Autour de la notion de réduction d'une matrice

Tristan Vaccon

octobre 2013

Table des matières

Lien avec les leçons d'algèbre de l'agrégation

Ce qui suit a un lien profond avec la leçon numéro 119, "Exemples d'actions de groupes sur les espaces de matrices.", en particulier ce qui constituerait les premières parties de cette leçon.

Les liens avec la leçon 140, "Systèmes d'équations linéaires. Systèmes échelonnés. Résolution. Exemples et applications.", sont eux-aussi important.

On notera aussi quelques intersections, plus ou moins conséquentes, avec les leçons 106, "Groupe linéaire d'un espace vectoriel de dimension finie E , sous-groupes de GL(E). Applications.", 108, "Exemples de parties génératrices d'un groupe. Applications." et 120, "Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.".

1 Quelques mots sur les changements de base

Voir le livre d'Anette Paugam [1] si l'on a le moindre doute sur la manière d'écrire un changement de base, et ce qu'il signifie.

Une bonne manière de se rappeler ce que signifie $P = Mat_e(f)$ est de dissocier les espaces selon la base qu'on leur donne. En effet, faire le changement de base de E muni de la base e vers E muni de la base f se répercute sur l'endomorphisme f par ce qui suit :

$$(E, f) \leftarrow^{f'} (E, f) = (E, f) \leftarrow^{Id} (E, e) \leftarrow^{f} (E, e) \leftarrow^{Id} (E, f).$$

Ceci s'obtient naturellement étant donné que f s'applique à des éléments de (E, e) et renvoie des éléments de (E, e), et similairement pour f'.

De plus, il n'est pas difficile d'obtenir la matrice P de $Id\ :\ (E, f) \to (E, e)$: il s'agit (comme pour toute matrice) d'écrire les images des éléments de la base de "départ" dans la base d'"arrivée". Ici, comme on a affaire à l'identité de E, il s'agit d'écrire les éléments de f dans la base e. D'où $P = Mat_e(f)$. De la même manière, on a $Mat_f(e)$ pour la matrice de $Id\ :\ (E, e) \to (E, f)$, et comme, bien heureusement, $(E, f) \leftarrow^{Id} (E, e) \leftarrow^{Id} (E, f) = (E, f) \leftarrow^{Id} (E, f)$, on a bien $Mat_f(e) = Mat_e(f)^{-1} = P^{-1}$.

D'où l'écriture finale, $M' = P^{-1}MP$ (et $M = PMP^{-1}$).

2 Premières informations : le pivot de Gauss

Le but de cette partie est de comprendre quelles sont les conséquences de l'algorithme du pivot de Gauss : on y verra notamment une interprétation en terme d'action de groupe et d'orbites.

2.1 Opérations élémentaires

Elles correspondent toutes à la multiplication, à droite ou à gauche, par une matrice inversible.

2.1.1 transvections

Ceci correspond aux opérations sur les lignes $L_i \leftarrow L_i + \lambda L_j$, ou la même chose sur les colonnes.

Cela correspond, pour une matrice à A, à regarder PA, avec P donnée par :

$$\begin{pmatrix} 1 & & & O & \\ & 1 & & \lambda & \\ & & . & & \\ & O & & 1 & \\ & & & & 1 \end{pmatrix} = I_n + \lambda E_{ij}$$

On a le même genre de résultat sur les colonnes.

Théorème 2.1. *Les transvections engendrent $SL_n(\mathbb{K})$.*

C'est une conséquence du pivot de Gauss.

2.1.2 dilatations

Ceci correspond aux opérations sur les lignes $L_i \leftarrow \lambda L_i$, ou la même chose sur les colonnes.

Cela correspond, pour une matrice à A, à regarder PA, avec P donnée par :

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & O & \\ & & & \lambda & & \\ & & O & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = I_n + (\lambda - 1)E_{ii}$$

On a le même genre de résultat sur les colonnes.

Théorème 2.2. *Les transvections plus les dilatations engendrent $GL_n(\mathbb{K})$. Les transvections plus les homothéties engendrent aussi $GL_n(\mathbb{K})$.*

Maintenant, si l'on considère $GL(E)$ où E est un espace vectoriel de dimension finie, sans fixer de base, on a des résultats légèrement différents.

Définition 2.3. $f \in GL(E)$ est une dilatation s'il existe une base B de E telle que la matrice $Mat_B(E)$ soit une matrice de dilatation (*i.e.* de la forme donnée plus haut).

Théorème 2.4. *Les dilatations engendrent $GL(E)$, si $char(k) \neq 2$.*

Démonstration. On sait que si l'on fixe une base B, matrices de dilatations et matrices de transvections engendrent $GL_n(k) = GL((E, B))$. On va montrer que toute transvection est un produit de matrices de dilatation.

Soit $\lambda \in k$, $\lambda \neq 0$. Il suffit de le faire pour $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$.

Si $\lambda \neq 1$, on remarque que :

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & \lambda^{-1} \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}.$$

Or, dans la partie de droite, la première matrice correspond à un endomorphisme qui est une dilatation : elle a deux valeurs propres distinctes, 1 et λ^{-1}, donc dans une base qui la diagonalise, on a une matrice de dilatation. La seconde étant une matrice de dilatation, on a bien le résultat.

Pour $\lambda = 1$, on peut remarque que $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}^2$, et ainsi conclure : toute matrice de transvection est un produit de dilatations.

Avec les résultats connus sur $GL_n(k)$, on peut conclure pour $GL(E)$.

□

2.1.3 permutations

Ceci correspond aux opérations sur les lignes $L_i \leftrightarrow L_j$, ou la même chose sur les colonnes.

Cela correspond, pour une matrice à A, à regarder PA, avec P donnée par :

$$\begin{pmatrix} 1 & & & & O & \\ & \ddots & & & & \\ & & {}^1 0_1 & & 1 & \\ & & & \ddots & & \\ & & 1 & & {}^1 0_1 & \\ & & & & & \ddots \; 1 \end{pmatrix} = I_n + E_{ij} + E_{ji} - E_{ii} - E_{jj}$$

On remarquera que ces matrices sont dans $\pm SL_n(\mathbb{K})$, donc s'écrive (et plutôt facilement) comme produit de matrices de transvections et de dilatations.

2.2 Les informations lisibles

2.2.1 Échelonnement

Définition 2.5. Une matrice est dite échelonnée si la fonction $i \mapsto ind(i)$, où $ind(i)$ est l'indice de la ligne i, est strictement croissante.

Remarque. Il n'est pas suffisant d'être triangulaire pour être échelonnée.

Il n'y a pas de terme pour "échelonné à permutation près", qui est pourtant suffisant pour toutes les applications que l'on souhaite.

2.2.2 Rang et noyau

Multiplier par une matrice inversible à gauche conserve le noyau, par une matrice inversible à droite, c'est l'image qui est conservée. On peut trouver tout de même dans le premier cas une base de l'image.

Calcul du rang immédiat lors de l'échelonnement.

Calcul du noyau : action sur les colonnes, les répercuter sur $(e_1, \ldots, e_n)$, base de la source. On lira les images de ces vecteurs sur la matrice, et on pourra ainsi lire le noyau.

2.2.3 Image

Pour ce qui est du calcul d'une base de l'image, on peut agir sur les colonnes pour échelonner la matrice selon les colonnes. Ceci ne change rien à l'image, et on pourra alors en calculer une base. On peut se souvenir des opérations faites pour avoir une préimage de cette base.

2.3 Orbites

2.3.1 Action par translation

On fait agir $GL_n(\mathbb{K})$ sur $M_{n,p}(\mathbb{K})$ par translation : $P \cdot A = PA$.

Alors deux matrices sont *équivalentes à gauche* si elles sont dans la même orbite. Ceci a lieu *si et seulement si elles ont le même noyau.*

De même, on fait agir $GL_p(\mathbb{K})$ sur $M_{n,p}(\mathbb{K})$ par translation : $P \cdot A = AP^{-1}$.

Alors deux matrices sont *équivalentes à droite* si elles sont dans la même orbite. Ceci a lieu *si et seulement si elles ont la même image.*

Dans chaque orbite, on a un représentant échelonné en ligne (respectivement, en colonnes).

2.3.2 Action par équivalence sur un corps

Orbites sous la double action des matrices inversibles à gauche et à droite : classes d'équivalence. Il y en a $n+1$, classées par leur rang :

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Celle de rang n correspond à $GL_n(k)$.

Proposition 2.6. *$GL_n(k)$ est engendré par les matrices de transformations élémentaires, ou éventuellement juste les transvections et les dilatations, et $SL_n(k)$ juste les transvections.*

Remarque. On connait les orbites pour l'action par équivalence sur un anneau euclidien (et on peut étant donné une matrice, reconnaître son orbite), elles sont données par la *forme normale de Smith* (et l'algorithme des facteurs invariants).

Remarque. On connait les orbites pour l'action par similitude $P \cdot M = P^{-1}MP$. Elles sont données par la réduction de Frobenius, et, si le corps est algébriquement clos, par la réduction de Jordan. Nous allons dans la suite montrer comment calculer "à la main" dans quelle orbite nous nous trouvons

Remarque. Voir [6] pour des propriétés topologiques amusantes de ces orbites.

3 Réduction “effective” d'une matrice

3.1 Autour de la diagonalisabilité

Ce que l'on a vu suffit pour considérer le problème de la réduction des matrices diagonalisables : on trouve une base de vecteurs propres par ce qui précéde, et écrire A dans cette base la rend, par définition, diagonale.

3.1.1 Caractéristique nulle

On peut décider si une matrice est diagonalisable : $\frac{\chi_A}{\chi_A \wedge \chi_A'}(A) = 0$.

On peut décider si une matrice est à valeur propres simples : $Disc(A) \neq 0$. En particulier, de telles matrices forment un ouvert (lorsqu'on a un corps de base de caractéristique nulle).

3.1.2 Corps finis

Proposition 3.1. *A est diagonalisable sur $\mathbb{F}_q$ si et seulement si $X^q - X$ annule A.*

Démonstration. Définition de $\mathbb{F}_q$ d'un côté, $X^q - X$ y étant scindé à racines simples. Diagonalisation de l'autre. □

3.2 Étudier $XI_n - A$

Pour diagonaliser ou calculer la décomposition de Jordan d'une matrice, on doit calculer une base pour chacun de ses vecteurs propres, ou respectivement pour ses espaces caractéristiques. Une bonne manière de faire sans avoir à répéter les mêmes calculs plusieurs fois est d'échelonner directement $XI_n - A$.

Pour ce qui est des espaces propres, on a la définition suivante :

Définition 3.2. Le noyau de $\lambda I_n - A$, si non trivial, est l'espace propre associé à la valeur propre λ de A.

On peut faire le premier pas de calcul de tout ces noyaux en même temps en échelonnant $XI_n - A$!

Il faut faire bien attention aux choix des pivots et ainsi, aux déterminants des matrices par lesquelles on multiplie : ils doivent être inversibles (en particulier, non nul quelque soit l'évaluation en X...) ! On ne prend donc des matrices de transvections et de dilatations que dans $GL_n(k[X])$. Cela gêne surtout pour les dilatations, et en fait, on peut se passer d'elles....

Remarque. On peut se passer d'effectuer les permutations. On peut tout à fait lire l'échelonnement.

On obtient, à permutation des lignes près, des matrices de la forme :

$$\widetilde{M(X)} = \begin{bmatrix} \alpha_1(X) & & * \\ & \ddots & \\ 0 & & \alpha_n(X) \end{bmatrix}$$

En général, $a_n(X) = \chi_A(X)$, sauf si l'on a trouvé des valeurs propres ou des facteurs de ce polynôme, qui seront sur les autres coefficients diagonaux.

Maintenant, trouver les espaces propres demande déjà de trouver les λ tels que $\lambda I_n - A$ ne soit pas injective. Ceci peut se faire en trouvant les zéros de $\det(XI_n - A)$, mais aussi en considérant directement notre matrice échelonnée, $\widetilde{M(X)}$, où l'on peut déjà avoir obtenu des facteurs du polynôme caractéristique sur les coefficients diagonaux, ce qui peut faciliter le calcul de son rang en fonction de λ.

Une fois trouvées ces valeurs propres, il suffit pour chaque λ valeur propre, de calculer le noyau de $\widetilde{M(\lambda)}$, qui est déjà échelonnée, et l'on obtient les espaces propres.

3.3 Cas trigonalisable non-diagonalisable

Il faut faire le calcul des espaces caractéristiques ou des blocs de Jordan. Ceci est moins évident.

3.3.1 Cas d'un bloc de Jordan de taille 2

On a trouvé un vecteur propre F_1. pour la valeur propre λ. On voudrait trouver F_2 tel que $AF_2 = F_1 + \lambda F_2$. Ceci correspondrait à créer un bloc diagonal de la forme $\begin{bmatrix} \lambda & 1 \\ & \lambda \end{bmatrix}$.

Or $AF_2 = F_1 + \lambda F_2$ équivaut à $(A - \lambda I_n)F_2 = F_1$.

Il suffit donc d'utiliser ce que l'on connait du calcul de l'image pour résoudre ce problème :

- On peut lire sur $\widetilde{M(\lambda)}$ une base de l'image.
- On peut essayer d'écrire F_1 dans cette base (par réduction de Gauss par exemple). Si l'on réussit, alors on a directement trouvé F_2 (en effectuant les mêmes opérations sur une préimage de cette base).
- Sinon, c'est que l'équation n'a pas de solution...

Exemple. Soit $A = \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & -2 \\ 3 & -1 & -2 \end{pmatrix}$. Trigonalisons cette matrice.

Après quelques laborieux calculs, on trouve :

$$\widetilde{M(T)} = \begin{pmatrix} 0 & \frac{8}{9} & \frac{1}{3}T^2 + \frac{2}{9}T - \frac{5}{9} \\ -2 & T-1 & 2 \\ 0 & 0 & \frac{9}{16}T^3 - \frac{9}{16}T^2 - \frac{9}{16}T + \frac{9}{16} \end{pmatrix} = \begin{pmatrix} 0 & \frac{8}{9} & \frac{1}{3}T^2 + \frac{2}{9}T - \frac{5}{9} \\ -2 & T-1 & 2 \\ 0 & 0 & \frac{9}{16}(T+1)(T-1)^2 \end{pmatrix}.$$

Il est alors évident que les espaces propres E_1 et E_{-1} sont de dimension 1.

On trouve directement $E_1 = Vect((1,0,1)^t)$ et $E_{-1} = Vect((1/2, 1/2, 1)^t)$.

Enfin, il faut maintenant chercher F tel que $(A - I_n)F = E_1$. On lit sur $\widetilde{M(1)}$ que $(A - I_n)e_1, (A - I_n)e_2$ est une base de l'image de $A - I_n$. Or, $(A - I_n)e_1 = (1,2,3)^t$, $(A - I_n)e_2 = (-1,0,-1)^t$. Clairement, $-e_2$ convient.

Ainsi, dans la base (E_{-1}, E_1, e_2), A s'écrit :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

3.3.2 Cas général

Le principe est le même, mais il faut trouver les tailles des blocs de Jordan. On peut utiliser la technique précédente pour chaque vecteur propre qui correspond à une valeur propre multiple.

Pour réussir à former un bloc $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$, une fois connu le vecteur propre correspondant à la première colonne, e_1, il suffit de trouver e_2 tel que $(A-\lambda I_n)e_2 = e_1$, puis ensuite e_3 tel que $(A-\lambda I_n)e_3 = e_2$.

Remarque. On a un bloc de Jordan par valeur propre.

Exemple. Soit $A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, trigonalisons cette matrice.

Après quelques laborieux calculs, on trouve :

$$\widetilde{M(T)} = \begin{pmatrix} 0 & -1 & -T^2+3T-2 \\ 0 & 0 & -T^3+3T^2-3T+1 \\ 1 & -1 & T-2 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -T^2+3T-2 \\ 0 & 0 & -(T-1)^3 \\ 1 & -1 & T-2 \end{pmatrix}$$

. Il est alors facile de voir que $\chi_A = (X-1)^3$ (et ainsi que la matrice est trigonalisable), mais aussi que $F_1 = e_1 + e_3$ génére l'unique droite propre, associée à la valeur propre 1.

Reste maintenant à trouver F_2 et F_3 pour constituer le reste de la décomposition en blocs de Jordan.

On remarque que $(A-I_n)e_1 = (-1,-1,-1)^t$ et $(A-I_n)e_2 = (1,0,1)^t$ forment une base de l'image de $(A-I_n)$.

Trouvons F_2 tel que $(A-I_n)F_2 = F_1$. On a directement $F_2 = e_2$ qui convient.

Trouvons F_3 tel que $(A-I_n)F_3 = F_2$. On voit que $F_3 = e_2 - e_1$ convient.

La base (e_1+e_3, e_2, e_2-e_1) convient donc.

4 Réduction simultanée de deux formes quadratiques

4.1 Un mot sur le théorème spectral

Théorème 4.1 (Théorème spectral)**.** *Soit a un produit scalaire sur $\mathbb{R}^n$, q sa forme quadratique associée, et soit $u \in M_n(\mathbb{R})$ auto-adjointe pour ce produit scalaire. Alors il existe $P \in O(q)$, orthonormale pour q, tel que $P^{-1}uP = D$ où $D \in M_n(\mathbb{R})$ est diagonale.*

Démonstration. Voici une idée de preuve.

- On montre que l'on peut étendre les scalaires à $\mathbb{C}$ pour pouvoir étudier a et q sur $\mathbb{C}$.
- On remarque que si x est vecteur propre de u associé à une valeur propre $\lambda \in \mathbb{C}$ de u (et il en existe vu que l'on est sur $\mathbb{C}$), alors $a(u(x), x) = \overline{\lambda} a(x, x) = \lambda a(x, x)$, et donc λ est réelle. Ensuite, on peut prendre $x + \overline{x}$ comme vecteur propre réel associé à la valeur propre réelle λ de u, et on peut revenir aux réels.
- On remarque que l'orthogonal, selon a, d'un espace-propre de u est stable par u.
- On peut, enfin, procéder par récurrence sur n. Le cas $n = 1$ est évident. Ensuite, pour un $u \in M_n(\mathbb{R})$ et un a comme avant, on peut utiliser Gram-Schmidt sur E_λ pour $\lambda \in \mathbb{R}$ comme précédement, et utiliser l'hypothèse de récurrence sur l'orthogonal de E_λ.

□

On en déduit le bien connu corollaire immédiat :

Corollaire 4.2. *Soit $A \in S_n(\mathbb{R})$. Alors il existe $P \in O_n(\mathbb{R})$ tel que $P^{-1}AP = D = {}^tPAP$ avec $D \in M_n(\mathbb{R})$ diagonale.*

Remarque. On voit que dans le corollaire, le fait que $P \in O_n(\mathbb{R})$ fournit un changement de base ${}^tPAP$ pour la forme quadratique associée à A.

Par contre, si l'on ne prend pas pour a le produit scalaire usuel, le théorème ne donne qu'une diagonalisation de l'endomorphisme, et rien sur la forme quadratique. Le théorème suivant permet de palier à ce problème.

4.2 Réduction simultanée

Théorème 4.3 (Réduction simultanée de deux formes quadratiques)**.** *Soit $A \in S_n^{++}(\mathbb{R})$ et $B \in S_n(\mathbb{R})$. Alors il existe $P \in GL_n\mathbb{R}$ tel que :*

- *P est orthonormale pour la forme quadratique définie par A :*

$$ {}^tPAP = I_n $$

- *P "diagonalise" la forme quadratique définie par B, il existe $D \in M_n(\mathbb{R})$ diagonale telle que :*

$$ {}^tPBP = D $$

Démonstration. On remarque que $A^{-1}B$ est auto-adjoint pour le produit scalaire définit par A.

En effet, ${}^t(A^{-1}B)A = B * A^{-1} * A = B$ et $A * (A^{-1}B) = B$.

On peut donc appliquer le théorème 4.1 : il existe $P \in O(q_A)$, orthonormale pour q_A tel que $P^{-1}A^{-1}BP = D$ pour D une matrice diagonale de $M_n(\mathbb{R})$.

Le fait que P soit orthonormale pour q_A s'écrit ${}^tPAP = I_n$, ou encore, ${}^tP = P^{-1}A^{-1}$. Mais alors, $D = P^{-1}A^{-1}BP = {}^tPBP$. Ainsi P convient. $\square$

Remarque. Attention aux preuves trop rapides : "on applique le théorème spectral 4.1 à B qui est symétrique, avec le produit scalaire défini par A".

D'une part, le résultat ne correspondrait pas (on veut ${}^tPBP = D$ et non $P^{-1}BP$), et d'autre part, B n'a *aucune* raison d'être auto-adjointe pour A (et ne le sera pas en général).

Pour en savoir plus sur ce problème, et notamment les applications du théorème de réduction simultanée de deux formes quadratiques (ainsi que le problème de l'effectivité de la recherche de P), voir [1]

5 Décomposition de Dunford effective

5.1 Dunford effectif

Le but de ce développement est de démontrer le théorème suivant :

Théorème 5.1. *Il existe un algorithme effectif pour calculer la décomposition de Dunford d'une matrice (i.e. sans calculer les valeurs propres).*

L'essentiel des idées de la preuve se trouve dans le Risler-Boyer [7], mais celui-ci comporte une erreur grave dans l'écriture de sa preuve.

Après une remarque/proposition préliminaire, la démonstration consiste en deux lemmes et une proposition qui sont suivis de la méthode effective de calcul.

Dans tout ce qui suit, on se place sur $\mathbb{K} \subset \mathbb{C}$, et tout les calculs effectifs pourront être fait dans $\mathbb{K}$. Cela est en particulier vrai pour $\mathbb{K} = \mathbb{Q}$.

5.2 Préliminaires

Remarque. Soit $P \in \mathbb{K}[X]$. Alors on peut calculer $sqfree(P)$ dans $\mathbb{K}$.

Démonstration. En effet, on suppose que $deg(P) > 1$, et on écrit $P = \prod_{i=1}^{r}(X - \lambda_i)^{n_i}$, avec $r \in \mathbb{N}^*$, $\sum n_i = deg(P)$, $\lambda_i \in \mathbb{C}$ tous distincts, et on suppose pour simplifier les écritures que P est unitaire.

Alors, soit $Q = sqfree(P) = \prod_{i=1}^{r}(X - \lambda_i)$. On a $Q = P/(P \wedge P')$.

En effet, on a directement que $val_{x-\lambda_i}(P/(P \wedge P')) = val_{x-\lambda_i}(P) - min(val_{x-\lambda_i}(P), val_{x-\lambda_i}(P')) = n_i - min(n_i, n_i - 1) = 1$, et d'autre part, $P/(P \wedge P')$ divise P, donc est de la forme $\prod_{i=1}^{r}(X - \lambda_i)^{l_i}$. On a donc directement $l_i = 1$ et $Q = P/(P \wedge P')$.

De plus, $P \wedge P'$ se calcule sur $\mathbb{K}$ par algorithme d'Euclide, et divise bien P dans $\mathbb{K}[X]$, donc on a bien $P/(P \wedge P') \in \mathbb{K}[X]$ qui peut se calculer dans $\mathbb{K}[X]$. On a donc bien montré ce que l'on voulait. □

Remarque. Ce qui précède marche aussi bien pour n'importe quel corps de caractéristique nulle.

Lemme 5.2. *Si $B \in \mathbb{K}[A] \cap GL_n(\mathbb{K})$, alors $B^{-1} \in \mathbb{K}[A]$.*

Démonstration. En effet, on $B^{-1} \in \mathbb{K}[B]$ par Cayley-Hamilton, et $B \in \mathbb{K}[A]$ donc $B^{-1} \in \mathbb{K}[A]$. □

Corollaire 5.3.

$$\mathbb{K}[A]^* = \mathbb{K}[A] \cap GL_n(\mathbb{K}).$$

Lemme 5.4. *Soit $U \in GL_n(\mathbb{K})$, $N \in M_n(\mathbb{K})$ nilpotente, tels que $UN = NU$. Alors $U - N$ est inversible.*

Démonstration. Soit $k \in \mathbb{N}^*$ tel que N^{k+1}. Alors

$$(Id - U^{-1}N)(Id + u^{-1}N + \cdots + (U^{-1}N)^k) = Id - (U^{-1}N)^{k+1} = Id.$$

□

En conséquence de ce qui précède :

Proposition 5.5. *Soit $A \in M_n(\mathbb{K})$. On pose $P = sqfree(\chi_A)$. Alors $P'(A) \in \mathbb{K}[A]^*$.*

Démonstration. P étant sans facteurs carrés, on a $P \wedge P' = 1$. Il existe donc $U, V \in \mathbb{K}[X]$ tels que $UP' + VP = 1$.

Mais alors, $U(A)P'(A) = Id - N$ avec $N = V(A)P(A)$ qui est nilpotente. En effet, $P(A)$ est nilpotente car $P(A)^l = 0$ pour $l = \max_i(n_i)$ et $\chi_A = \prod_{i=1}^r (X - \lambda_i)^{n_i}$ (sur $\mathbb{C}$).

Avec le lemme précédent, $U(A)P'(A) \in GL_n(\mathbb{K})$, donc $P'(A) \in GL_n(\mathbb{K})$. □

5.3 Méthode de calcul

On pose maintenant $A \in M_n(\mathbb{K})$, $\chi_A = \prod_{i=1}^r (X - \lambda_i)^{n_i}$ comme plus haut, et $P = sqfree(\chi_A)$.

Nous allons définir une suite de matrices $(A_n)_{n \in \mathbb{N}}$ telle que celle-ci soit constante à partir d'un certain rang, égale au D diagonalisable de la décomposition de Dunford.

Théorème 5.6. *On pose $A_0 = A$ et $A_{n+1} = A_n - P(A_n)(P'(A_n))^{-1}$. Alors la suite $(A_n)_{n\in\mathbb{N}}$ est bien définie, et telle que pour tout $n \in \mathbb{N}$, $P'(A_n) \in \mathbb{K}[A]^*$.*

De plus, pour tout $n \in \mathbb{N}$, $P(A_n) = P(A)^{2^n} B_n$, avec $B_n \in \mathbb{K}[A]$.

Montrons d'abord que si le théorème qui précède est vrai, nous avons bien A_n constante à partir d'un certain rang, et diagonalisable.

Du fait que $P(A)$ soit nilpotente (on l'a vu plus haut), il existe $r \in \mathbb{N}^*$ tel que $P(A)^{2^r} = 0$. Mais alors, $P(A_n) = 0$ pour tout $n \geqslant r$, et, vu sa définition, $A_n = A_r$.

Mais alors, si on pose $D = A_r$, on a $P(D) = 0$ et D est sans facteur carré, donc D est diagonalisable sur $\mathbb{C}$.

De plus, on peut écrire :

$$D - A = \sum_{i=0}^{r-1}(A_{i+1} - A_i) = \sum_{i=0}^{r-1}(-P(A_i)P'(A_i)^{-1}) = P(A)^{2^i} B_i(A)P'(A_i)^{-1}.$$

Or $P(A)^{2^i}$ est nilpotente et comme toutes les matrices sont dans $\mathbb{K}[A]$, tout commute, donc en particulier, chacun des termes de la somme, et la somme en elle-même, est nilpotente.

On peut donc bien écrire $A = D + N$ avec $D, N \in \mathbb{K}[A]$, D diagonalisable (sur $\mathbb{C}$) et N nilpotente. De plus, tout ce qui précède permet bien de voir que l'on peut calculer D, en un nombre fini de calculs sur $M_n(\mathbb{K})$, ce que l'on souhaitait démontrer.

Finalement, donnons une démonstration du dernier théorème :

Démonstration. Nous allons montrons la bonne définition et toutes les propriétés vérifiées par la suite grâce à une récurrence sur n.

Pour $n = 0$, il n'y a rien à ajouter à ce que l'on a déjà vu (en particulier à la proposition 5.5).

On suppose le résultat connu pour un certain $n \in \mathbb{N}$.

On écrit $P(A_{n+1}) = P(A_n + Y) = P(A_n) + YP'(A_n) + Y^2\tilde{P}(A_n, Y)$ par la formule de Taylor à l'ordre 2 pour les polynômes, pour un certain polynôme $\tilde{P}$.

Par définition de la suite, $Y = -P(A_n)P'(A_n)^{-1}$, et $P(A_n) = P(A)^{2^n} B_n$.

Mais alors,

$$P(A_{n+1}) = P(A)^{2^{n+1}}(P'(A_n)^{-1})^2 B_n^2 \tilde{P}(A_n, -P(A_n)P'(A_n)^{-1}).$$

On a bien $(P'(A_n)^{-1})^2 B_n^2 \tilde{P}(A_n, -P(A_n)P'(A_n)^{-1}) \in \mathbb{K}[A]$. On peut donc écrire $P(A_{n+1}) = P(A)^{2^{n+1}} B_{n+1}$ pour un certain $B_{n+1} \in \mathbb{K}[A]$.

De plus, aussi par la formule de Taylor, pour un certain $Q \in \mathbb{K}[X]$:

$$P'(A_{n+1}) - P'(A_n) = (A_{n+1} - A_n)Q(A_n).$$

Vue la définition de A_{n+1}, cela peut être réécrit : $P'(A_{n+1}) - P'(A_n) = P(A_n)P'(A_n)^{-1}Q(A_n)$.

Or, $P(A_n)$ est nilpotente et tout commute car dans $\mathbb{K}[A]$, donc $P(A_n)P'(A_n)^{-1}Q(A_n)$ est nilpotente, et par le lemme 5.4, $P'(A_{n+1}) \in \mathbb{K}[A]^*$.

La récurrence est donc achevée, et le théorème montré. $\square$

Aurait aussi été possible

Décomposition de Jordan, de Bruhat, de Frobenius, réduction des matrices normales, plus d'actions par congruence (théorème de Sylvester et algorithme de Gauss), réduction des quadriques...

Références

[1] Paugam, Annette Agrégation de mathématique : Questions délicates en Algèbre et Géométrie

[2] Grifone, Joseph Algèbre linéaire

[3] Mneimé & Testard Introduction à la théorie des groupes de Lie classiques

[4] Beck, Malick & Peyré Objectif Agreg

[5] Francinou, Gianella, Nicolas Oraux X-ENS : Algèbre tome 1

[6] Francinou, Gianella, Nicolas Oraux X-ENS : Algèbre tome 2

[7] Risler, Boyer Algèbre pour la licence 3 : groupes, anneaux, corps (Dunod, 2006).