Annexes

PIANKO Yanis

Concentration

Table des matières

I Annexe : Décomposition de Dunford

Rappelons rapidement deux « visions » ou interprétation de la décomposition de Dunford additive, une à travers les matrices et l'autre à travers les endomorphismes. Soit $A \in \mathcal{M}_p(\mathbb{K})$ et u l'endomorphisme canoniquement associé. On suppose que le polynôme caractéristique $\chi_u\chi_A = \prod_{i=1}^{n}(X-\lambda_i)^{\alpha_i}$ de A est scindé.

Définition

On appelle sous-espace caractéristique de A [resp. u] associé à la valeur propre λ_i le sous-espace vectoriel $\operatorname{Ker}(A-\lambda_i I_p)^{\alpha_i}$ [resp. $\operatorname{Ker}(u-\lambda_i I)^{\alpha_i}$], noté F_{λ_i}.

On note p_i la dimension de F_{λ_i}. Attention, $1 \leqslant p_i \leqslant \alpha_i$ mais on n'a pas nécessairement $\alpha_i = p_i$.

Vision matricielle

Proposition

D'après le théorème de décomposition des noyaux, et le théorème de Cayley-Hamilton,

$$\oplus_{i=1}^{n} \operatorname{Ker}(A-\lambda_i I_p)^{\alpha_i} = \mathbb{K}^p$$

Dans une base adaptée à cette décomposition, la matrice de u (semblable à A, on note P la matrice de passage) est donc :

$$A' = \begin{pmatrix} A'_1 & & \\ & \ddots & \\ & & A'_n \end{pmatrix}$$

Chaque bloc A'_i est trigonalisable, et son spectre est $\{\lambda_i\}$. Ainsi, dans une autre base adaptée bien choisie (on note Q la matrice de passage), la matrice de u est :

$$A'' = \begin{pmatrix} \lambda_1 I_{p_1} + T_1 & & \\ & \ddots & \\ & & \lambda_n I_{p_n} + T_n \end{pmatrix}$$

où pour tout $i \in [\![1,n]\!], T_i$ triangulaire supérieure stricte. Ainsi,

$$A = \underbrace{Q^{-1}\begin{pmatrix} \lambda_1 I_{p_1} & & \\ & \ddots & \\ & & \lambda_n I_{p_n} \end{pmatrix} Q}_{D} + \underbrace{Q^{-1}\begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_n \end{pmatrix} Q}_{N}$$

et D est diagonalisable, N nilpotente.

Vision d'endomorphismes

Proposition

D'après le théorème de décomposition des noyaux, et le théorème de Cayley-Hamilton,

$$\oplus_{i=1}^{n} \operatorname{Ker}(u-\lambda_i I)^{\alpha_i} = E$$

On pose $u_i = u_{|F_{\lambda_i}} \in \mathcal{L}(F_{\lambda_i})$. Alors, $\forall x \in F_{\lambda_i}$, par définition,

$$(u(x) - \lambda_i x)^{\alpha_i} = (u_i(x) - \lambda_i x)^{\alpha_i} = 0$$

Ainsi, l'endomorphisme $u_i - \lambda_i I$ est nilpotent. On peut le noter n_i, et alors :

$$\forall i \in [\![1, n]\!], u_i = \lambda_i I + n_i$$

Une application linéaire est entièrement déterminée par ses restrictions à des sous-espaces vectoriels supplémentaires. En posant d et n les applications telles que leurs restrictions respectives sur chaque F_{λ_i} soit $\lambda_i I$ et n_i, on a alors :

$$d + n = u$$

L'existence de cette décomposition est au programme. L'unicité ne l'est pas mais se fait avec des notions au programme, donc à connaître.

Proposition

En posant π_i la projection canonique sur F_{λ_i}, on a alors $d = \sum_{i=1}^{n} \lambda_i \pi_i$, ce qui permettra de justifier par le lemme des noyaux que d (et donc $n = u - d$) est un polynôme en u.

De là découlent assez naturellement la commutativité de d et n, puis l'unicité :

$$n_1 + d_1 = n_2 + d_2$$
$$n_1 - n_2 = d_2 - d_1$$

L'argument important est que n_1 et n_2 commutent, ainsi que d_2 et d_1, car polynômes en u. Ainsi, $n_1 - n_2$ est nilpotent et $d_2 - d_1$ est diagonalisable.
La seule matrice diagonalisable est la matrice nulle.

II Annexe : Algorithme de Décomposition

Cette annexe présente l'origine de l'algorithme de Décomposition de Dunford utilisé, en l'exploitant plus en détail.

Soit $u \in \mathcal{L}(E)$. On suppose que $\mathbb{K}$ est algébriquement clos, ce qui nous assure que le polynôme caractéristique P_u de u est scindé, soit $P_u(X) = \prod_{k=1}^{n} (X - \lambda_k)^{\alpha_k}$, avec $\alpha_k \in \mathbb{N}^*$ et les $\lambda_k \in \mathbb{K}$ deux à deux distincts. Le polynôme de u s'écrit alors $\pi_u(X) = \prod_{k=1}^{n} (X - \lambda_k)^{\beta_k}$, avec $1 \leqslant \beta_k \leqslant \alpha_k$.

On note $u = d + v$ la décomposition de Dunford de u avec d polynôme en u diagonalisable qui commute avec v nilpotente.

Le polynôme minimal de d est $\pi_d = \prod_{k=1}^{n} (X - \lambda_k) = P$. Il est en effet scindé à racines simples puisque d est diagonalisable, et son polynôme caractéristique est $P_d = P_u$. En effet, N_k est l'espace propre associé à la valeur propre λ_k de d et $\dim(N_k) = \alpha_k$ pour tout k compris entre 1 et n.

L'idée consiste à déterminer d comme solution de $P(w) = 0$ où l'inconnue w est dans $\mathbb{K}[u] \subset \mathcal{L}(E)$. **On va s'inspirer de la méthode de Newton pour résoudre cette équation**, ce qui nous conduit à envisager une suite $(w_k)_{k\in\mathbb{N}}$ d'endomorphisme de E défini par :

$$\begin{cases} w_0 = u \\ \forall k \in \mathbb{N}, w_{k+1} = w_k - P(w_k)\left(P'(w_k)\right)^{-1} \end{cases}$$

où $P'(X) = \sum_{k=1}^{n} \prod_{\substack{j=1 \\ j\neq k}}^{n} (X - \lambda_j)$ est le polynôme dérivé de P.

— Dans le cas où $n = 1$, d est alors diagonalisable avec une seule valeur propre, c'est donc $d = \lambda Id$. En effet, on a $P(X) = X - \lambda, P'(X) = 1$ et $w_1 = u - (u - \lambda Id)\, Id = \lambda I$, puis $w_k = \lambda Id$ pour tout $k \geqslant 1$ par récurrence.
Pour ce qui suit, on suppose $n \geqslant 2$, et **il nous faut justifier l'existence de tels endomorphismes $w_k \in \mathbb{K}[u]$ tels que $P'(w_k)$ soit inversible d'inverse dans $\mathbb{K}[u]$**. On commence par $w_0 = u$.

— Nous allons montrer que l'endomorphisme $P'(u)$ est inversible et son inverse est dans $\mathbb{K}[u]$.
Comme le polynôme P est scindé à racines simples, les polynômes P et P' n'ont pas de racines communes, donc π_u et P' n'ont pas de racines communes (les racines de π_u sont celles de P) et en conséquence sont premiers entre eux. Le théorème de Bézout nous dit alors qu'il existe deux polynômes A et B dans $\mathbb{K}[X]$ tels que $A\pi_u + BP' = 1$, ce qui nous donne en évaluant en u :

$$A(u)\pi_u(u) + B(u)P'(u) = Id \Rightarrow B(u)P'(u) = Id$$

L'endomorphisme $P'(u)$ est donc inversible d'inverse $B(u) \in \mathbb{K}[u]$.

— Nous allons montrer que l'endomorphisme $P(u)$ est nilpotent.
Pour $m = \max_{1\leqslant i\leqslant n} \beta_i$, le polynôme $P^m(X) = \prod_{k=1}^{n} (X - \lambda_k)^m$ est un multiple de $\pi_u(X) = \prod_{k=1}^{n} (X - \lambda_k)^{\beta_k}$, c'est donc un polynôme annulateur de u. On a donc $(P(u))^m = P^m(u) = 0$ et $P(u)$ est nilpotent.

— Soient a et b deux endomorphismes de E qui commutent et tels que a soit inversible et b nilpotent. Nous allons montrer que l'endomorphisme $a-b$ est inversible.
En écrivant que $a-b=a\left(Id-a^{-1}b\right)$, il suffit de montrer que $Id-a^{-1}b$ est inversible. Comme b commute avec a, il commute aussi avec a^{-1} ($ab=ba\Rightarrow a^{-1}aba^{-1}=a^{-1}baa^{-1}$) et si de plus b est nilpotent, il en est de même alors de $a^{-1}b$ puisque $\left(a^{-1}b\right)^r=\left(a^{-1}\right)^r b^r$ pour tout entier naturel r.
En désignant par $r\geqslant 1$ l'indice de nilpotence de $\left(a^{-1}b\right)$, on a :

$$\left(Id-a^{-1}b\right)\sum_{k=0}^{r-1}\left(a^{-1}b\right)^k=Id-\left(a^{-1}b\right)^r=Id$$

ce qui signifie que $Id-a^{-1}b$ est inversible d'inverse $\displaystyle\sum_{k=1}^{r-1}\left(a^{-1}b\right)^r$.

— Nous allons montrer que l'on peut construire une suite $(w_k)_{k\in\mathbb{N}}$ d'endomorphismes de E telle que $w_0=u$ et :

$$\forall k\in\mathbb{K},\begin{cases} w_k\in\mathbb{K}[u]\\ P'(w_k)\in GL(E),\left(P'(w_k)\right)^{-1}\in\mathbb{K}[u]\\ P(w_k)\text{ est nilpotent}\\ w_{k+1}=w_k-P(w_k)\left(P'(w_k)\right)^{-1}\end{cases}$$

On procède par récurrence sur $k\geqslant 0$. Pour $k=0$, on dispose de $w_0=u\in\mathbb{K}[u]$ et on vient de voir que $P'(u)\in GL(E)$ avec $\left(P'(u)\right)^{-1}=B(u)\in\mathbb{K}[u]$ et $P(u)$ est nilpotent. On peut donc poser :

$$\begin{aligned} w_1&=w_0-P(w_0)\left(P'(w_0)\right)^{-1}=u-P(u)\left(P'(u)\right)^{-1}\\ &=u-P(u)B(u)\in\mathbb{K}[u]\end{aligned}$$

Supposons les endomorphismes $w_0,\ldots,w_k$ construits. On peut alors poser :

$$w_{k+1}=w_k-P(w_k)\left(P'(w_k)\right)^{-1}$$

et $w_{k+1}\in\mathbb{K}[u]$.
Comme $n\geqslant 2$, en utilisant la formule de Taylor pour les polynômes :

$$P'(Y)-P'(X)=\sum_{j=1}^{n-1}\frac{1}{j!}P^{(j+1)}(X)(Y-X)^j=(Y-X)\,Q(X,Y)$$

On déduit qu'il existe deux polynômes R_k et S_k dans $\mathbb{K}[X]$ tels que :

$$\begin{aligned} P'(w_{k+1})-P'(w_k)&=(w_{k+1}-w_k)\,R_k(u)=-P(w_k)\left(P'(w_k)\right)^{-1}R_k(u)\\ &=P(w_k)S_k(u)\end{aligned}$$

Comme $P(w_k)$ est nilpotent et polynômial en u, il commute avec $S_k(u)$, donc $P'(w_{k+1})-P'(w_k)$ est aussi nilpotent et :

$$P'(w_{k+1})=P'(w_k)+\left(P'(w_{k+1}-P'(w_k))\right)=a_k-b_k$$

est inversible puisque $a_k=P'(w_k)$ est inversible et commute avec $b_k=P'(w_k)-P'(w_{k+1})$ qui est nilpotent. L'inverse de $P'(w_{k+1})$ est alors :

$$\begin{aligned}\left(P'_{(k+1)}\right)^{-1}&=\left(a_k\left(Id-a_k^{-1}\right)b_k\right)^{-1}=\left(Id-a_k^{-1}b_k\right)^{-1}a_k^{-1}\\ &=\left(\sum_{j=0}^{r_k-1}\left(a_k^{-1}b_k\right)^j\right)a_k^{-1}\end{aligned}$$

Cet inverse est polynomial en u puisque a_k^{-1} et b_k sont dans $\mathbb{K}[u]$.

Il reste enfin à vérifier que $P(w_{k+1})$ est nilpotent. En utilisant la formule de Taylor pour les polynômes :

$$\begin{aligned} P(Y) &= P(X) + (Y-X)P'(X) + (Y-X)^2 \sum_{j=2}^{n} \frac{1}{j!} P^{(j)}(X)(Y-X)^{j-2} \\ &= P(X) + (Y-X)P'(X) + (Y-X)^2 Q(X,Y) \end{aligned}$$

(On a supposé $n \geqslant 2$), on déduit qu'il existe un polynôme R_k dans $\mathbb{K}[X]$ tels que :

$$P(w_{k+1}) = P(w_k) + (w_{k+1} - w_k)P'(w_k) + (w_{k+1} - w_k)^2 R_k(u)$$

avec :

$$P(w_k) + (w_{k+1} - w_k)P'(w_k) = P(w_k) - P(w_k)\left(P'(w_k)\right)^{-1} P'(w_k) = 0$$

ce qui nous donne :

$$P(w_{k+1}) = \left(P(w_k)\left(P'(w_k)\right)^{-1}\right)^2 R_k(u) = P(w_k)S_k(u) \quad (*)$$

où $S_k \in \mathbb{K}[X]$ et $P(w_{k+1})$ est nilpotent comme $P(w_k)$ (qui commute avec $S_k(u)$). La suite $(w_k)_{k\in\mathbb{N}}$ est donc bien définie.

— Nous allons montrer que la suite $(w_k)_{k\in\mathbb{N}}$ est stationnaire sur d à partir d'un certain rang, ce qui permettra de répondre à la question posée.
Il s'agit tout d'abord de montrer que $P(w_k) = 0$ à partir d'un rang. De $(*)$, on déduit que pour tout entier k, l'endomorphisme $P(w_k)$ est multiple de $(P(u))^{2^k}$ dans $\mathbb{K}[u]$. En effet, c'est vrai pour $k = 0$, $(P(w_0) = P(u))$ et en supposant le résultat acquis pour $k \geqslant 0$, soit $P(w_k) = (P(u))^{2^k} Q_k(u)$ avec $Q_k \in \mathbb{K}[X]$, on a dans $\mathbb{K}[u]$ qui est commutatif :

$$\begin{aligned} P(w_{k+1}) &= (P(w_k))^2 \left(\left(P'(w_k)\right)^{-1}\right)^2 R_k(u) \\ &= \left((P(u))^{2^k} Q_k(u)\right)^2 \left(\left(P'(w_k)\right)^{-1}\right)^2 R_k(u) = (P(u))^{2^{k+1}} Q_{k+1}(u) \end{aligned}$$

Et comme $P(u)$ est nilpotent, on aura $P(w_k) = 0$ pour k assez grand, ce qui équivaut à dire que $w_{k+1} = w_k$ et $w_{k+p} = w_k$ pour tout $p \geqslant 0$.

Soit $k_0 \geqslant 0$ le plus petit entier tel que $P(w_{k_0}) = 0$. Comme w_{k_0} est annulé par le polynôme P qui est scindé à racines simples, cet endomorphisme est diagonalisable.
Si $k_0 = 0$, l'endomorphisme $u = w_0$ est diagonalisable, donc $d = u$ et $n = 0$.
Si $k_0 \geqslant 1$, en écrivant que :

$$\begin{aligned} u - w_{k_0} = w_0 - w_{k_0} &= \sum_{j=0}^{k_0-1} (w_j - w_{j+1}) \\ &= \sum_{j=0}^{K_0-1} P(w_j)\left(P'(w_j)\right)^{-1} = \sum_{j=0}^{k_0-1} (P(u))^{2^j} Q_j(u)\left(P'(w_j)\right)^{-1} \end{aligned}$$

On voit que $u - w_{k_0}$ est multiple de $P(u)$ dans $\mathbb{K}[u]$, donc nilpotent.
On a donc une décomposition $u = w_{k_0} + (u - w_{k_0})$ avec w_{k_0} qui commute avec $u - w_{k_0}$, qui est nilpotent, c'est donc la décomposition de Dunford de u. En particulier, $w_{k_0} = u$.

Le calcul des w_k peut se faire sans connaître les valeurs propres de u. En effet, en remarquant que $P_u \wedge P_u' = \prod_{k=1}^{n} (X-\lambda_k)^{\alpha_k - 1}$, *on déduit que :*

$$P(X) = \prod_{k=1}^{n} (X-\lambda_k) = \frac{P_u(X)}{P_u \wedge P_u'}$$

le pgcd de P_u et P_u' pouvant se calculer en utilisant l'algorithme d'Euclide (divisions euclidiennes). De même pour l'inverse $B(u)$ de $P'(u)$, on utilise l'algorithme d'Euclide dans la relation de Bézout $AP_u + BP = 1$

⋆ ⋆ ⋆

III Annexe : Rappels et Compléments sur l'exponentielle matricielle

Proposition

L'exponentielle matricielle est une fonction $\mathcal{C}^1(\mathcal{M}_n(\mathbb{C}), \mathcal{M}_n(\mathbb{C}))$.

L'idée est de copier la démonstration du théorème de dérivation terme à terme des séries entières d'une variable réelle ou complexe, en faisant attention au fait que l'anneau $\mathcal{M}_p(\mathbb{K})$ n'est pas commutatif.

Pour $k \geqslant 1$, l'application $\varphi_k : X \to X^k$ est de classe $\mathcal{C}^\infty$ sur $\mathcal{M}_p(\mathbb{K})$ (ses composantes sont polynômiales).
On vérifie facilement par récurrence sur $k \geqslant 1$ que :

$$\begin{aligned}\varphi_k(X+H) = (X+H)^k &= X^k + X^{k-1}H + X^{k-2}HX^2 + \ldots + HX^{k-1} + o(\|H\|) \\ &= X^k + \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=k-1}} X^i H X^j + \|H\|\varepsilon(H)\end{aligned}$$

où $\lim\limits_{H\to 0} \varepsilon(H) = 0$.
En effet, pour $k = 1$, c'est vrai, et en supposant le résultat acquis pour $k \geqslant 1$, on a :

$$\begin{aligned}(X+H)^{k+1} &= \left(X^k + \sum_{\substack{0\leqslant i,j\leqslant k_1 \\ i+j=k-1}} X^i H X^j + \|H\|\varepsilon(H)\right)(X+H) \\ &= X^{k+1} + \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=k-1}} X^i H X^{j+1} + X^k H + \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=k-1}} X^i H X^{j+1} H + \|H\|\varepsilon(H)(X+H) \\ &= X^{k+1} + \sum_{\substack{0\leqslant i,j\leqslant k \\ i+j=k}} X^i H X^j + \|H\|\varepsilon_1(H)\end{aligned}$$

avec :

$$\begin{aligned}\|\varepsilon_1(H)\| &= \left\| \frac{1}{\|H\|} \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=K-1}} X^i H X^{j+1} H + \varepsilon(H)(X+H) \right\| \\ &\leqslant k\|X\|^k\|H\| + \|\varepsilon(H)\|(\|X\| + \|H\|) \xrightarrow[H\to 0]{} 0\end{aligned}$$

La différentielle de φ_k en X est donc définie par :

$$\forall H \in \mathcal{M}_p(\mathbb{K}), \mathrm{d}\varphi_k(X)(H) = \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=k-1}} X^i H X^j$$

(Pour $k = 0$, on a $\mathrm{d}\varphi_0(X) = 0$).

Avec $\|\mathrm{d}\varphi_k(X)(H)\| \leqslant k\|X\|^{k-1}\|H\|$, on déduit que $\frac{1}{k!}\|\mathrm{d}\varphi_k(X)\| \leqslant \frac{\|X\|^{k-1}}{(k-1)!}$ et la série $\sum \frac{1}{k!}\mathrm{d}\varphi_k(X)$ est uniformément convergente sur tout compact de l'espace de Banach $\mathcal{L}(\mathcal{M}_p(\mathbb{K}))$. Il en résulte que exp est de classe $\mathcal{C}^1$ sur $\mathcal{M}_p(\mathbb{K})$ avec :

$$\mathrm{d}(\exp)(X)(H) = \sum_{k=1}^{+\infty} \frac{1}{k!}\mathrm{d}\varphi_k(X)(H) = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{\substack{0\leqslant i,j\leqslant k-1 \\ i+j=k-1}} X^i H X^j$$

Proposition

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, la matrice e^A est inversible d'inverse e^{-A}.

La fonction $\psi : t \to e^{tA}e^{-tA}$ est dérivable sur $\mathbb{R}$ de dérivée :

$$\psi'(t) = Ae^{tA}e^{-tA} + e^{tA}(-A)e^{-tA} = (A - A)\psi(t) = 0$$

En effet, A commute avec e^{tA}, ce qui entraîne que ψ est constante, et donc $\psi(t) = \psi(0) = I_n$ pour tout t.

Remarque. On déduit de ce que l'on vient de démontrer que pour tout matrice $A \in \mathcal{M}_n(\mathbb{K})$, les solutions du système différentiel $Y' = AY$, où $Y \in \mathcal{C}^1(\mathbb{R}, \mathbb{K}^n)$ sont les fonctions $Y : t \to e^{tA}Y_0$, où $Y_0 \in \mathbb{K}^n$. En effet ces fonctions sont solutions et pour toutes fonction $Y \in \mathcal{C}^1(\mathbb{R}, \mathbb{K}^n)$ solution de $Y' = AY$, en notant $Z(t) = e^{-tA}Y(t)$, on a :

$$Z'(t) = -e^{-tA}AY(t) + e^{-tA}Y'(t) = e^{-tA}(-AY(t) + Y'(t)) = 0$$

Donc $Z(t) = Y_0$ et $Y(t) = e^{tA}Y_0$.

Proposition

Soient A, B dans $\mathcal{M}_n(\mathbb{K})$. Les matrices A et B commutent si et seulement si

$$\forall t \in \mathbb{R}, e^{t(A+B)} = e^A e^B$$

Supposons que A et B commutent. La fonction $\psi : t \to e^{t(A+B)}e^{-tA}e^{-tB}$ est dérivable sur $\mathbb{R}$ et de dérivée :

$$\begin{aligned}\psi'(t) &= (A + B)\, e^{t(A+B)}e^{-tA}e^{-tB} - e^{t(A+B)}Ae^{-tA}e^{-tB} - e^{t(A+B)}e^{-tA}Be^{-tB} \\ &= (A + B - A - B)\psi(t) = 0\end{aligned}$$

Puisque tous les endomorphismes considérés commutent. Ainsi, ψ est constante, soit $\psi(t) = \psi(0) = I_n$ pour tout réel t. L'inverse de e^{-tA} est e^{tA} donc pour tout réel t, on a :

$$e^{t(A+B)} = e^{tA}e^{tB}$$

L'unicité du développement en série entière au voisinage de 0 d'une fonction développable en série entière de $]-r, r[$ dans l'algèbre de Banach $\mathcal{M}_n(\mathbb{K})$ nous donne une démonstration de la condition nécessaire.

Pour A, B fixés dans $\mathcal{M}_n(\mathbb{K})$ et tout réel t on a $e^{t(A+B)} = \sum_{k=0}^{+\infty} \frac{t^k}{k!}(A+B)^k$ et $e^{tA}e^{tB} = \sum_{k=0}^{+\infty} \frac{t^k}{k!}A^k \sum_{k=0}^{+\infty} \frac{t^k}{k!}B^k =$ $I_n + t(A + B) + \frac{t^2}{2}(A^2 + 2AB + B^2) + o(t^2)$. L'égalité est donc réalisée si et seulement si tous les coefficients de ces deux développements en série entière coïncident, ce qui entraîne en particulier $(A + B)^2 = A^2 + 2AB + B^2$, soit $AB = BA$.

Proposition

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que son polynôme caractéristique soit scindée sur $\mathbb{K}$ et $A = D + N$ sa décomposition de Dunford, avec D diagonalisable et N nilpotente d'indice de nilpotence q. On a $e^A = e^D e^N = e^D \sum_{k=0}^{q-1} \frac{1}{k!}N^k$ et la décomposition de Dunford de e^A est donnée par $e^D + e^D\left(e^N - I_n\right)$ avec e^D et $e^D(e^N - I_n)$ nilpotente.

Les arguments importants pour cette démonstration sont que e^D et $\left(e^N - I_n\right)$ commutent, $e^N - I_n$ est une **somme finie de matrices nilpotentes qui commutent** donc nilpotente.

IV Annexe : Raffinement de la surjectivité de l'exponentielle sur $GL_n(\mathbb{C})$

On va montrer, de plusieurs manières différentes, que $\forall A \in \mathcal{M}_n(\mathbb{C}), \exists P \in \mathbb{C}[X], A = \exp(P(A))$. On utilisera des méthodes différentes que celle utilisée dans le sujet.

Exercice de topologie

1) En dimension 1 :
   (a) Montrer que $H = \exp(\mathbb{C})$ est un sous-groupe ouvert de $\mathbb{C}^*$.
   (b) Montrer que H est aussi fermé dans $\mathbb{C}^*$. Conclure que $\exp : \mathbb{C} \longrightarrow \mathbb{C}^*$ est surjective.

2) En dimension quelconque. Soit $A \in \mathcal{M}_n(\mathbb{C})$. On veut montrer qu'il existe $P \in \mathbb{C}[X]$, tel que $\exp(P(A)) = A$.
   (a) L'application exp est-elle un morphisme du groupe additif $\mathcal{M}_n(\mathbb{C})$ sur le groupe multiplicatif $GL_n(\mathbb{C})$?
   (b) On note $\mathbb{C}[A]^* = \mathbb{C}[A] \cap GL_n(\mathbb{C})$. Justifier que $\exp : \mathbb{C}[A] \longrightarrow \mathbb{C}[A]^*$ est bien définie et que $\mathbb{C}[A]^*$ est un ouvert de $\mathbb{C}[A]$.
   (c) Montrer que exp est de classe $\mathcal{C}^1$ sur $\mathcal{M}_n(\mathbb{C})$.
   (d) Montrer que $H = \exp(\mathbb{C}[A])$ est un sous-groupe ouvert de $\mathbb{C}[A]^*$, puis conclure.

Solution :

1) (a) Pour tout $a, b \in \mathbb{C}, \exp(a+b) = \exp(a)\exp(b)$. La fonction $\exp : \mathbb{C} \longrightarrow \mathbb{C}^*$ est donc un morphisme de groupes, par suite $\exp(\mathbb{C})$ est un sous-groupe de $\mathbb{C}^*$.
Montrons que 1 est intérieur à H. La fonction exp est holomorphe de dérivée elle-même, elle est donc de classe $\mathcal{C}^1$ si on la regarde comme une fonction de deux variables réelles, sa différentielle en 0 est la multiplication par $\exp(0) = 1$, c'est donc l'identité sur $\mathbb{R}^2$. D'après le théorème d'inversion locale, exp est donc localement inversible, il existe un voisinage ouvert $\mathcal{V}$ de 1 dont tous les éléments sont des exponentielles, donc $1 \in \mathcal{V} \subset H$, ce qui montre que 1 est intérieur à H.
Si $a \in H$, $a\mathcal{V}$ est un voisinage de a, ouvert car image de $\mathcal{V}$ par l'homéomorphisme $h \to ah$ (sa réciproque est $h \to a^{-1}h$), et inclus dans H car H est stable par multiplication. H est donc ouvert.
   (b) On partitionne $\mathbb{C}^*$ en ses classes modulo H ($x \sim y \iff xy^{-1} \in H$). Chacune de ses classes bH est ouverte comme image de l'ouvert H par $h \to bh$. Le complémentaire de H dans $\mathbb{C}^*$ est donc la réunion des ouverts bH avec $b \notin H$, c'est donc une partie ouverte de $\mathbb{C}^*$, ce qui donne H fermé dans $\mathbb{C}^*$.
Concluons : H est à la fois fermé et ouvert dans $\mathbb{C}^*$, qui est connexe (car c'est un espace vectoriel), et H n'est pas vide, donc $H = \mathbb{C}^*$.

2) (a) Dès que $n \geqslant 2$, comme deux matrices ne commutent pas forcément, on n'a plus $\exp(A+B) = \exp(A)\exp(B)$, l'application $\exp : \mathcal{M}_n(\mathbb{C}) \longrightarrow GL_n(\mathbb{C})$ n'est donc plus un morphisme ! On ne peut donc pas généraliser ainsi la preuve. Il nous faut de la commutativité. On pense alors aux algèbres de polynômes.
   (b) Si $M = P(A)$ où P est un polynôme,

$$\lim_{n\to+\infty} \sum_{k=0}^{n} \frac{(P(A))^k}{k!}$$

Comme $\mathbb{C}[A]$ est une $\mathbb{C}$-algèbre de dimension finie ($\mathbb{C}[A]$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ de dimension finie sur $\mathbb{C}$), c'est donc une partie fermée de $\mathcal{M}_n(\mathbb{C})$. Par suite $\exp(M)$

est un polynôme en A, puisque limite d'une suite de polynômes en A. On sait en plus que $\exp(M)$ est inversible, donc $\exp : \mathbb{C}[A] \longrightarrow \mathbb{C}[A]^*$ est bien définie.
On a $\mathbb{C}[A]^* = \{M \in \mathbb{C}[A], \det M \neq 0\}$, donc c'est une partie ouverte de $\mathbb{C}[A]$ comme image réciproque de l'ouvert $\mathbb{C}^*$ par l'application continue det (car polynomiale).

(c) On a montré que exp était différentiable dans le sujet.

(d) À partir de là, il n'y a plus qu'à imiter la preuve de la question 1. Comme tous les polynômes en A commutent entre eux, $\exp : \mathbb{C}[A] \longrightarrow \mathbb{C}[A]^*$ est un morphisme de groupes, et $H = exp(\mathbb{C}[A])$ est un sous-groupe de $C[A]^*$. La fonction exp est de classe $\mathcal{C}^1$, sa différentielle en 0 est l'identité donc est inversible :

$$\exp\left(H'\right) = I_n + H' + o\left(\left\|H'\right\|^2\right)$$

On en déduit l'existence d'un voisinage ouvert $\mathcal{V}$ tel que $I_n \in \mathcal{V} \subset H$. Pour tout $M \in H, M\mathcal{V}$ est un voisinage ouvert de M dans H, car $B \to MB$ est un homéomorphisme sur $GL_n(\mathbb{C})$ (c'est une application linéaire donc continue car en dimension finie), d'inverse $B \to M^{-1}B$).
Le groupe H est donc un sous-groupe ouvert de $\mathbb{C}[A]^*$, il est donc aussi fermé (même preuve qu'en 1.b).
Si M et N sont dans $\mathbb{C}[A]^*$, la fonction polynomiale non nulle de $\mathbb{C}$ dans $\mathbb{C}$ $z \to \det((1-z)M + zN)$ n'admet qu'un nombre fini de zéros et ne s'annule ni en 0, ni en 1, donc il existe un chemin continu γ joignant 0 à 1 qui évite ces zéros dans $\mathbb{C}$. Alors l'arc paramétré $t \to (1-\gamma(t))M + \gamma(t)N$ joint M et N et est à valeurs dans $\mathbb{C}[A]^*$, ce qui prouve que $\mathbb{C}[A]^*$ est connexe par arcs.
Finalement la partie H est non vide, ouverte et fermée dans le connexe $\mathbb{C}[A]^*$, donc $H = \mathbb{C}[A]^*$, ce qui achève la preuve.

Extension du cas inversible diagonalisable

Remarque : c'est sensiblement la méthode traitée dans le sujet

Lemme

Soit $A \in GL_n(\mathbb{C})$ une matrice diagonalisable. Il existe un polynôme $Q \in \mathbb{C}_{n-1}[X]$ tel que $Q(A)$ soit diagonalisable et $e^{Q(A)} = A$.

Démonstration. Comme A est inversible et diagonalisable, $\exists P \in GL_n(\mathbb{C})$ tel que $A = P\text{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}$ avec $\lambda_i \neq 0 \forall i \in [\![1, n]\!]$.
L'exponentielle est surjective de $\mathbb{C}$ sur $\mathbb{C}^*$, il existe des nombres complexes $\mu_1, \ldots, \mu_n$ tels que $\lambda_k = e^{\mu_k}$ pour tout $k \in [\![1, n]\!]$.
Le théorème d'interpolation de Lagrange nous dit qu'il existe un polynôme $Q \in \mathbb{C}_{n-1}[X]$ tel que $\mu_k = Q(\lambda_k)$ pour tout k compris entre 1 et n (en fait, $Q \in \mathbb{C}_{p-1}[X]$ si on a p valeurs propres distinctes).
Alors, en posant $\Delta = P\text{diag}(\mu_1, \ldots, \mu_n)P^{-1}$, on a :

$$e^{\Delta} = Pe^{\text{diag}(\mu_1,\ldots,\mu_n)}P^{-1} = P\text{diag}(e^{\mu_1,\ldots,e^{\mu_n}})P^{-1} = A$$

$$\text{et} \quad \Delta = \text{diag}(Q(\lambda_1), \ldots, Q(\lambda_n))P^{-1} = Q(A)$$

On peut alors démontrer grâce à la décomposition de Dunford la proposition suivante :

Proposition

Pour toute matrice $A \in GL_n(\mathbb{C})$, il existe un polynôme $Q \in \mathbb{C}[X]$ tel que $e^{Q(A)} = A$ (l'exponentielle matricielle réalise une surjection de $M_n(\mathbb{C})$ sur $GL_n(\mathbb{C})$).

Démonstration. Soit $A \in GL_n(\mathbb{C})$. On a la décomposition de Dunford $A = D + N$ avec D diagonalisable qui commute avec N nilpotente. De plus, on sait que D et N sont des polynômes en A. Comme D a les mêmes valeurs propres que A, elle est inversible et le lemme précédent nous dit qu'il existe $Q_1 \in \mathbb{C}[X]$ tel que $\Delta = Q_1(D)$ soit diagonalisable et $e^{\Delta} = D$. La matrice D étant un polynôme en A, Δ l'est aussi.
Il faut chercher une matrice $X = \Delta + Y$ avec Y nilpotente commutant avec Δ telle que $e^X = A = D+N$. On a nécessairement $D = e^{\Delta}$, $N = e^{\Delta}\left(e^Y - I_n\right)$, soit $e^Y = e^{-\Delta}N + I_n = D^{-1}N + I_n$. Comme N et D commutent, N et D^{-1} commutent, donc $D^{-1}N$ nilpotent et ainsi $I_n + DN^{-1}$. D^{-1} est polynômiale en A (Cayley-Hamilton) et ainsi Y est polynômiale en A. Δ et Y commutent, $\Delta + Y$ polynômial en A.

$$e^{\Delta+Y} = e^{\Delta}e^{Y} = D\left(D^{-1}N + I_n\right) = D + N = A$$

Réduction de Jordan

Le théorème de réduction de Jordan nous dit que toute matrice $A \in GL_n(\mathbb{C})$ est semblable à une matrice diagonale par blocs de la forme :

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix}$$

avec $J_k = \lambda_k I_n + V_k$ pour tout entier k compris entre 1 et p, la matrice V_k étant nilpotente et λ_k étant valeur propre de A. Comme la matrice A est inversible, tous les λ_k sont non nuls et on peut trouver des matrices à coefficients complexes X_k telles que $e^{X_k} = J_k$.

En écrivant que $A = PJP^{-1}$ avec P inversible et en définissant la matrice X par $X = P \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & X_p \end{pmatrix} P^{-1}$,

on a :

$$e^X = P \begin{pmatrix} e^{X_1} & 0 & \cdots & 0 \\ 0 & e^{X_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{X_p} \end{pmatrix} P^{-1} = PJP^{-1} = A$$

⋆ ⋆ ⋆

V Annexe : Ensemble Racine Carré

Soient un entier naturel n non nul et M une matrice carrée d'ordre n à coefficients dans un anneau A. Un élément R de $\mathcal{M}_n(A)$ est une racine carrée de M si $R^2 = M$. On note Rac M l'ensemble des racines carrées de M. Il est possible de dénombrer Rac M s'il est fini, sinon de remarquer qu'il est constitué de classes de similitude.

Proposition

Si R est une racine carrée de M alors R est inversible si et seulement si M l'est.
Si une matrice est inversible, les racines carrées de son inverse sont les inverses de ses racines carrées.

Le théorème spectral nous indique que toute matrice symétrique est orthodiagonalisable dans $\mathcal{M}_n(\mathbb{R})$. On montre dans le sujet qu'elle est positive si et seulement si ses valeurs propres sont des réels positifs ou nuls.

Proposition

Si une matrice S est diagonalisable alors son carré a même sous-espaces propres (associés aux carrés des valeurs propres de S).

Corollaire

Parmi les racines carrées d'une matrice symétrique positive M, une et une seule est symétrique positive : la matrice S qui a mêmes sous-espaces propres que M et dont les valeurs propres associées sont les racines carrées de celles de M. De plus, lorsque M est définie positive, S l'est aussi.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On peut déterminer Rac A dans certains cas particuliers :

Cas 1 : n valeurs propres distinctes

A est diagonalisable, $A = PDP^{-1}$ avec D diagonale. Si $R^2 = A, R$ commute avec A, donc laisse stable les droites propres de A (car A est à spectre simple), R est donc aussi diagonalisable dans la base d'espaces propres de A, d'où $R = PSP^{-1}$ avec S diagonale. On a alors $S^2 = D$, ce qui donne $s_i^2 = d_i$, où (s_i) et (d_i) sont les coefficients diagonaux de S et D. S'il existe un $d_i < 0$, alors il n'y a pas de racine carrée. Sinon, $s_i = \pm\sqrt{d_i}$ pour tout i.
Réciproquement les matrices $P\text{diag}(\pm\sqrt{d_1}, \dots, \pm\sqrt{d_n})P^{-1}$ sont bien des racines carrées de A.

Conclusion :
- si A admet une valeur propre strictement négative, A n'admet pas de racines carrées (réelles).
- si toutes les valeurs propres de A sont strictement positives, A admet $2n$ racines carrées.
- si 0 est valeur propre de A et que ses autres valeurs propres sont positives, alors A admet $2n-1$ racines carrées.

Cas 2 : $A = I_n$

Si $R^2 = I_n, R$ annule le polynôme $X^2 - 1$ scindé à racines simples. R est donc diagonalisable, semblable à $\text{diag}(\varepsilon_1, \dots, \varepsilon_n)$ avec les ε_i valant ± 1.
Réciproquement, si $P \in GL_n(\mathbb{R})$, alors $R = P\text{diag}(\varepsilon_1, \dots, \varepsilon_n)P^{-1}$ vérifie bien $R^2 = I_n$. Donc $\text{Rac}(I_n) = \left\{P\text{diag}(\varepsilon_1, \dots, \varepsilon_n)P^{-1}, P \in GL_n(\mathbb{R})\right\}$. C'est la réunion de $n+1$ classes de similitude (cela correspond au nombre de 1 présents sur la diagonale de $\text{diag}(\varepsilon_1, \dots, \varepsilon_n)$).

Cas 3 : $A = 0$

Si $R \in \text{Rac } A$, alors $R^2 = 0, R$ est nilpotente, donc semblable à une diagonale de blocs de Jordan J_k (voir compléments sur réduction de Jordan, en utilisant Dunford ou alors le complément sur la réduction de Frobenius), avec :

$$J_k = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix} \in \mathcal{M}_k(\mathbb{R})$$

si $k \in \mathbb{N}^*$, et $J_1 = (0)$. Nécessairement la taille des blocs est inférieure ou égale à 2, sinon $R^2 \neq 0$. R est donc semblable à une matrice $R_k = \text{diag}(\underbrace{J_2, \ldots, J_2}_{k \text{ fois}}, 0, \ldots, 0)$. Pour des raisons de taille, il y a au maximum $\text{E}(\frac{n}{2})$ blocs J_2 dans R_k, donc $0 \leqslant k \leqslant \text{E}(\frac{n}{2})$. Réciproquement, toute matrice semblable à R_k vérifie bien $R_k^2 = 0$.

On remarque que $\text{rg}\, R_k = k$, donc si $k \neq k'$, R_k et R'_k ne sont pas semblables. Rac 0 est donc constituée de $\text{E}(\frac{n}{2}) + 1$ classes de similitude.

Cas 4 : $A = -I_n$

Commençons par une remarque : si $n = 2$, on interprète $-I_n$ comme la matrice de rotation d'angle π, on se dit alors que les matrices de rotation d'angle $\dfrac{\pi}{2}$ vont jouer un rôle.

- Si $R^2 = -I_n$, alors $(\det R)^2 = (-1)^n$, donc nécessairement n est pair.
- R annule le polynôme $X^2 + 1$ scindé à racines simples sur $\mathbb{C}$. La matrice R est donc diagonalisable, semblable sur $\mathbb{C}$ à $\text{diag}(\varepsilon_1 i, \ldots, \varepsilon_n i)$ avec les ε_i dans $\{\pm 1\}$. Comme R est réelle, ses valeurs propres complexes sont conjuguées. Parmi les ε_i, il y en a donc autant qui prennent la valeur 1 que la valeur -1. Quitte à conjuguer par une matrice de permutation, on a donc $R \sim \text{diag}(i, -i, \ldots, i, -i)$.
- Maintenant :
$$R_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sim \text{diag}(i, -i)$$
puisque $R_{\frac{\pi}{2}}$ a pour polynôme caractéristique $X^2 + 1 = (X - i)(X + i)$. Par produit de blocs, et par transitivité de la relation de similitude, on en déduit que $R \sim \text{diag}(R_{\frac{\pi}{2}}, \ldots, R_{\frac{\pi}{2}})$.
Attention, pour le moment les 2 matrices sont semblables sur $\mathbb{C}$. Mais deux matrices réelles semblables sur $\mathbb{C}$, le sont aussi sur $\mathbb{R}$. (à redémontrer)
- Réciproquement, si n est pair et $P \in GL_n(\mathbb{R})$, $R = P\text{diag}(R_{\frac{\pi}{2}}, \ldots, R_{\frac{\pi}{2}})P^{-1}$ vérifie bien $R^2 = -I_n$.

Faisons le bilan :
- Si n est impair, Rac $(-I_n)$ est vide.
- Si n est pair, Rac $(-I_n) = \left\{P\text{diag}(R_{\frac{\pi}{2}}, \ldots, R_{\frac{\pi}{2}})P^{-1}, P \in GL_n(\mathbb{R})\right\}$ et l'on obtient une seule classe de similitude.

- Soit R une racine carrée de la matrice A diagonalisable. L'idée est de se ramener aux cas précédents par diagonalisation simultanée possible grâce à la commutation.
Notons $\lambda_1, \ldots, \lambda_k$ les valeurs propres de A et $p_1, \ldots, p_k$ leur multiplicité respective. Notons u et v les endomorphismes de R^n canoniquement associés à A et R. Puisque u et v commutent, v laisse stable les sous-espaces propres de u. Ainsi dans une base de vecteurs propres de u (elle existe puisque u diagonalisable), la matrice de u est diagonale et celle de v est diagonale par blocs. Voici la traduction matricielle : si P désigne la matrice de passage de la base canonique de R^n à la

base de vecteurs propres, on a $A = P\text{diag}(\lambda_1 I_{p_1}, \dots, \lambda_k I_{p_k})P^{-1}$ et $R = P\text{diag}(R_1, \dots, R_k)P^{-1}$ avec $R_i \in \mathcal{M}_{p_i}(\mathbb{K})$.
Comme $A = R^2$, par produit des blocs, on tire que pour tout $i \in [\![1; k]\!]$:

$$\lambda_i I_{p_i} = R_i^2$$

Les matrices R_i sont donc des racines carrées des matrices $\lambda_i I_{p_i}$.
— Si $\lambda_i = 0$, on est ramené à Rac 0.
— Si $\lambda_i > 0$, on est ramené à Rac I_{p_i} car $I_{p_i} = \left(\frac{1}{\sqrt{\lambda_i}} R_i\right)^2$.
— Si $\lambda_i < 0$, on est ramené à Rac$(-I_{p_i})$ car $-I_{p_i} = \left(\frac{1}{\sqrt{-\lambda_i}} R_i\right)^2$.

On peut mener une étude topologique de cet espace, en munissant $\mathcal{M}_n(\mathbb{R})$ d'une norme.

Proposition

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Rac A est une partie fermée d'intérieur vide de $\mathcal{M}_n(\mathbb{R})$.

Démonstration. Rac A est un ensemble algébrique, c'est-à-dire une intersection de zéros de fonctions polynomiales non nulles. À ce titre, il est fermé et d'intérieur vide (c'est une petite partie au sens topologique de Baire).
Une preuve détaillée figure dans l'épreuve d'algèbre du concours CCP section MP de 2005. Voici les grandes lignes de la preuve.
Un ensemble algébrique est fermé comme intersections de fermés (qui sont des images réciproques du fermé $\{0\}$ par une application continue). Un ensemble de zéros d'une fonction polynomiale P non nulle est d'intérieur vide. En effet s'il admet un point intérieur, P s'annule sur une boule ouverte qui est égale (en choisissant la norme infinie) à un produit cartésien d'intervalles ouverts, ce qui implique alors que le polynôme est nul. On conclut ensuite puisqu'une intersection de parties d'intérieur vide est encore d'intérieur vide.

Proposition

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Rac A est une partie non bornée de $\mathcal{M}_n(\mathbb{R})$.

Démonstration. Montrons que pour $n \geqslant 2$, Rac I_n n'est pas bornée. On choisit la norme infinie sur $\mathcal{M}_n(\mathbb{R})$, qui vaut le max des valeurs absolues des coefficients. Pour tout entier p non nul,

$$R_p = \begin{pmatrix} -1 & p \\ 0 & 1 \end{pmatrix}$$

est une racine carrée de I_2, de norme égale à p. Si $n > 2$, la matrice $\text{diag}(R_p, 1, \dots, 1)$ est encore une racine carrée de I_n de norme p, ce qui prouve que pour $n \geqslant 2$, Rac I_n n'est pas bornée.

⋆ ⋆ ⋆

VI Annexe : Algèbre bilinéaire, congruence et Cholesky

Définition

Une forme bilinéaire sur E est une application :

$$\varphi \left| \begin{array}{l} E \times E \to \mathbb{K} \\ (x,y) \mapsto \varphi(x,y) \end{array} \right.$$

telle que, pour tous x, y dans E, les applications $z \to \varphi(x,z)$ et $z \to \varphi(z,y)$ sont linéaires.

Une telle forme bilinéaire est dite symétrique [resp. alternée] si $\varphi(x,y) = \varphi(y,x)$ pour tous x, y dans E [resp. si $\varphi(x,y) = -\varphi(y,x)$ pour tous x, y dans E].

Définition

Soient E un espace vectoriel réel de dimension $n \in \mathbb{N}^*, \mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\varphi : E^2 \longrightarrow \mathbb{R}$ une application bilinéaire. On appelle matrice de φ dans la base $\mathcal{B}$, la matrice A définie par

$$\forall (i,j) \in [\![1,n]\!]^2, A_{i,j} = \varphi(e_i, e_j)$$

A désigne une matrice de $\mathcal{M}_p(\mathbb{R})$.

— Il est possible de montrer qu'une application φ de $E \times E$ dans $\mathbb{K}$ est une forme bilinéaire sur E si, et seulement si, il existe une matrice $B = ((b_{ij}))_{1 \leqslant i,j \leqslant p}$ dans $\mathcal{M}_p(\mathbb{K})$ et des formes linéaires $\ell_1, \ldots, \ell_p$ linéaires linéairement indépandantes telles que :

$$\forall (x,y) \in E \times E, \varphi(x,y) = \sum_{1 \leqslant i,j \leqslant p} b_{ij} \ell_i(x) \ell_j(y)$$

— On peut montrer que, pour tous vecteurs x et y de coordonnées X et Y dans $\mathcal{B}$, $\varphi(x,y) = {}^tXAY$.

Il suffit de le montrer pour des vecteurs quelconques de la base $\mathcal{B}$. Or, ${}^te_iAe_j = A_{i,j} = \varphi(e_i, e_j)$. On conclut par bilinéarité.

— Vérifier que φ est symétrique si et seulement si sa matrice A dans une base quelconque $\mathcal{B}$ de E est symétrique [resp. alternée si, et seulement si, sa matrice A dans une base quelconque $\mathcal{B}$ de E est antisymétrique].

Il suffit de remarquer que $\varphi(e_i, e_j) = A_{i,j}$ et que $\varphi(e_j, e_i) = A_{j,i}$.

— Une matrice symétrique A est la matrice d'un produit scalaire si elle est définie positive, comme la forme linéaire, d'où la notation.

En effet, par définition, A est définie positive si et seulement si $\forall x \in \mathbb{R}^n \backslash \{0\}, \varphi(x,x) = \langle Ax, x \rangle > 0$.

Définition

Une forme quadratique sur E est une application :

$$q \left| \begin{array}{l} E \to \mathbb{K} \\ x \mapsto \varphi(x,x) \end{array} \right.$$

où φ est une forme bilinéaire sur E.

On représente une forme bilinéaire et sa forme quadratique associée par la même matrice.

De la même façon que l'on définit des classes de similitude sur l'espace des matrices pour classer les matrices représentant les mêmes applications linéaires dans des bases différentes (matrices semblables), on peut définir des classes de similitude classifiant les matrices représentant les mêmes formes linéaires dans des bases différentes. Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Si A_1 et A_2 sont les matrices d'une forme bilinéaire φ sur E dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement, montrer qu'on a alors

$$A_2 = {}^tPA_1P$$

On dit alors que A_1 et A_2 sont congruentes.

On peut vérifier que la relation de congruence est bien une relation d'équivalence sur $\mathcal{M}_p(\mathbb{R})$ (laissé en exercice). Ainsi, deux matrices congruentes ont même rang (le rang de la forme bilinéaire associée), et que si il existe une matrice positive dans la classe d'équivalence associée à φ, alors, φ est positive et donc toutes les matrices de la classe d'équivalence le sont.

On peut utiliser ces notions pour obtenir une jolie démonstration de la décomposition de Cholesky.

— Soient $A, B \in \mathcal{M}_p(\mathbb{R})$ respectivement symétrique définie positive et symétrique positive. On pose $S \in \text{Rac}\, A \bigcap S_p^+(\mathbb{R})$. Alors AB est semblable à ${}^tSBS$, où S est la matrice de passage. On en déduit que AB est diagonalisable et que son spectre est contenu dans $\mathbb{R}_+$.

— On pose $\varphi : (X, Y) \in (\mathbb{R}^p)^2 \mapsto {}^tXAY$. φ est un produit scalaire si et seulement si A est symétrique définie positive. On considère la base canonique $(e_1, \ldots, e_p)$, qu'il est possible d'orthonormaliser au sens de Gram-Schmidt en une base orthonormée pour φ notée $\mathcal{B} = (\varepsilon_1, \ldots, \varepsilon_p)$. Si l'on note $P \in GL_p(\mathbb{R})$ la matrice de passage de la base $\mathcal{B}$ à la base canonique, comme I_p est la matrice de φ dans la base des ε_i, on a :

$$A = {}^tPI_pP = {}^tPP$$

La matrice P est triangulaire supérieure et ses coefficients diagonaux sont strictement positifs : en effet, ce sont les produits scalaires $\varphi(e_i, \varepsilon_i)$ et ils sont strictement positifs d'après les règles d'orthonormalisation de Gram-Schmidt.

C'est la décomposition de Cholesky.

Remarque : La réciproque est vraie, au sens suivant : si $A \in \mathcal{M}_n(\mathbb{R})$ et s'il existe une matrice $T \in \mathcal{M}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs tel que $A = {}^tTT$, alors $A \in S_n^{++}(\mathbb{R})$.

Remarque : Soient $A \in GL_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à résoudre le système linéaire $Ax = b$. Si A est symétrique définie positive, on peut calculer sa décomposition de Cholesky et procéder à une méthode de descente-remontée afin de déterminer x. Sinon, le système linéaire est équivalent à ${}^tAAx = {}^tAb$, où ${}^tAA$ est bien une matrice symétrique définie positive : on peut donc de même calculer la décomposition de Cholesky de ${}^tAA$, et procéder à une méthode de descente-remontée ; cette méthode nécessite les calculs supplémentaires de ${}^tAA$ et de ${}^tAb$, tous deux asymptotiquement négligeables devant le coût de la décomposition de Cholesky (pourvu qu'on dispose d'une méthode efficace pour calculer le produit matriciel ${}^tAA$).

⋆ ⋆ ⋆

VII Annexe : Espaces hermitiens et décomposition polaire

Définition

Étant donné un espace vectoriel complexe E, une forme sesquilinéaire est une application

$$\varphi \left| \begin{array}{l} E^2 \to \mathbb{C} \\ (x,y) \mapsto \varphi(x,y) \end{array} \right.$$

semi-linéaire à gauche et linéaire à droite (si $\lambda \in \mathbb{C}, x, y \in E$, on a $\varphi(\lambda x, y) = \overline{\lambda}\varphi(x,y)$ et $\varphi(x, \lambda y) = \lambda\varphi(x,y)$).

Elle est dite hermitienne si $\varphi(x,y) = \overline{\varphi(y,x)}$. La forme quadratique hermitienne associée est alors $q : \ x \to \varphi(x,x) \in \mathbb{R}$. Si $q(x) > 0$ pour x non nul, on dit que φ est définie positive : c'est un produit scalaire hermitien. La norme hermitienne associée est $x \to \sqrt{q(x)}$.

Le produit scalaire hermitien canonique de $\mathbb{C}^p$ est donné par :

$$\varphi(X,Y) = \overline{x_1}y_1 + \ldots + \overline{x_p}y_p,$$

avec $X = {}^t(x_1, \ldots, x_p)$ et $Y = {}^t(y_1, \ldots, y_p) \in \mathbb{C}^p$. Il y a beaucoup de parallèles avec les espaces euclidiens : notions d'orthogonalité, Cauchy-Schwarz, etc.

Considérons E un espace hermitien. Un endomorphisme u de E possède un unique endomorphisme adjoint u^* vérifiant pour tout $(x,y) \in E^2$, $\langle u(x), y\rangle = \langle x, u^*(y)\rangle$. On dit que u est hermitien ou auto-adjoint si $u^* = u$. Dans ces conditions, si $A = (a_{ij})_{1 \leqslant i,j \leqslant p}$ désigne la matrice de u dans une base orthonormale de E, la matrice A^* de coefficient (i,j) égal à $\overline{a_{j,i}}$ est la matrice de u^* dans la même base. La matrice A est dite hermitienne si $A = A^*$. Dans ces conditions, A est hermitienne si et seulement si u est hermitien.
On appelle unitaire un endomorphisme ou une matrice inversible dont l'inverse est égale à son adjoint.

La théorie des endomorphismes hermitiens est semblable en tout point à celle des endomorphismes symétriques, dont ils sont la généralisation sur des $\mathbb{C}$-espaces vectoriels. Le théorème spectral s'exprime de la sorte :

Théorème

Un endomorphisme hermitien est diagonalisable dans une base orthonormée et son spectre est réel ;
Une matrice hermitienne A est unitairement semblable à une matrice diagonale réelle, autrement dit, il existe $U \in \mathbb{U}_p(\mathbb{C})$ et D matrice diagonale réelle telles que

$$A = UDU^{-1} = UDU^*$$

On généralise la notion d'endomorphismes ou de matrices hermitiennes positives et définies positives, les ensembles étant notés $\mathcal{H}_p^+(\mathbb{C})$ et $\mathcal{H}_p^{++}(\mathbb{C})$.
Enfin, on peut ainsi généraliser la décomposition polaire aux espaces hermitiens :

Proposition

- Toute matrice $A \in GL_p(\mathbb{C})$ peut s'écrire de manière unique $A = UH$ où U est une matrice unitaire et H une matrice hermitienne définie positive.
- Toute matrice $A \in \mathcal{M}_p(\mathbb{C})$ peut s'écrire $A = UH$ où U est une matrice unitaire et H une matrice hermitienne positive.

Les preuves sont similaires à celles dans le cas réel. On peut montrer que :

Proposition

L'application

$$\left|\begin{array}{l}\mathbb{U}_p(\mathbb{C})\times\mathcal{H}_p^{++}(\mathbb{C})\to GL_p(\mathbb{C})\\(U,H)\mapsto UH\end{array}\right.$$

est un homéomorphisme.

Démonstration. Soit $M\in GL_p(\mathbb{C})$. On peut montrer que M^*M est hermitienne définie positive. Donc il existe $U\in\mathbb{U}_p(\mathbb{C})$ telle que $M^*M=U\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_p\end{pmatrix}U^*$ et $\lambda_1,\dots,\lambda_p$ p réels strictement positifs.

En posant $H=U\begin{pmatrix}\sqrt{\lambda_1}&&\\&\ddots&\\&&\sqrt{\lambda_p}\end{pmatrix}U^*$, on a $H=H^*$, c'est-à-dire H hermitienne. De plus, pour tout $X\in\mathbb{C}^*$ non nul, on a :

$$X^*HX=X^*U\begin{pmatrix}\sqrt{\lambda_1}&&\\&\ddots&\\&&\sqrt{\lambda_p}\end{pmatrix}U^*X=(U^*X)^*\begin{pmatrix}\sqrt{\lambda_1}&&\\&\ddots&\\&&\sqrt{\lambda_p}\end{pmatrix}U^*X>0$$

C'est-à-dire H est hermitienne définie positive. On pose $V=MH^{-1}$, alors :

$$V^*V=\left(MH^{-1}\right)^*MH^{-1}=\left(H^{-1}\right)^*M^*MH^{-1}=\left(H^{-1}\right)^*U\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_p\end{pmatrix}U^*H^{-1}$$

D'où :

$$V^*V=(U^*)^{-1}\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}}&&\\&\ddots&\\&&\frac{1}{\sqrt{\lambda_p}}\end{pmatrix}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_p\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}}&&\\&\ddots&\\&&\frac{1}{\sqrt{\lambda_p}}\end{pmatrix}U^{-1}$$

Or $U\in\mathbb{U}_n(\mathbb{C})$ donc $V^*V=I_n$, soit $V\in\mathbb{U}_n(\mathbb{C})$ et on a bien l'existence d'une décomposition polaire.

On considère une décomposition polaire quelconque $M=VH$, alors $M^*M=(VH)^*VH=H^*V^*VH$, or V est unitaire et H hermitienne donc $M^*M=H^2$. On note m et h les deux endomorphismes de $\mathbb{C}^n$ dont les matrices dans la base canonique de $\mathbb{C}^n$ sont respectivement M^*M et H. Si $\mu_1,\dots,\mu_k$ sont les valeurs propres de m (qui sont réelles positives) et $E_{\mu_1},\dots,E_{\mu_k}$ sont les sous-espaces propres associés, alors les E_{μ_i} sont stables par h (puisque m et h commutent) ; on peut donc considérer $h_i=h_{|E_{\mu_i}}$ pour tout $1\leqslant i\leqslant k$. Comme h_i est hermitien, h_i est diagonalisable et toute valeur propre λ de h_i est réelle positive et vérifie $\lambda^2=\mu_i$ donc $h_i=\sqrt{\mu_i}id_{E_{\mu_i}}$.

Ainsi, h est complètement détermiée par m, i.e H par M, ce qui assure l'unicité de la décomposition. L'application

$$\left|\begin{array}{l}\mathbb{U}_n(\mathbb{C})\times\mathcal{H}_n^{++}(\mathbb{C})\to GL_n(\mathbb{C})\\(U,H)\mapsto UH\end{array}\right.$$

est continue. Réciproquement, considérons une suite $(M_p)_p$ de $GL_n(\mathbb{C})$ qui converge vers une matrice $M \in GL_n(\mathbb{C})$. On pose $M = UH$ et $M_p = U_pH_p$ pour tout p. Puisque le groupe $\mathbb{U}_n(\mathbb{C})$ est compact, la suite $(U_p)_p$ admet une sous-suite convergente $(U'_{\varphi(p)})_p$ dont on note U_0 la limite ; alors la suite $(H'_{\varphi(p)})_p$ converge vers une matrice hermitienne positive H_0 et, comme $H_0 = MU_0^{-1}$ est inversible, H_0 est définie positive. L'unicité de la décomposition polaire donne alors $U = U_0$ et $H = H_0$. Cela signifie que la suite $(U_p)_p$ n'admet que U pour valeur d'adhérence et comme $\mathbb{U}_n(\mathbb{C})$ est compact, $(U_p)_p$ converge vers U ; il en résulte que $(H_p)_p$ converge vers H. L'application réciproque est donc bien continue.

⋆ ⋆ ⋆