Une démonstration élémentaire et constructive de la forme normale canonique

par Patrick Teller

La forme canonique d'un endomorphisme peutêtre établie par des méthodes de la théorie des modules [1] ou en utilisant la dualiité [5]; on trouvera ici une démonstration élémentaire, uniquement matricielle, inspirée par [2], à cette différence que ce dernier texte donne l'impression recourir à des « trucs » alors que tout est en fait très naturel.

La représentation de sous-espaces invariants par un endomorphisme, avec la nature de l'endomorphisme induit, au moyen de triplets (A,T,$\tilde{A}$), qui prolonge la notion de vecteur et de valeur propres, m'a été suggérée par [3].

Dans tout ce qui suit K désigne un corps quelconque.

1 Sous-espace vectoriels stables par une matrice

Nous appellerons l'énoncé suivant Théorème, bien qu'il soit évident, parce qu'il sera continuellement utilisé

Théorème 1.

Soit une matrice $A \in \mathcal{M}_n(K)$ et une matrice $T \in \mathcal{M}_{n,r}(K)$, de rang r, dont on désignera les colonnes par $(T_1, ..., T_r)$, le sous-espace vectoriel vect(T_i,i=1,..,r) est stable par A si et seulement si il existe une matrice $\tilde{A} \in \mathcal{M}_r(K)$ telle que $AT=T\tilde{A}$.

Cette égalité signifie que l'image par A du vecteur T_j est égale à $\sum_{i=1}^{r} \tilde{A}_{ij} T_i$.

Dans la suite on désignera par T le sous-espace vectoriel dont une base est formée par les colonnes T_1,..,T_r et on désignera la matrice $\tilde{A}$ du nom de matrice induite.

2 Sur les matrices compagnons

Définition 2.

On appelle matrice compagnon du polynôme $X^k + \sum_{i=0}^{k-1} a_i X^i$ la matrice $C=\begin{pmatrix} 0 & ... & & -a0 \\ 1 & ... & & -a1 \\ 0 & 1 & & .... \\ . & 0 & . & ... \\ 0 & ... & ... & -a_{k-2} \\ & ... & 0\ 1 & -\mathrm{ak}-1 \end{pmatrix}$

(en fait il y a plusieurs types de matrices compagnons mais seul celui-ci sera utilisé ici); pour la suite on posera a_k=1 ce qui permettra d'écrire le polynôme sous la forme $\sum_{i=0}^{k-1} a_i X^i$; il est bien connu que $\sum_{i=0}^{k} a_i C^i$=0.

Théorème 3.

Une matrice M est semblable à une matrice compagnon si et seulement si son polynome caractéristique est son polynome minimal.

Démonstration. □

Définition 4. *Assemblage de Matrices Compagnons*

Soit deux matrices compagnons C et D on appellera assemblage de C et D toute matrice compagnon G tel que G est semblable à $\begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$. Lorsqu'une telle matrice existe elle est unique, on la notera C$\circ D$.

Proposition 5. *Un Théorème Chinois linéaire*

Soit deux matrices compagnons C et D, de polynomes minimaux $\pi_C(X)$ et $\pi_D(X)$, la matrice $\begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$ est semblable à une matrice compagnon si seulement si les polynomes minimaux sont premiers entre eux.

De plus si $\pi_C(X) \wedge \pi_D(X)=1$ et $\pi_C(X)U(X)+\pi_D(X)V(X)=1$ et si MT'=T'C et MT"=T"D, alors MT=T(C°D), où T=$\pi_C(M)U(M)0(_c$@T") $+\pi_D(M)V(M)$(T'@0_d), où 0_c@T" désigne la matrice (0_1 .... 0_c T_1'' ... . T_d'') *et T'@0_d désigne de même la matrice* (T_1' .... T_c' 0_1 ... . 0_d)

Démonstration.

Il est immédiat que le polynome minimal π_G de G=$\begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$ est $\pi_C \vee \pi_D$ tandis que le polynome caractéristique χ_G est égal à $\chi_C \times \chi_D$.

Si $\pi_C(X) \wedge \pi_D(X)=1$ alors il existe U(X) et V(X) tels que $\pi_C(X)U(X)+\pi_D(X)V(X)=1$ (on notera que U(X) est premier avec $\pi_D(X)$ et V(X) est premier avec $\pi_C(X)$); on en déduit que, pour tout vecteur Z, $\pi_C(M)U(M)Z+\pi_D(M)V(M)Z=Z$, d'où, si on pose $T_1=\pi_C(M)U(M)(0_c$@T") $+\pi_D(M)V(M)$(T'@0_d), et si on considère un polynôme P(X), $P(M)T_1=P(M)(\pi_C(M)U(M)$ (0_c@T") $+\pi_D(M)V(M)$(T'@0_d).

Supposons que $P(M)T_1=0$.

Si $P(X) \equiv Q(X)[\pi_C(X)]$ $P(M)T_1=Q(M)\pi_D(M)V(M)$(T'@0_d) qui est nul si et seulement si $\pi_C(X)$ divise $Q(X)\pi_D(X)V(M)$, ce qui entraine que $\pi_C(X)$ divise Q(X), d'où $P(X) \equiv 0[\pi_C(X)]$.

De même $P(X) \equiv 0[\pi_D(X)]$.

Donc Vect(T_1,MT_1, ..., M^{k-1}),où k=deg($\pi_C(X)\pi_D(X)$) est stable par M, et $\pi_C(M)\pi_D(M)T_1=\pi_C(M)\pi_D(M)(\pi_C(M)U(M)0_c$@T"$+\pi_D(M)V(M)$T'@$0_d$)$=0+0=0$. □

Proposition 6.

Soit $M \in \mathcal{M}_n(K)$ de polynome minimal $P(X)^r$, où P(X) est irréductible dans K[X] et de degré p, et $T_1 \in \mathrm{Ker}(P^r(M))$ $\backslash \mathrm{Ker}(P^{r-1}(M))$ alors, si on considère la matrice $T_{P(X)}=(T_1, \mathrm{MT}_1, M^2T_1, ..., M^{\mathrm{rp}-1}T_1)$, $MT_{P(X)}=T_{P(X)}C_{P(X)}$, où $C_{P(X)}$ est la matrice compagnon du polynôme $P(X)^r$.

Démonstration.

Par hypothèse $P^r(M)T_1=0$ et $P^{r-1}(M)T_1 \neq 0$, l'ensemble des polynômes Q(X) tels que $Q(M)(T_1)=0$ est un idéal (donc principal), qui contient $P^r(X)$ et donc, comme P(X) est irréductible, est engendré par une puissance de P(X), enfin comme $P^{r-1}(M)T_1 \neq 0$, il s'agit de $P^r(X)$.

Par suite $(T_1, \mathrm{MT}_1, M^2T_1, ..., M^{\mathrm{rp}-1}T_1)$ forme une famille libre et la restriction de l'action de M au sous-espace Vect($T_1, \mathrm{MT}_1, M^2T_1, ..., M^{\mathrm{rp}-1}T_1$) est la matrice compagnon du polynôme $P(X)^r$ □

D'où finalement:

Théorème 7.

Soit $A \in \mathcal{M}_n(K)$ de polynome minimal $\prod_{i=1}^{s} P_i(X)^{r_i}$, où les $P_i(X)$ sont irréductibles et premiers entre eux deux à deux alors il existe une matrice T telle que AT=TC, où C=$C_{P_1(X)} \circ .... \circ C_{P_s(X)}$

Théorème 8.

Soit $A \in \mathcal{M}_n(K)$, de polynôme minimal P(X), et C la matrice compagnon de P(X), alors il existe un sous-espace stable T, de dimension k, tel que la restriction de A à T est la matrice C, ce qui signifie que T est monogène, engendré par un vecteur x et ses images par A et que la restriction de A à T admet P(X) comme polynôme minimal.

Proposition 9.

Soit $C=\begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & . & -a_1 \\ 0 & 1 & & ... & ... \\ ... & 0 & & ... & ... \\ ... & ... & & 0 & ... \\ 0 & 0 & ... & 0 & -a_{k-1} \end{pmatrix}$, *alors quel que soit* $0<j<k$ *la dernière ligne de la matrice* $\sum_{i=j}^{k} a_i C^{i-j}$ *est* $Lj=(0,0...,1,0,...0)$ *où le 1 est en position j.*

Démonstration.

par récurrence sur k-j en initialisant à k-j=1 □

3 Les puissances de la matrice $M=\begin{pmatrix} A & U \\ 0 & B \end{pmatrix}$

Proposition 10.

Soit $M=\begin{pmatrix} A & U \\ 0 & B \end{pmatrix}$, *si on pose* $P(X)=\sum_{i=0}^{k} a_i X^i$, alors $P(M)=\begin{pmatrix} P(A) & P^*(A,U,B) \\ 0 & P(B) \end{pmatrix}$, *où* $P^*(A,U,B)=\sum_{j=0}^{k-1}(\sum_{i=j+1}^{k} a_i A^{i-1-j})\mathrm{UB}^j$.

Démonstration.

On montre par une simple récurrence que $\forall i \in N^*, M^i=\begin{pmatrix} A^i & \Gamma_i(A,U,B) \\ 0 & B^i \end{pmatrix}$, *où* $\Gamma_i(A,U,B)=\sum_{t=0}^{i-1} A^{i-1-t}\mathrm{UB}^t$; *puis* $P^*(A,U,B)=\sum_{i=1}^{k} a_i\Gamma_i(A,U,B)=P^*(A,U,B)=\sum_{i=1}^{k} a_i\Gamma_i(A,U,B)=\sum_{i=1}^{k} a_i(\sum_{j=0}^{i-1} A^{i-1-j}\mathrm{UB}^j)=\sum_{j=0}^{k-1}(\sum_{i=j+1}^{k} a_i A^{i-1-j})\mathrm{UB}^j$. □

Proposition 11.

Soit $MT_1=T_1C$, *où C est la matrice compagnon du polynôme P(X) et* T_2 *un supplémentaire de* T_1, *(U,V) telles que* $M\begin{pmatrix} T_1 & T_2 \end{pmatrix}=\begin{pmatrix} T_1 & T_2 \end{pmatrix}\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}$, *alors si on pose* $U=\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_n \end{pmatrix}$ *on a* $\sum_{j=0}^{k-1} U_j V^j=0$.

Démonstration.

Comme $\begin{pmatrix} T_1 & T_2 \end{pmatrix}$ *est de rang n M et* $\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}$ *sont semblables d'où* $P\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}=0$ *et* $\sum_{j=0}^{k-1}(\sum_{i=j+1}^{k} a_i C^{i-1-j})\mathrm{UV}^j=0$. *En particulier, comme la dernière ligne de la matrice* $(\sum_{i=j+1}^{k} a_i C^{i-1-j})$ *est* L_{j+1}, *si on pose* $U=\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_n \end{pmatrix}$ *la première ligne de la matrice* $\sum_{j=0}^{k-1}(\sum_{i=j+1}^{k} a_i C^{i-1-j})\mathrm{UV}^j$ *est égale à* $\sum_{j=0}^{k-1} U_j V^j$, *d'où* $\sum_{j=0}^{k-1} U_j V^j=0$. □

Théorème 12.

Avec les notations ci-desssus le sous-espace T_1 possède un supplémentaire stable par A.

Démonstration.

Le supplémentaire que nous recherchons pourra s'écrire sous la forme $T_1S_1+T_2S_2$, avec S_2 nécessairement inversible.

Nous voulons une égalité $A\begin{pmatrix} T_1 & T_2 \end{pmatrix}\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix}=\begin{pmatrix} T_1 & T_2 \end{pmatrix}\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & A'' \end{pmatrix}$ qui exprimera la stabilité du sous-espace engendré par les colonnes de $T_1S_1+T_2S_2$.

Comme $\begin{pmatrix} T_1 & T_2 \end{pmatrix}$ est inversible il est nécessaire et suffisant que $\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & A'' \end{pmatrix}=\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix}$, ce qui équivalent à $\begin{pmatrix} C & 0 \\ 0 & A'' \end{pmatrix}=\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix}^{-1}\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix} \iff \begin{pmatrix} C & 0 \\ 0 & A'' \end{pmatrix}=\begin{pmatrix} I & -S_1S_2^{-1} \\ 0 & {S_2}^{-1} \end{pmatrix}\begin{pmatrix} C & U \\ 0 & V \end{pmatrix}\begin{pmatrix} I & S_1 \\ 0 & S_2 \end{pmatrix} \iff \begin{pmatrix} C & 0 \\ 0 & A'' \end{pmatrix}=\begin{pmatrix} C & \mathrm{CS_1+US_2-S_1S_2^{-1}VS_2} \\ 0 & S_2^{-1}\mathrm{VS_2} \end{pmatrix} \iff \begin{cases} S_2^{-1}\mathrm{VS_2}=A'' \\ \mathrm{CS_1}S_2^{-1}+U-S_1S_2^{-1}V=0 \end{cases}$. Soit alors une matrice inversible arbitraire S_2 et $A''=S_2^{-1}\mathrm{VS_2}$, il est alors nécessaire et suffisant de déterminer X tel que CX-XV=-U, puis nous poserons S_1=$\mathrm{XS_2}$.

Posons comme plus haut $U=\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_k \end{pmatrix}$ et $X=\begin{pmatrix} X1 \\ X2 \\ .... \\ \mathrm{Xk-1} \\ 0 \end{pmatrix}$ (le choix de X_k=0 est cohérent parce que nous sommes dans le cas où l'équation de Sylvester n'a pas de solution unique, il sera validé si nous trouvons effectivement une solution).

$$\mathrm{CS_1}-S1V=-\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_k \end{pmatrix}\iff\begin{pmatrix} 0 & ... & & -a_0 \\ 1 & ... & & -a_1 \\ 0 & 1 & & .... \\ . & 0 & . & ... \\ 0 & ... & ... & -a_{k-2} \\ & & ... \ 0 \ 1 & -a_{k-1} \end{pmatrix}\begin{pmatrix} X1 \\ X2 \\ ... \\ ... \\ \mathrm{Xk-1} \\ 0 \end{pmatrix}-\begin{pmatrix} X1 \\ X2 \\ .... \\ \mathrm{Xk-1} \\ 0 \end{pmatrix}V=-\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_k \end{pmatrix}\iff\begin{pmatrix} -X_1V \\ X_1-X_2V \\ X_2-X_3V \\ .... \\ X_{k-1} \end{pmatrix}=$$

$$-\begin{pmatrix} U_1 \\ U_2 \\ .... \\ U_k \end{pmatrix}\iff\begin{cases} 0=U_1+U_2+U_3V+\cdots+U_kV^{k-1} \\ X_1=-U_2-U_3V-\cdots-U_kV^{k-1} \\ ..... \\ \mathrm{Xk-2=-Uk-1-UkV} \\ \mathrm{Xk-1=-Uk} \end{cases}$$

. Or nous avons vu plus haut que $0=U_1+U_2+U_3V+\cdots++U_kV^{k-1}$ donc le système est bien vérifié par $X=\begin{pmatrix} -(U_2+U_3V+\cdots++U_kV^{k-1}) \\ -(U_3+U_4V+\cdots++U_kV^{k-2}) \\ .... \\ -U_k \\ 0 \end{pmatrix}$. □

4 Le Théorème de Frobenius

Théorème 13.

Soit un endomorphisme f du K-espace vectoriel E de polynôme minimal π il existe une suite (F1,..,Fp) de sous-espace vectoriels stables par f tels que $E=\oplus_{i=1}^{p}F_i$, la restriction de f à chaque sous-espace F_i est cyclique de polynôme minimal P_i,$\forall i, P_{i+1}|P_i$; P_1 est le polynôme minimal de f .

De plus la suite des polynômes (P_i), appelés invariants de similitude de f, est entièrement déterminée par f et deux endomorphismes sont semblables si et seulement ils ont les mêmes invariants de similitude.

Démonstration.

L'existence de F1 est établie par le Théorème 4, la construction de la suite des F_i est assurée par le Théorème 8.

Dans notre construction le polynôme P_2 est déterminé par la matrice V, qui est induite (à une similitude près) par f sur F_2 et il en est de même pour les suivants.

La dernière affirmation découle de la décomposition en somme directe. □

Maintenant que c'est terminé je serai plus direct: la démonstration de ce résultat comme un cas de la Théorie des Modules est intéressante car elle le replace dans un cadre bien plus large mais elle n'aide en rien à la prise en mains de ce domaine de l'Algèbre, la démonstration habituellement présentée à notre époque, fondée sur l'usage de la dualité me semble aseptisée, elle est propre, bien faite, et « on ne se salit pas les mains »; quant à la démonstration de R.H. Hartwig, qui m'a donné le point de départ, je dois reconnaître qu'elle m'a toujours semblé artificielle de par les nombreuses formules justes mais « parachutées » qu'elle utilise.

Bibliographie:

[1] G. Birkhof, S. Mac Lane, Algèbre, les Grands Théorèmes, Gauthier-Villars, 1971.

[2] F.R. Gantmacher, Matrix Theory, Chelsea Publishing Company, 1977

[3] P. R. Halmos, Eigenvectors and adjoints, Linear Algebra Appl. 4:11-15 (1971).

[4] R.E. Hartwig, Roth's Removal Rule and the Rational Canonical Form,

[5] V.W. Prasolov, Problems and Theorems in Linear Algebra, Translation of Mathematical Monographs, American Mathematical Society, 1994

[6] P . Teller, l'Equation AM=MB, http://lalgebrisant.fr/images/pdfArticles/EquatAMMB.pdf

Paris, fin Janvier 2021