SEMAINE 4

RÉDUCTION DES ENDOMORPHISMES (DEUXIÈME PARTIE)

EXERCICE 1 :

Une matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ est dite **stochastique** lorsqu'elle vérifie les deux conditions suivantes :

(i) $\forall i \in [\![1,n]\!]\ \forall j \in [\![1,n]\!] \quad a_{ij} \in [0,1]$;

(ii) $\forall i \in [\![1,n]\!] \quad \displaystyle\sum_{j=1}^{n} a_{ij} = 1$ (la somme des éléments de chaque ligne vaut 1).

Elle est dite **stochastique stricte** si, de plus, les coefficients a_{ij} sont tous strictement positifs.

On notera $\mathcal{S}_n$ l'ensemble des matrices stochastiques de $\mathcal{M}_n(\mathbb{R})$, et $\mathcal{S}_n^*$ celui des matrices stochastiques strictes.

1. Montrer que les ensembles $\mathcal{S}_n$ et $\mathcal{S}_n^*$ sont stables par produit.

2. Si $A \in \mathcal{S}_n$, montrer que 1 est valeur propre de A.

3. Si $A \in \mathcal{S}_n^*$, montrer que $\mathrm{Ker}(A - I_n)$ est de dimension un.

4. Montrer que les valeurs propres d'une matrice stochastique sont toutes de module inférieur ou égal à 1, et que les valeurs propres autres que 1 d'une matrice stochastique stricte sont de module strictement inférieur à 1.

5. Soit $A \in \mathcal{S}_n$, soit λ une valeur propre de A. Montrer qu'il existe $i \in [\![1,n]\!]$ tel que

$$|\lambda - a_{ii}| \leq 1 - a_{ii} .$$

- -

1. Soient $A = (a_{ij})$ et $B = (b_{ij})$ stochastiques. On a $AB = (c_{ik})$, où $c_{ik} = \displaystyle\sum_{j=1}^{n} a_{ij} b_{jk}$.

- Il est clair que $c_{ik} \geq 0$ pour tout couple d'indices (i,k), l'inégalité étant stricte si A et B sont dans $\mathcal{S}_n^*$.
- On a $b_{jk} \leq 1$ pour tout (j,k), donc $c_{ik} = \displaystyle\sum_{j=1}^{n} a_{ij} b_{jk} \leq \sum_{j=1}^{n} a_{ij} = 1$ pour tout couple (i,k).
- Enfin,

$$\sum_{k=1}^{n} c_{ik} = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) = \sum_{j=1}^{n} \left(a_{ij} \sum_{k=1}^{n} b_{jk} \right) = \sum_{j=1}^{n} a_{ij} = 1 .$$

On peut aussi remarquer qu'une matrice A vérifie la propriété **(ii)** *si et seulement si* $AJ = J$, *où* J *est la matrice dont tous les coefficients valent 1. Il est alors immédiat que cette propriété* **(ii)** *est "stable par produit".*

2. Si $A \in \mathcal{S}_n$, alors $AX = X$, où X est le vecteur dont toutes les coordonnées valent 1, donc 1 est valeur propre de A.

3. Soit $B = A - I_n$, soit C la matrice carrée d'ordre $n-1$ extraite de B en ôtant la dernière ligne et la dernière colonne :

$$C = \begin{pmatrix} a_{11} - 1 & a_{12} & \dots & a_{1,n-1} \\ a_{21} & a_{22} - 1 & \dots & a_{2,n-1} \\ \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} - 1 \end{pmatrix} .$$

Montrons que C est inversible. Cela repose sur le fait que $C = (c_{ij})$ est **à diagonale strictement dominante**, c'est-à-dire que, pour tout $i \in [\![1, n-1]\!]$, on a $|c_{ii}| > \sum_{j \neq i} |c_{ij}|$: en effet,

$$|c_{ii}| - \sum_{\substack{1 \leq j \leq n-1 \\ j \neq i}} |c_{ij}| \; = \; |a_{ii} - 1| - \sum_{\substack{1 \leq j \leq n-1 \\ i \neq j}} a_{ij} \; = \; 1 - \sum_{j=1}^{n-1} a_{ij} = a_{in} > 0 \; .$$

Soit donc $X = (x_1, \cdots, x_{n-1}) \in \operatorname{Ker} C$, supposé non nul ; on a, pour tout $i \in [\![1, n-1]\!]$, $\sum_{j=1}^{n-1} c_{ij} x_j = 0$. Soit $i_0 \in [\![1, n-1]\!]$ tel que $|x_{i_0}| = \max_{1 \leq i \leq n-1} |x_i|$, on a alors, pour $i = i_0$, $c_{i_0 i_0} x_{i_0} = -\sum_{j \neq i_0} c_{i_0 j} x_j$, mais c'est impossible car

$$\left| \sum_{j \neq i_0} c_{i_0 j} x_j \right| \leq \sum_{j \neq i_0} |c_{i_0 j}| |x_j| \leq |x_{i_0}| \left(\sum_{j \neq i_0} |c_{i_0 j}| \right) < |x_{i_0}| |c_{i_0 i_0}| \; .$$

La matrice C est donc inversible,c'est-à-dire de rang $n-1$, donc $B = A - I_n$ est de rang au moins égal à $n-1$, donc exactement $n-1$ puisqu'on sait que 1 est valeur propre de A, et donc $\dim \operatorname{Ker}(A - I_n) = 1$.

On a ainsi prouvé le **théorème d'Hadamard** (*moi, froid ? jamais...*) : *toute matrice à diagonale strictement dominante est inversible.*

4. Soit $A \in \mathcal{S}_n$, soit $\lambda \in \mathbb{C}$ tel que $|\lambda| > 1$; alors la matrice $B = A - \lambda I_n = (b_{ij})$ est à diagonale strictement dominante : en effet,

$$|b_{ii}| = |a_{ii} - \lambda| \geq \big| |a_{ii}| - |\lambda| \big| = |\lambda| - a_{ii} > 1 - a_{ii} = \sum_{j \neq i} a_{ij} = \sum_{j \neq i} |b_{ij}| \; .$$

La matrice B est donc inversible, et $\lambda \notin \operatorname{Sp}(A)$.

De même, si $A \in \mathcal{S}_n^*$ et si $|\lambda| = 1$ avec $\lambda \neq 1$, alors la matrice $B = A - \lambda I_n$ est encore à diagonale strictement dominante, puisque

$$|b_{ii}| = |a_{ii} - \lambda| > \big| |a_{ii}| - |\lambda| \big| = 1 - a_{ii} = \sum_{j \neq i} a_{ij} = \sum_{j \neq i} |b_{ij}|$$

(l'inégalité est stricte puisque l'égalité $\big| |u| - |v| \big| = |u - v|$ a lieu si et seulement si les complexes u et v sont "colinéaires de même sens", c'est-à-dire l'un des deux nuls ou $\frac{v}{u} \in \mathbb{R}_+^*$ et ce n'est pas le cas ici : $\lambda \notin \mathbb{R}_+$). Donc $B = A - \lambda I_n$ est inversible, et λ n'est pas valeur propre de A.

5. Par contraposition, c'est toujours le même raisonnement : si on avait $\forall i \in [\![1, n]\!] \quad |\lambda - a_{ii}| > 1 - a_{ii}$, la matrice $B = A - \lambda I_n$ serait à diagonale strictement dominante, donc inversible, et λ ne serait pas valeur propre de A.

On a ainsi obtenu une localisation des valeurs propres : si A est une matrice stochastique,

$$\mathrm{Sp}(A) \subset \bigcup_{i=1}^{n} D(a_{ii}, 1-a_{ii}) \qquad\qquad (D = \textit{disque fermé}) \; .$$

En fait, cette dernière question se généralise facilement à une matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$ *quelconque ; avec les mêmes méthodes, on montre, que si on pose* $r_i = \sum_{j \neq i} |a_{ij}|$ *pour tout* $i \in [\![1, n]\!]$, *on a*

$$\mathrm{Sp}(A) \subset \bigcup_{i=1}^{n} D(a_{ii}, r_i) \; .$$

EXERCICE 2 :

Soit E un $\mathbb{C}$-espace vectoriel de dimension n, soit $\mathcal{F}$ un ensemble d'endomorphismes de E qui commutent deux à deux.

Montrer l'existence d'une "décomposition de Dunford simultanée", c'est-à-dire d'une liste d'entiers naturels non nuls $n_1, \cdots, n_p$ avec $\sum_{i=1}^{p} n_i = n$ et d'une base $\mathcal{B}$ de E tels que, dans la base $\mathcal{B}$, tout élément f de $\mathcal{F}$ soit représenté par une matrice diagonale par blocs de la forme $M_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1 I_{n_1} + N_1, \cdots, \lambda_p I_{n_p} + N_p)$, les λ_i étant des nombres complexes et chaque matrice $N_i \in \mathcal{M}_{n_i}(\mathbb{C})$ étant triangulaire supérieure avec des zéros sur la diagonale.

- -

Première méthode :

Introduisons la notion suivante : un sous-espace vectoriel V de E sera dit $\mathcal{F}$**-indécomposable** s'il est $\mathcal{F}$-stable (c'est-à-dire stable par chaque élément de $\mathcal{F}$) et si on ne peut pas le décomposer en $V = V_1 \oplus V_2$, avec V_1 et V_2 tous deux $\mathcal{F}$-stables et non réduits à $\{0\}$.

Si $\mathcal{F}$ est une partie quelconque de $\mathcal{L}(E)$, on montre (par récurrence forte sur la dimension de E) qu'il existe au moins une décomposition de E en somme directe de sous-espaces $\mathcal{F}$-indécomposables : $E = \bigoplus_{i=1}^{p} E_i$. Si les éléments de $\mathcal{F}$ commutent deux à deux, il en est de même des endomorphismes qu'ils induisent sur chaque E_i $(1 \leq i \leq p)$.

Pour terminer l'exercice, il reste alors à prouver le lemme suivant :

Lemme : si E *est* $\mathcal{F}$*-indécomposable, alors il existe une base de* E *dans laquelle les éléments de* $\mathcal{F}$ *ont tous des matrices de la forme* $\begin{pmatrix} \lambda & & (X) \\ & \ddots & \\ (0) & & \lambda \end{pmatrix}$, *c'est-à-dire* $\lambda I + N$ *avec* $\lambda \in \mathbb{C}$ *et* N *triangulaire supérieure avec des zéros sur la diagonale.*

(*Démonstration du lemme*) : Soit $f \in \mathcal{F}$, notons μ son polynôme minimal.

Supposons $\mu = \mu_1\mu_2$ avec μ_1 et μ_2 premiers entre eux et non constants. De $\mu(f) = 0$, on déduit (*lemme des noyaux*) $E = E_1 \oplus E_2$ avec $E_1 = \operatorname{Ker}\mu_1(f)$ et $E_2 = \operatorname{Ker}\mu_2(f)$. Tout élément g de $\mathcal{F}$ commute avec f, donc laisse stables E_1 et E_2. L'espace E étant supposé $\mathcal{F}$-indécomposable, l'un des sous-espaces E_i est réduit à $\{0\}$. Mais si l'on suppose par exemple $E_1 = \{0\}$, alors $E = E_2$ donc $\mu_2(f) = 0$ ce qui contredit la minimalité de μ.

On a ainsi prouvé que tout élément f de $\mathcal{F}$ a un polynôme minimal de la forme $(X - \lambda)^m$, donc est de la forme $\lambda \operatorname{id}_E + \nu$ avec ν nilpotent.

Les endomorphismes de $\mathcal{F}$ commutent deux à deux, donc sont cotrigonalisables (*exercice classique : on montre d'abord, par récurrence sur la dimension de E, l'existence d'un vecteur propre commun, puis on fait une nouvelle récurrence sur* $\dim E$*, comme dans l'exercice* **3** *question* **3** *de la semaine* **3***, pour construire une base de trigonalisation commune*). Chacun admettant une seule valeur propre, dans une base de trigonalisation commune, leurs matrices sont de la forme indiquée. (*fin de la dém. du lemme*)

Pour terminer l'exercice, il suffit de partir d'une décomposition $E = \bigoplus_{i=1}^{p}$ de E en sous-espaces $\mathcal{F}$-indécomposables, de construire une base $\mathcal{B}_i$ dans chaque E_i qui trigonalise tous les endomorphismes induits, et de concaténer ces différentes bases.

Deuxième méthode *proposée par Charles-Antoine GOFFIN, étudiant en MP* :*

Admettons toujours comme "classique" le fait qu'une famille d'endomorphismes commutant deux à deux dans un $\mathbb{C}$-espace vectoriel de dimension finie est cotrigonalisable, et raisonnons par récurrence forte sur $n = \dim E$:

- pour $n = 1$, c'est évident ;
- soit $n \geq 2$, si c'est vrai pour tout $k < n$, soit $\mathcal{F}$ une famille d'endomorphismes d'un $\mathbb{C}$-espace vectoriel E de dimension n qui commutent deux à deux.
  - ▷ si chacun des endomorphismes de la famille $\mathcal{F}$ a une seule valeur propre, c'est-à-dire est de la forme $\lambda \operatorname{id}_E + \nu$ avec ν nilpotent, comme ils sont cotrigonalisables, il existe bien une base dans laquelle ils ont tous une matrice de la forme $\lambda I_n + N$, avec N triangulaire supérieure avec des zéros sur la diagonale, et c'est terminé ;
  - ▷ sinon, au moins un des endomorphismes u de la famille $\mathcal{F}$ a plusieurs valeurs propres distinctes, soit λ une de ces valeurs propres, soit V le sous-espace caractéristique associé, soit W la somme de tous les autres sous-espaces caractéristiques de u. On a $E = V \oplus W$. Les sous-espaces V et W sont laissés stables par tous les endomorphismes de la famille $\mathcal{F}$ puisque V est le noyau d'un polynôme en u (et W une somme de...*idem*). Les endomorphismes de V et de W induits par les éléments de $\mathcal{F}$ commutent deux à deux et on peut leur appliquer l'hypothèse de récurrence puisque $\dim V < n$ et $\dim W < n$. Il ne reste plus qu'à concaténer les bases de V et de W ainsi construites et c'est fini.

EXERCICE 3 :

1. Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice inversible. Montrer l'existence d'un polynôme P de $\mathbb{C}[X]$ tel que $P(A)^2 = A$.

2. Montrer qu'une matrice $A \in \mathrm{GL}_n(\mathbb{C})$ est semblable à son inverse si et seulement si elle est le produit de deux involutions.

Indication : si $A^{-1} = P^{-1}AP$, on vérifiera que P^2 commute avec A, puis on introduira une matrice $Q \in \mathbb{C}[P^2]$ telle que $Q^2 = P^2$.

- -

1.a. Etudions d'abord le cas où A admet une seule valeur propre $\lambda \in \mathbb{C}^*$. Dans ce cas, $A = \lambda I + N$ avec N nilpotente, disons d'indice p ($N^p = 0$ et $N^{p-1} \neq 0$).

Supposons d'abord $\lambda = 1$. Soit

$$\sqrt{1+x} = 1 + a_1 x + \ldots + a_{p-1}x^{p-1} + o(x^{p-1}) = S(x) + o(x^{p-1})$$

le développement limité à l'ordre $p-1$ de la fonction $]-1, +\infty[\to \mathbb{R}$, $x \mapsto \sqrt{1+x}$ en zéro (sa partie régulière S est un polynôme de $\mathbb{R}[X]$ de degré inférieur ou égal à $p-1$). On a alors, au voisinage de zéro (pour une variable x réelle) :

$$(\sqrt{1+x})^2 = 1 + x = Q(x) + o(x^{p-1}) ,$$

où Q est le polynôme S^2 tronqué à l'ordre $p-1$: donc $1 + x = S(x)^2 + o(x^{p-1})$ et cette relation entre fonctions polynomiales, avec l'unicité du développement limité, montre que, dans $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, S^2 est congru à $1+X$ modulo X^p, notons $S(X)^2 = 1 + X + X^p R(X)$ avec $R \in \mathbb{R}[X]$.

On a donc $S(N)^2 = I + N + N^p R(N) = I + N = A$, soit $P(A)^2 = A$ où P est le polynôme défini par $P(X) = S(X-1)$.

Si $\lambda \neq 1$, on écrit $A = \lambda I + N = \lambda B$, avec $B = I + \dfrac{1}{\lambda}N = \dfrac{1}{\lambda}A$ et il existe un polynôme Q de $\mathbb{C}[X]$ tel que $Q(B)^2 = B$, donc $\lambda\, Q\left(\dfrac{1}{\lambda}A\right)^2 = A$. En notant μ une racine carrée complexe de λ et en posant $P(X) = \mu\, Q\left(\dfrac{X}{\lambda}\right)$, on a $P(A)^2 = A$.

1.b. Si A est une matrice inversible quelconque, notons u l'endomorphisme de $\mathbb{C}^n$ canoniquement associé : on décompose suivant les sous-espaces caractéristiques : si $\mu = \displaystyle\prod_{k=1}^{m}(X - \lambda_k)^{\beta_k}$ est le polynôme minimal de u (les λ_k étant distincts non nuls), d'après le lemme des noyaux, on a $\mathbb{C}^n = \displaystyle\bigoplus_{k=1}^{m} F_k$, avec $F_k = \mathrm{Ker}(u - \lambda_k\, \mathrm{id}_{\mathbb{C}^n})^{\beta_k}$, la restriction v_k de u à F_k admettant $(X - \lambda_k)^{\beta_k}$ comme polynôme annulateur (et même plus précisément comme polynôme minimal), ce qui signifie que $v_k = \lambda_k\, \mathrm{id}_{F_k} + \nu_k$, où ν_k est un endomorphisme nilpotent de F_k. Traduction matricielle : la matrice A est semblable à une matrice D diagonale par blocs : $D = \mathrm{diag}(J_1, \ldots, J_m)$ avec, pour tout k, $J_k = \lambda_k I_{\alpha_k} + N_k$, la matrice N_k étant nilpotente d'indice β_k (α_k est la dimension du sous-espace caractéristique F_k, c'est aussi la multiplicité de la valeur propre λ_k dans le polynôme caractéristique).

Bref, pour tout $k \in [\![1, m]\!]$, il existe un polynôme P_k tel que $P_k(J_k)^2 = J_k$ d'après la partie **1.a.** Il reste à montrer l'existence d'un polynôme P (indépendant de k) tel que

$$\forall k \in [\![1, m]\!] \qquad P(J_k) = P_k(J_k) \; . \qquad \textbf{(*)}$$

Mais cette condition **(*)** équivaut à

$$\forall k \in [\![1, m]\!] \qquad \mu_k \mid P - P_k \; ,$$

où $\mu_k = (X - \lambda_k)^{\beta_k}$ est le polynôme minimal de J_k. Les polynômes μ_k étant premiers entre eux deux à deux, l'existence d'un tel polynôme P résulte du théorème chinois : le système de congruences

$$P \equiv P_k \quad [\mu_k] \qquad (1 \leq k \leq m)$$

admet pour ensemble de solutions dans $\mathbb{C}[X]$ une classe de congruence modulo $\mu = \displaystyle\prod_{k=1}^{m} \mu_k$.

Démonstration par récurrence sur m : si μ_1 et μ_2 sont premiers entre eux, d'après Bézout, il existe des polynômes U_1 et U_2 tels que $P_1 - P_2 = V_2\mu_2 - V_1\mu_1$. Le polynôme $P_0 = P_1 + V_1\mu_1 = P_2 + V_2\mu_2$ est une "solution particulière" du système de congruences $\begin{cases} P \equiv P_1 & [\mu_1] \\ P \equiv P_2 & [\mu_2] \end{cases}$. Un polynôme quelconque P vérifie alors ce système si et seulement si $\begin{cases} P \equiv P_0 & [\mu_1] \\ P \equiv P_0 & [\mu_2] \end{cases}$, ce qui équivaut à $\begin{cases} \mu_1 \mid P - P_0 \\ \mu_2 \mid P - P_0 \end{cases}$, soit à $\mu_1\mu_2 \mid P - P_0$, donc à $P \equiv P_0\ [\mu_1\mu_2]$. Voilà qui amorce la récurrence, je laisse le lecteur courageux poursuivre ces chinoiseries.

Soit donc P un polynôme vérifiant **(*)** : on a alors $P(D)^2 = D$ et, puisque $A = SDS^{-1}$ avec S inversible,

$$P(A)^2 = P(SDS^{-1})^2 = \big(S\,P(D)\,S^{-1}\big)^2 = S\,P(D)^2\,S^{-1} = SDS^{-1} = A \; .$$

2. • Si A est le produit de deux involutions ($A = UV$ avec $U^2 = V^2 = I$), alors A est inversible et

$$A^{-1} = V^{-1}U^{-1} = VU = V(UV)V^{-1} = VAV^{-1} \; ,$$

donc A et A^{-1} sont semblables.

• Si A est inversible et si $A^{-1} = P^{-1}AP$ avec P inversible, alors $P = APA$ puis

$$P^2A = (APA)^2A = APA^2PA^2 = APA(APA)A = APAPA$$

et

$$AP^2 = A(APA)^2 = A^2PA^2PA = A(APA)APA = APAPA \; ,$$

donc A et P^2 commutent. D'après la question **1.**, il existe un polynôme $F \in \mathbb{C}[X]$ tel que $F(P^2)^2 = P^2$. Posons $Q = F(P^2)$, ainsi $Q^2 = P^2$.

La matrice Q est un polynôme en P^2, donc un polynôme en P ; elle commute donc avec P et avec P^{-1}. En posant $U = QP^{-1}$, on a alors

$$U^2 = (QP^{-1})^2 = QP^{-1}QP^{-1} = Q^2(P^{-1})^2 = Q^2(P^2)^{-1} = I \ :$$

U est une involution.

Posons enfin $V = U^{-1}A = PQ^{-1}A$; ainsi, $A = UV$, il reste à prouver que V est une involution :

$$\begin{aligned} V^2 &= PQ^{-1}APQ^{-1}A = QP^{-1}APQ^{-1}A \qquad \textit{(*)} \\ &= P^{-1}QAQ^{-1}PA \qquad \textit{(**)} \\ &= P^{-1}APA \qquad \textit{(***)} \\ &= A^{-1}A = I \qquad \textit{cqfd} . \end{aligned}$$

(*) : car $PQ^{-1} = (QP^{-1})^{-1} = U^{-1} = U = QP^{-1}$;

(**) : car P et Q commutent ;

(***) : car A commute avec P^2, donc aussi avec Q qui est un polynôme en P^2.

EXERCICE 4 :

1. Soit $N \in \mathcal{M}_n(\mathbb{C})$ une matrice nilpotente. Montrer l'existence d'une matrice $M \in \mathcal{M}_n(\mathbb{C})$ telle que $\exp(M) = I_n + N$.

2. Montrer que l'application $\exp : \mathcal{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ est surjective.

- -

1. Soit r l'indice de nilpotence de N ($N^{r-1} \neq 0$ et $N^r = 0$). Soit le polynôme P, partie régulière du développement limité à l'ordre $r-1$ de la fonction $f : x \mapsto \ln(1+x)$ en zéro :

$$P(X) = X - \frac{X^2}{2} + \ldots + (-1)^r \frac{X^{r-1}}{r-1} = \sum_{k=1}^{r-1} (-1)^{k+1} \frac{X^k}{k} .$$

Soit le polynôme Q, partie régulière du développement limité à l'ordre $r-1$ de la fonction $g : x \mapsto e^x$ en zéro :

$$Q(X) = 1 + X + \frac{X^2}{2!} + \ldots + \frac{X^{r-1}}{(r-1)!} = \sum_{k=0}^{r-1} \frac{X^k}{k!} .$$

La troncature à l'ordre $r-1$ du polynôme composé $Q \circ P$ est la partie régulière du développement limité à l'ordre $r-1$ en zéro de la fonction composée $g \circ f : x \mapsto 1+x$ (*cours de MPSI sur les développements limités*), le polynôme $(Q \circ P)(X) - (1+X)$ a donc une valuation au moins égale à r :

$$(Q \circ P)(X) = 1 + X + X^r\, R(X) \ , \quad \text{avec} \quad R \in \mathbb{C}[X] .$$

Ainsi, $(Q \circ P)(N) = Q\big(P(N)\big) = I_n + N$ puisque $N^r = 0$. Mais, le polynôme P étant de valuation un, on a $P(N) = NA = AN$, où A est un polynôme en N : $A = \displaystyle\sum_{k=1}^{r-1} \frac{(-1)^{k+1}}{k} N^{k-1}$. Donc $\big(P(N)\big)^r = N^r A^r$ (puisque A et N commutent), soit $\big(P(N)\big)^r = 0$. Finalement,

$$\exp\big(P(N)\big) = \sum_{k=0}^{\infty} \frac{\big(P(N)\big)^k}{k!} = \sum_{k=0}^{r-1} \frac{\big(P(N)\big)^k}{k!} = Q\big(P(N)\big) = I_n + N \ .$$

2. Toute matrice de la forme $\lambda I_n + N$ avec $\lambda \in \mathbb{C}^*$ et N nilpotente, admet “un logarithme” : en effet, $\lambda I_n + N = \lambda(I_n + N')$ avec $N' = \dfrac{1}{\lambda} N$ nilpotente. Si $M' \in \mathcal{M}_n(\mathbb{C})$ vérifie $\exp(M') = I_n + N'$ et si α est un nombre complexe tel que $e^\alpha = \lambda$, alors la matrice $M = \alpha I_n + M'$ vérifie $\exp(M) = \lambda I_n + N$.

Si $A \in \mathrm{GL}_n(\mathbb{C})$, on peut trouver une matrice inversible $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $J = P^{-1}AP$ soit diagonale par blocs, de la forme

$$J = \mathrm{diag}(\lambda_1 I_{n_1} + N_1, \ldots, \lambda_p I_{n_p} + N_p) \ , \quad \text{avec}$$

λ_1, ..., λ_p nombres complexes non nuls ;
n_1, ..., n_p entiers naturels non nuls tels que $n_1 + \ldots + n_p = n$;
$N_i \in \mathcal{M}_{n_i}(\mathbb{C})$ nilpotente
(décomposition suivant les sous-espaces caractéristiques, *cf.* détails dans l’exercice **3.**, question **1.b.**).

Pour tout $i \in [\![1, p]\!]$, il existe une matrice $M_i \in \mathcal{M}_{n_i}(\mathbb{C})$ telle que $\exp(M_i) = \lambda_i I_{n_i} + N_i$. Soit la matrice diagonale par blocs

$$M = \mathrm{diag}(M_1, \ldots, M_p) \in \mathcal{M}_n(\mathbb{C}) \ .$$

On a $\exp(M) = J$, puis $\exp(PMP^{-1}) = PJP^{-1} = A$.

EXERCICE 5 :

Soit $\mathbb{K}$ un corps de caractéristique nulle. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice, on note

$$\chi_A(X) = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$$

son polynôme caractéristique. La matrice $C(X) = {}^t\mathrm{Com}(A - XI_n)$ peut être considérée comme une matrice à coefficients dans $\mathbb{K}_{n-1}[X]$ (le justifier) et peut aussi s'écrire comme "polynôme à coefficients matriciels" :

$$C(X) = {}^t\mathrm{Com}(A - XI_n) = C_{n-1} + XC_{n-2} + \cdots + X^{n-2}C_1 + X^{n-1}C_0 = \sum_{k=0}^{n-1} X^k C_{n-1-k}\ .$$

1. Montrer que, pour tout $k \in [\![0, n-1]\!]$, on a $\mathrm{tr}(C_{n-1-k}) = -(k+1)a_{k+1}$.

2. Expliciter C_0. Exprimer C_k en fonction de C_{k-1} pour $k \in [\![1, n-1]\!]$. En déduire un algorithme de calcul des coefficients du polynôme caractéristique.

Source : solution empruntée à Yvan GOZARD, dans la RMS (Revue de Mathématiques Spéciales) 9/10 de mai-juin 1994.

- -

Les coefficients de la matrice des cofacteurs de $A - XI_n$, ou de sa transposée $C(X)$, sont des déterminants de matrices carrées d'ordre $n-1$ extraites de $A - XI_n$, ce sont donc des polynômes de degré inférieur ou égal à $n-1$.

1. On a $\chi_A(X) = \det(A - XI_n)$. Si l'on note $\Gamma_j(X)$ le j-ième "vecteur-colonne" de la matrice $A - XI_n$, les règles de dérivation d'un déterminant donnent

$$\begin{aligned}
\chi_A'(X) &= \sum_{j=1}^{n} \det\big(\Gamma_1(X), \cdots, \Gamma_{j-1}(X), \Gamma_j'(X), \Gamma_{j+1}(X), \cdots, \Gamma_n(X)\big) \\
&= -\sum_{j=1}^{n} M_{jj}(X) = -\,\mathrm{tr}\,\big(C(X)\big) = -\,\mathrm{tr}\left(\sum_{k=0}^{n-1} X^k C_{n-1-k}\right)
\end{aligned}$$

en notant $M_{jj}(X)$ le mineur d'indices (j,j) de la matrice $A - XI_n$, qui est le coefficient d'indices (j,j) de la matrice $C(X)$.

On a donc $\chi_A'(X) = -\sum_{k=0}^{n-1} \big(\mathrm{tr}\, C_{n-1-k}\big) X^k$. En identifiant avec $\chi_A'(X) = \sum_{k=0}^{n-1} (k+1) a_{k+1} X^k$, on obtient

$$\forall k \in [\![0, n-1]\!] \qquad \mathrm{tr}(C_{n-1-k}) = -(k+1)\, a_{k+1}\ .$$

2. On a $(A - XI_n)\, C(X) = \chi_A(X)\, I_n$, soit

$$(A - XI_n)\left(\sum_{k=0}^{n-1} X^k C_{n-1-k}\right) = \sum_{k=0}^{n} a_k X^k I_n\ .$$

En identifiant les "coefficients" (matriciels) dans cette identité polynomiale, on obtient les relations

$$\begin{array}{lrcll} (\mathbf{1}): & A\,C_{n-1} & = & a_0\,I_n & \\ (\mathbf{2}): & A\,C_{n-k-1}-C_{n-k} & = & a_k\,I_n & \qquad (1\leq k\leq n-1) \\ (\mathbf{3}): & -C_0 & = & a_n\,I_n & \end{array}$$

On a ainsi, d'après **(3)**, $C_0=-a_n\,I_n=(-1)^{n-1}I_n$.

Pour $k\in [\![1,n-1]\!]$, la relation **(2)** donne $C_{n-k}=A\,C_{n-k-1}+\dfrac{1}{k}\,\mathrm{tr}(C_{n-k})\,I_n$ grâce à la relation obtenue à la question **1.** et aussi $\mathrm{tr}(A\,C_{n-k-1})-\mathrm{tr}(C_{n-k})=n\,a_k=-\dfrac{n}{k}\,\mathrm{tr}(C_{n-k})$, on en déduit $\mathrm{tr}(C_{n-k})=\dfrac{k}{k-n}\,\mathrm{tr}(A\,C_{n-k-1})$, puis enfin

$$C_{n-k}=A\,C_{n-k-1}-\frac{1}{n-k}\,\mathrm{tr}(A\,C_{n-k-1})\,I_n\ .$$

En résumé, les matrices C_k $(0\leq k\leq n-1)$ peuvent être calculées de proche en proche par les relations

$$\begin{cases} C_0\ =\ (-1)^{n-1}I_n \\ C_k\ =\ A\,C_{k-1}-\dfrac{1}{k}\,\mathrm{tr}(A\,C_{k-1})\,I_n \qquad \text{pour}\quad 1\leq k\leq n-1\ . \end{cases}$$

On en déduit les coefficients du polynôme caractéristique puisque $a_k=-\dfrac{1}{k}\,\mathrm{tr}(C_{n-k})$ si $1\leq k\leq n-1$, et $a_0=\dfrac{1}{n}\,\mathrm{tr}(A\,C_{n-1})$ d'après la relation **(1)** non encore exploitée.

Il s'agit de la **méthode de Faddeev** *qui donne lieu, pour des "grosses" matrices, à des calculs numériques plus rapides que le calcul du polynôme caractéristique comme déterminant.*

Une autre façon de retrouver cet algorithme de calcul des coefficients du polynôme caractéristique est d'utiliser les formules de Newton, qui permettent de relier les susdits coefficients (fonctions symétriques élémentaires des valeurs propres λ_i de A si on se place dans une clôture algébrique de $\mathbb{K}$) aux nombres $\displaystyle\sum_{i=1}^{n}\lambda_i^k=\mathrm{tr}(A^k)$, voir par exemple Jean-Marie ARNAUDIÈS et Henri FRAYSSE, Tome 1, Algèbre, exercice XV.4.7, ISBN 2-04-016450-2.