \end{pmatrix}.$$

Comme $f = g + \lambda \text{id}_V$, on en déduit

$$\{f\}_B^B = J(\lambda, m_1) \oplus \dots \oplus J(\lambda, m_r). \quad \text{Finalement, } \dim(E(\lambda)) = \dim \ker(f - \lambda \text{Id}_V) = \dim \ker(g) = n - \dim \text{Im}(g) = n - \text{rang}(g) = n - (n-r) = r : \text{ en effet, il y a dans } \{g\}_B^B \text{ (n-r) colonnes non-nulles, et elles sont linéairement indépendantes, donc } \text{rang}(g) = n-r. \quad \#$$

7.57 Théorème (Forme normale de Jordan): Soit $f: V \rightarrow V$

triagonalisable, et $\lambda_1, \dots, \lambda_k$ les valeurs propres de f .

Alors il existe une base B de V avec

$$\{f\}_B^B = B(\lambda_1) \oplus \dots \oplus B(\lambda_k),$$

où $B(\lambda_i) = J(\lambda_i, m_{i,1}) \oplus \dots \oplus J(\lambda_i, m_{i,r_i})$ comme en 7.56. Cette décomposition est unique à l'ordre des blocs près.

$$\{f\}_B^B = \left(\begin{array}{ccccccc} J(\lambda_1, m_{1,1}) & & & & & & \\ & \ddots & & & & & \\ & & J(\lambda_1, m_{1,r_1}) & & & & \\ & & & \ddots & & & \\ & & & & J(\lambda_k, m_{k,1}) & & \\ & & & & & \ddots & \\ & & & & & & J(\lambda_k, m_{k,r_k}) \end{array} \right)$$

$\left. \begin{array}{c} B(\lambda_1) \\ \vdots \\ B(\lambda_k) \end{array} \right\}$

Si $C_A(x) = (x - \lambda_1)^{n_1} \dots (x - \lambda_k)^{n_k}$ et

$m_A(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k}$, alors

- Le nombre de blocs $J(\lambda_i, ?)$ est égal à $r_i = \dim(E(\lambda_i))$
- Le plus gros bloc $J(\lambda_i, ?)$ est de taille m_i
- Le nombre de λ_i dans la diagonale est $n_i = m_{i,1} + \dots + m_{i,r_i}$.

7.58 Remarque: il est facile de vérifier: si $\dim V \leq 6$, alors $C_f(x)$, $m_f(x)$ et $\dim(E(\lambda))$ pour chaque valeur propre λ

(c'est-à-dire les infos $\lambda_1, \dots, \lambda_k$, $n_1, \dots, n_k$, $m_1, \dots, m_k$ et $r_1, \dots, r_k$) déterminent la forme normale de Jordan de f , à l'ordre des blocs près;

En dimension 3: supposons $f: V \rightarrow V$ linéaire, alors on a exactement l'une des options suivantes:

(a) f est diagonalisable.

(b) f ne l'est pas, et a 2 valeurs propres λ_1, λ_2 , avec $m_{\lambda_1}(\lambda_1) = 2$: $\exists B$, $\{f\}_B^B = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$

(ici $m_f(x) = -C_f(x) = -(x-\lambda_1)^2(x-\lambda_2)$)

(c) f n'est pas diagonalisable, et a une seule valeur propre λ .

alors on a B avec

$$\{f\}_B^B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{ou} \quad \{f\}_B^B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

$$(m_f(x) = (x-\lambda)^2)$$

$$(m_f(x) = (x-\lambda)^3)$$

En particulier, si $A, B \in M_3(K)$ sont diagonalisables, alors $m_A(x) = m_B(x)$

En dimension 4: ce n'est plus vrai:

$$\begin{matrix} \downarrow \\ A \approx B. \end{matrix}$$

$$\left(\begin{array}{c|c} \lambda & 1 \\ \hline \lambda & \lambda \\ \hline \lambda & 1 \\ \lambda & \lambda \end{array} \right) \quad \text{et} \quad \left(\begin{array}{c|c} \lambda & 1 \\ \lambda & \lambda \\ \hline \lambda & 1 \\ \lambda & \lambda \end{array} \right) \quad \text{ou même polynôme minimal } (x-\lambda)^2.$$

7.59 Théorème: f est diagonalisable si et seulement si $m_f(x)$ a toutes ses racines dans K , et si

$$m_f(x) = (x-\lambda_1) \cdots (x-\lambda_k), \quad \lambda_1, \dots, \lambda_k \text{ distincts.}$$

preuve: les blocs de Jordan ont alors tous la taille 1. $\#$