Concours commun Mines-Ponts

DEUXIÈME ÉPREUVE. FILIÈRE MP

A. Quelques exemples

1) Pour tout réel θ, posons $S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$. Pour tout réel θ, S_θ est la matrice dans une base orthonormée de $\mathbb{R}^2$ muni de sa structure euclidienne canonique, de la symétrie orthogonale par rapport à la droite d'angle polaire $\dfrac{\theta}{2}$. Donc, pour tout réel θ, $S_\theta^2 = I_2$. La matrice $A = I_2$ admet donc une infinité de racines carrées deux à deux distinctes.

Soit X une racine carrée de A dans $\mathscr{M}_2(\mathbb{C})$ qui est un polynôme en A. Puisque A est une matrice scalaire, il en est de même de X. Donc, il existe $\lambda \in \mathbb{C}$ tel que $X = \lambda I_2$. L'égalité $X^2 = I_2$ fournit $\lambda^2 = 1$ puis $X = \pm I_2$. Réciproquement, les matrices I_2 et $-I_2$ sont de racines carrées de $A = I_2$ qui sont des polynômes en A.

2) $A = E_{1,3}$ puis $A^2 = 0_3$. Donc, A est nilpotente d'indice 2. Une matrice X qui est un polynôme en A est donc une matrice de la forme $X = \alpha I_3 + \beta E_{1,3}$, $(\alpha, \beta) \in \mathbb{C}^2$. Pour une telle matrice,

$$X^2 = \alpha^2 I_3 + 2\alpha\beta E_{1,3}$$

et donc, la famille $(I_3, E_{1,3})$ étant libre,

$$X^2 = A \Leftrightarrow \alpha^2 I_3 + 2\alpha\beta E_{1,3} = E_{1,3} \Leftrightarrow \alpha^2 = 0 \text{ et } 2\alpha\beta = 1$$

ce qui est impossible. Donc, aucune éventuelle racine carrée de $A = E_{1,3}$ n'est un polynôme en A.

Maintenant, pour $\alpha \in \mathbb{C}^*$, $\left(\alpha E_{1,2} + \dfrac{1}{\alpha} E_{2,3}\right)^2 = \alpha^2 E_{1,2}^2 + E_{1,2}E_{2,3} + E_{2,3}E_{1,2} + \dfrac{1}{\alpha^2} E_{2,3}^2 = E_{1,3}$. Donc, les matrices

$$\alpha E_{1,2} + \frac{1}{\alpha} E_{2,3} = \begin{pmatrix} 0 & \alpha & 0 \\ 0 & 0 & \dfrac{1}{\alpha} \\ 0 & 0 & 0 \end{pmatrix},\ \alpha \in \mathbb{C}^*,$$

sont des racines carrées deux à deux distinctes de $A = E_{1,3}$. Ainsi, $A = E_{1,3}$ admet une infinité de racines carrées deux à deux distinctes.

3) • Soit $A \in \mathscr{S}_n(\mathbb{R})$. D'après le théorème spectral, il existe $P \in O_n(\mathbb{R})$ et $D = \operatorname{diag}(\lambda_i)_{1 \leqslant i \leqslant n} \in \mathscr{D}_n(\mathbb{R})$ telles que $A = PD{}^tP$.
Supposons de plus que A soit définie positive. Donc, pour tout $i \in [\![1, n]\!]$, $\lambda_i > 0$. Soit alors $\Delta = \operatorname{diag}\left(\sqrt{\lambda_i}\right)_{1 \leqslant i \leqslant n}$ puis $R = P\Delta{}^tP$. R est orthogonalement semblable à une matrice diagonale et donc R est symétrique. De plus, les valeurs propres de R, à savoir les $\sqrt{\lambda_i}$, $1 \leqslant i \leqslant n$, sont des réels strictement positifs. Finalement, R est une matrice symétrique définie positive vérifiant

$$R^2 = \left(P\Delta{}^tP\right)^2 = P\Delta^2{}^tP = PD{}^tP = A.$$

R est une racine carrée de A qui est une matrice symétrique définie positive. Ceci montre l'existence d'une telle matrice.

• Par construction, la matrice R vérifie : si $\mathrm{Sp}(A) = (\lambda_1, \dots, \lambda_n)$, alors $\mathrm{Sp}(R) = \left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\right)$ et d'autre part, si on munit $\mathscr{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique, les colonnes de P constituent une base orthonormée de vecteurs propres associée à la famille $(\lambda_1, \dots, \lambda_n)$ de valeurs propres de A et aussi une base orthonormée de vecteurs propres associée à la famille $\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\right)$ de valeurs propres de R.

Soit S une racine carrée de A qui est une matrice symétrique définie positive. Soit λ une valeur propre de A. λ est un réel strictement positif. Les polynômes $X - \sqrt{\lambda}$ et $X + \sqrt{\lambda}$ sont premiers entre eux (car $\lambda \neq 0$) et donc, d'après le théorème de décomposition des noyaux,

$$\operatorname{Ker}(A - \lambda) = \operatorname{Ker}\left(S^2 - \lambda I_n\right) = \operatorname{Ker}\left(S - \sqrt{\lambda} I_n\right) \oplus \operatorname{Ker}\left(S + \sqrt{\lambda} I_n\right) = \operatorname{Ker}\left(S - \sqrt{\lambda} I_n\right),$$

car $-\sqrt{\lambda} < 0$ et donc $-\sqrt{\lambda}$ n'est pas valeur propre de S. Par suite, puisque

$$\mathscr{M}_{n,1}(\mathbb{R}) = \overset{\perp}{\bigoplus_{\lambda \in \mathrm{Sp}(A)}} \mathrm{Ker}\,(A - \lambda I_n)\,,$$

une base orthonormée de $\mathscr{M}_{n,1}(\mathbb{R})$ constituée de vecteurs propres de A et associée à la famille $(\lambda_1, \ldots, \lambda_n)$ de valeurs propres de A est encore une base orthonormée de $\mathscr{M}_{n,1}(\mathbb{R})$ de vecteurs propres de S associée à la famille de valeurs propres $\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\right)$. Dit autrement, si P et Δ sont les matrices telles que $R = P\Delta^t R$, la matrice ${}^tPSP$ est la matrice Δ et donc $S = R$. Ceci montre l'unicité de R.

B. Existence et calcul d'une racine carrée

4) Soit $(i, j) \in [\![1, n]\!]^2$.

• Si $i > j$, le coefficient ligne i, colonne j, de U^2 est $\displaystyle\sum_{k=1}^{n} u_{i,k}u_{k,j} = 0$ car dans cette somme si $i > k$, $u_{i,k} = 0$ et si $k \geqslant i > j$, $u_{k,j} = 0$.

• Si $i = j$, le coefficient ligne i, colonne j, de U^2 est $\displaystyle\sum_{k=1}^{n} u_{i,k}u_{k,i} = u_{i,i}^2$.

• Si $i < j$, le coefficient ligne i, colonne j, de U^2 est $\displaystyle\sum_{k=1}^{n} u_{i,k}u_{k,i} = \sum_{k=i}^{j} u_{i,k}u_{k,j}$ car si $k < i$, $u_{i,k} = 0$ et si $k > j$, $u_{k,j} = 0$.

Donc $U^2 = T \Leftrightarrow \left\{\begin{array}{l} \forall i \in [\![1, n]\!],\ u_{i,i}^2 = t_{i,i} \\ \forall (i, j) \in [\![1, n]\!]^2 \ \left(i < j \Rightarrow \displaystyle\sum_{k=i}^{j} u_{i,k}u_{k,j} = t_{i,j}\right) \end{array}\right.$ ou encore

$U^2 = T \Leftrightarrow \left\{\begin{array}{l} \forall i \in [\![1, n]\!],\ u_{i,i}^2 = t_{i,i} \quad (1) \\ \forall (i, j) \in [\![1, n]\!]^2 \ \left(i < j \Rightarrow (u_{i,i} + u_{j,j})\,u_{i,j} = t_{i,j} - \displaystyle\sum_{k=i+1}^{j-1} u_{i,k}u_{k,j}\right) \quad (2) \end{array}\right.$. (Le cas où $j = i + 1$ est conventionnel : la somme $\displaystyle\sum_{k=i+1}^{j-1} u_{i,k}u_{k,j}$ est vide et sa valeur est 0).

Tout nombre complexe non nul z admet deux racines carrées distinctes non nulles et opposées l'une à l'autre. Sur ces deux racines carrées, ou bien l'une des deux a une partie réelle strictement positive, ou bien l'une des deux a une partie réelle nulle et une partie imaginaire strictement positive (dans le cas où z est un réel strictement négatif et uniquement dans ce cas). Puisque T est inversible, tous les $t_{i,i}$, $1 \leqslant i \leqslant n$, sont non nuls et on peut donc, pour chaque équation $u_{i,i}^2 = t_{i,i}$, choisir pour solution un nombre $u_{i,i}$ du type précédent, ce que l'on fait. Par construction, pour tout $(i, j) \in [\![1, n]\!]^2$, $u_{i,i} + u_{j,j} \neq 0$, car ce nombre a soit une partie réelle strictement positive, soit une partie imaginaire strictement positive.

Soit $i \in [\![1, n]\!]$. Quand $j = i + 1$, l'équation (2) s'écrit $(u_{i,i} + u_{i+1,i+1})\,u_{i,i+1} = t_{i,i+1}$ et se résout en $u_{i,i+1} = \dfrac{t_{i,i+1}}{u_{i,i} + u_{i+1,i+1}}$.

Soient $i \in [\![1, n]\!]$ puis $p \in [\![i + 1, n - 1]\!]$ (si cela est possible). Supposons avoir résolu les équations $(u_{i,i} + u_{j,j})\,u_{i,j} = t_{i,j} - \displaystyle\sum_{k=i+1}^{j-1} u_{i,k}u_{k,j}$ pour $i + 1 \leqslant j \leqslant p$ et donc avoir obtenu les $u_{i,j}$ pour $1 \leqslant i \leqslant n$ et $i + 1 \leqslant j \leqslant p$.

Quand $j = p+1$, les équations (2) s'écrivent $(u_{i,i} + u_{p+1,p+1})\,u_{i,p+1} = t_{i,p+1} - \displaystyle\sum_{k=i+1}^{p} u_{i,k}u_{k,j}$ et se résolvent en $u_{i,p+1} = \dfrac{1}{u_{i,i} + u_{p+1,p+1}}\left(t_{i,p+1} - \displaystyle\sum_{k=i+1}^{p} u_{i,k}u_{k,j}\right)$, les $u_{i,k}u_{k,j}$, $1 \leqslant i < k < j \leqslant p$ étant déjà connus.

On a résolu par récurrence le système de l'énoncé et donc toute matrice triangulaire inversible admet au moins une racine carrée.

5) A est triangulable dans $\mathscr{M}_n(\mathbb{C})$ et donc il existe $P \in GL_n(\mathbb{C})$ et $T \in \mathscr{T}_{n,s}(\mathbb{C})$ telles que $A = PTP^{-1}$. De plus, A est inversible et donc T est inversible. On pose $R = PUP^{-1}$ où U est la matrice de la question précédente. On a

$$R^2 = \left(PUP^{-1}\right)^2 = PU^2P^{-1} = PTP^{-1} = A.$$

Donc, A admet au moins une racine carrée. Si de plus, aucune valeur propre de A n'est un réel strictement négatif, il en est de même de T et donc les $t_{i,i}$ ne sont pas des réels strictement négatifs. Mais alors les $u_{i,i}$, $1 \leqslant i \leqslant n$, qui sont les valeurs propres de U et donc de R, ont tous une partie réelle strictement positive. Donc, A admet une racine carrée dont les valeurs propres ont des parties réelles strictement positives.

C. Algorithme de Newton

6) Soit $(A, B) \in \mathscr{M}_n(\mathbb{C})^2$.

$$\begin{aligned}
\|AB\| &= \sqrt{\sum_{1\leqslant i,j\leqslant n}\left|\sum_{k=1}^{n} a_{i,k}b_{k,j}\right|^2} \\
&\leqslant \sqrt{\sum_{1\leqslant i,j\leqslant n}\left(\sum_{k=1}^{n} |a_{i,k}|\,|b_{k,j}|\right)^2} \\
&\leqslant \sqrt{\sum_{1\leqslant i,j\leqslant n}\left(\sum_{k=1}^{n} |a_{i,k}|^2\right)\left(\sum_{l=1}^{n} |b_{l,j}|^2\right)} \text{ (d'après l'inégalité de CAUCHY-SCHWARZ)} \\
&= \sqrt{\sum_{1\leqslant i,j\leqslant n}\left(\sum_{1\leqslant k,l\leqslant n} |a_{i,k}|^2\,|b_{l,j}|^2\right)} = \sqrt{\sum_{1\leqslant i,j,k,l\leqslant n} |a_{i,k}|^2\,|b_{l,j}|^2} = \sqrt{\left(\sum_{1\leqslant i,k\leqslant n} |a_{i,k}|^2\right)\left(\sum_{1\leqslant j,l\leqslant n} |b_{l,j}|^2\right)} \\
&= \|A\|\|B\|.
\end{aligned}$$

Donc, $\|\ \|$ est une norme sous-multiplicative.

7) Posons $m_A = \prod_{i=1}^{k}(X-\lambda_i)^{\beta_i}$ où les λ_i sont les valeurs propres deux à deux distinctes de A dans $\mathbb{C}$ et les β_i sont des entiers naturels non nuls.

$$\begin{aligned}
m_A(B) \in GL_n(\mathbb{C}) &\Leftrightarrow \prod_{i=1}^{k}(B-\lambda_i I_n)^{\beta_i} \in GL_n(\mathbb{C}) \\
&\Leftrightarrow \det\left(\prod_{i=1}^{k}(B-\lambda_i I_n)^{\beta_i}\right) \neq 0 \\
&\Leftrightarrow \prod_{i=1}^{k}(\det(B-\lambda_i I_n))^{\beta_i} \neq 0 \\
&\Leftrightarrow \forall i \in [\![1,k]\!],\ \det(B-\lambda_i I_n) \neq 0 \\
&\Leftrightarrow \forall i \in [\![1,k]\!],\ \lambda_i \notin \mathrm{Sp}(B) \Leftrightarrow \mathrm{Sp}(A)\cap\mathrm{Sp}(B) = \varnothing.
\end{aligned}$$

Montrons par récurrence que pour tout $k \in \mathbb{N}$, $A^kM = MB^k$.

- L'égalité est vraie quand $k = 0$.
- Soit $k \geqslant 0$. Supposons que $A^kM = MB^k$. Alors,

$$A^{k+1}M = AA^kM = AMB^k = MBB^k = MB^{k+1}.$$

Le résultat est démontré par récurrence. Mais alors, pour tout polynôme $P = \sum a_kX^k$ de $\mathbb{C}[X]$

$$P(A)M = \sum a_kA^kM = \sum a_kMB^k = MP(B).$$

En particulier, si $P = m_A$, on obtient

$$0 = m_A(A)M = Mm_A(B).$$

Si $m_A(B)$ est inversible, alors $M = 0$. Par contraposition, puisque $M \neq 0$, $m_A(B)$ n'est pas inversible et donc A et B ont une valeur propre en commun.

8) Soit $\lambda \in \mathbb{C}$ une valeur propre commune à A et B. Soient X un vecteur propre de A associé à λ et Y un vecteur propre de B^T associé à λ (qui est également valeur propre de ${}^tB$) puis $M = XY^T$. M est un élément de $\mathscr{M}_{n,1}(\mathbb{R})$.

$$AM = AXY^T = \lambda XY^T = \lambda M$$

et

$$MB = XY^TB = X\left(B^TY\right)^T = \lambda XY^T = \lambda M.$$

Donc, $AM = MB$. De plus, $M = (x_iy_j)_{1\leqslant i,j\leqslant n}$. Puisque $X \neq 0$ et $Y \neq 0$, il existe $(i,j) \in [\![1,n]\!]^2$ tel que $x_iy_j \neq 0$ et donc $M \neq 0$.

En résumé, A et B ont une valeur propre en commun si et seulement si il existe $M \in \mathscr{M}_n(\mathbb{C}) \setminus \{0\}$ telle que $AM = MB$.

9) Soit $X \in \mathscr{M}_n(\mathbb{C})$. Pour $H \in \mathscr{M}_n(\mathbb{C})$,

$$\|F(X+H) - F(X) - (XH+HX)\| = \left\|(X+H)^2 - X^2 - (XH+HX)\right\| = \left\|H^2\right\| \leqslant \|H\|^2.$$

Donc, pour $H \neq 0$, $\dfrac{1}{\|H\|}\|F(X+H) - F(X) - (XH+HX)\| \leqslant \|H\|$ puis $\displaystyle\lim_{\substack{H\to 0\\ H\neq 0}} \frac{1}{\|H\|}(F(X+H) - F(X) - (XH+HX)) = 0$.

Finalement, $F(X+H) \underset{H\to 0}{=} F(X) + (XH+HX) + o(H)$ où de plus, l'application $H \mapsto XH+HX$ est linéaire. Ceci montre F est différentiable en X et que

$$\forall H \in \mathscr{M}_n(\mathbb{C}),\ dF_X(H) = XH + HX.$$

Ensuite,

$$\begin{aligned}
dF_X \notin GL\left(\mathscr{M}_n(\mathbb{C})\right) &\Leftrightarrow \mathrm{Ker}\left(dF_X\right) \neq \{0\}\\
&\Leftrightarrow \exists H \in \mathscr{M}_n(\mathbb{C}) \setminus \{0\}/\ XH = H(-X)\\
&\Leftrightarrow X \text{ et } -X \text{ ont une valeur propre en commun}
\end{aligned}$$

ou aussi dF_X est inversible si et seulement si X et $-X$ n'ont pas de valeur propre en commun.

Si X n'est pas inversible, alors 0 est valeur propre commune à X et $-X$ et donc dF_X n'est pas inversible. Par contraposition, si dF_X est inversible, alors X est une matrice inversible.

10) Les valeurs propres de X^* sont des nombres complexes dont la partie réelle est strictement positive. On sait que si $\mathrm{Sp}\left(X^*\right) = (\mu_1,\ldots,\mu_n)$, alors $\mathrm{Sp}\left(-X^*\right) = (-\mu_1,\ldots,-\mu_n)$ et donc les valeurs propres de $-X^*$ ont des parties réelles strictement négatives. On en déduit que X^* et $-X^*$ n'ont pas de valeur propre en commun puis que dF_{X^*} est inversible.

Soit $X = \displaystyle\sum_{1\leqslant i,j\leqslant n} x_{i,j}E_{i,j} \in \mathscr{M}_n(\mathbb{C})$. Pour $(i,j) \in [\![1,n]\!]^2$, $\dfrac{\partial F}{\partial x_{i,j}}(X) = dF_X\left(E_{i,j}\right) = XE_{i,j} + E_{i,j}X$. Pour $(i,j) \in [\![1,n]\!]^2$, l'application $\dfrac{\partial F}{\partial x_{i,j}}$ est linéaire sur $\mathscr{M}_n(\mathbb{C})$ et donc continue sur $\mathscr{M}_n(\mathbb{C})$.

L'application F est donc de classe C^1 sur $\mathscr{M}_n(\mathbb{C})$ puis l'application $X \mapsto dF_X$ est continue sur $\mathscr{M}_n(\mathbb{C})$. Par continuité du déterminant, on en déduit que l'application $X \mapsto \det\left(dF_X\right)$ est une application continue sur $\mathscr{M}_n(\mathbb{C})$ à valeurs dans $\mathbb{C}$. Puisque $\det\left(DF_{X^*}\right) \neq 0$, par continuité de l'application $X \mapsto \det\left(dF_X\right)$, il existe $r > 0$ tel que pour tout $X \in \overline{B}\left(X^*, r\right)$, $\det\left(dF_X\right) \neq 0$ ou encore dF_X inversible.

11) $G\left(X^*\right) = X^* - \left(dF_{X^*}\right)^{-1}\left(F\left(X^*\right)\right) = X^* - \left(dF_{X^*}\right)^{-1}\left(X^{*2} - A\right) = X^* - \left(dF_{X^*}\right)^{-1}(0) = X^*$ car $\left(dF_{X^*}\right)^{-1}$ est linéaire.

Soit alors $H \in B(0,r)$. $X^* + H \in \overline{B}\left(X^*, r\right)$ puis dF_{X^*+H} est inversible et

$$G\left(X^*+H\right) - G\left(X^*\right) = \left(X^*+H\right) - X^* - \left(dF_{X^*+H}\right)^{-1}\left(F\left(X^*+H\right)\right) = H - \left(dF_{X^*+H}\right)^{-1}\left(X^*H + HX^* + H^2\right).$$

Maintenant, $dF_{X^*+H}(H) = \left(X^*+H\right)H + H\left(X^*+H\right) = X^*H + HX^* + 2H^2$ et donc $H = \left(dF_{X^*+H}\right)^{-1}\left(X^*H + HX^* + 2H^2\right)$ puis

$$\begin{aligned}
G\left(X^*+H\right) - G\left(X^*\right) &= \left(dF_{X^*+H}\right)^{-1}\left(X^*H + HX^* + 2H^2\right) - \left(dF_{X^*+H}\right)^{-1}\left(X^*H + HX^* + H^2\right)\\
&= \left(dF_{X^*+H}\right)^{-1}\left(H^2\right).
\end{aligned}$$

Ensuite, $dF_{X^*} \circ \left(\mathrm{Id} + \left(dF_{X^*}\right)^{-1} \circ dF_H\right) = dF_{X^*} + dF_H$ puis pour $H' \in \mathscr{M}_n(\mathbb{C})$,

$$(\mathrm{d}F_{X^*}+\mathrm{d}F_H)(H') = X^*H' + H'X^* + H'H + HH' = (X^*+H)H' + H'(X^*+H) = \mathrm{d}F_{X^*+H}(H')$$

et donc $\mathrm{d}F_{X^*+H} = \mathrm{d}F_{X^*} + \mathrm{d}F_H = \mathrm{d}F_{X^*} \circ \left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_H\right)$ puis

$$(\mathrm{d}F_{X^*+H})^{-1} = \left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_H\right)^{-1} \circ (\mathrm{d}F_{X^*})^{-1}.$$

12) Soit $X \in B(X^*, r)$. Soit $H = X - X^*$ de sorte que $X = X^* + H$ où $H \in B(0, r)$. On munit $\mathscr{L}(\mathscr{M}_n(\mathbb{C}))$ d'une norme N. L'application $M \mapsto \mathrm{d}F_M$ est continue $\mathscr{M}_n(\mathbb{C})$ (car F est de classe C^1 sur $\mathscr{M}_n(\mathbb{C})$ d'après des théorèmes généraux). On en déduit que l'application $M \mapsto \left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_M\right)^{-1}$ est continue sur le compact $\overline{B}(0, r)$ (d'après des théorèmes généraux). En particulier, cette application est bornée sur le compact $\overline{B}(0, r)$ et donc sur $B(0, r)$. Donc, il existe $C_1 > 0$ tel que

$$\forall M \in B(0, r),\ N\left(\left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_M\right)^{-1}\right) \leqslant C_1.$$

L'application $\Phi : \begin{array}{ccc} (\mathscr{L}(\mathscr{M}_n(\mathbb{C})), N) \times (\mathscr{M}_n(\mathbb{C}), \|\ \|) & \to & (\mathscr{M}_n(\mathbb{C}), \|\ \|) \\ (\varphi, M) & \mapsto & \varphi(M) \end{array}$ est bilinéaire sur l'espace de dimension finie $\mathscr{L}(\mathscr{M}_n(\mathbb{C})) \times \mathscr{M}_n(\mathbb{C})$. « On sait » alors qu'il existe $C_2 > 0$ tel que

$$\forall (\varphi, M) \in \mathscr{L}(\mathscr{M}_n(\mathbb{C})) \times \mathscr{M}_n(\mathbb{C}),\ \|\Phi(\varphi, M)\| \leqslant C_2 N(\varphi) \|M\|.$$

Enfin, puisque $(\mathrm{d}F_{X^*})^{-1}$ est une endomorphisme de l'espace de dimension finie $\mathscr{M}_n(\mathbb{C})$, on sait qu'il existe une constante $C_3 > 0$ tel que

$$\forall M \in \mathscr{M}_n(\mathbb{C}),\ \left\|(\mathrm{d}F_{X^*})^{-1}(M)\right\| \leqslant C_3 \|M\|.$$

Donc, pour toute matrice H de $B(0, r)$,

$$\begin{aligned} \|G(X) - X^*\| = \|G(X^*+H) - G(X^*)\| &= \left\|\left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_H\right)^{-1} \circ (\mathrm{d}F_{X^*})^{-1}\left(H^2\right)\right\| \\ &\leqslant C_2 N\left(\left(\mathrm{Id} + (\mathrm{d}F_{X^*})^{-1} \circ \mathrm{d}F_H\right)^{-1}\right) \times \left\|(\mathrm{d}F_{X^*})^{-1}\left(H^2\right)\right\| \\ &\leqslant C_1 C_2 C_3 \left\|H^2\right\| \leqslant C_1 C_2 C_3 \|H\|^2 \text{ (d'après la question 6)} \\ &= C_1 C_2 C_3 \|X - X^*\|^2. \end{aligned}$$

Le nombre $C = C_1 C_2 C_3 > 0$ convient.

13) Puisque C peut être quelconque, le résultat de l'énoncé est faux quand $k = 0$. On va montrer par récurrence que pour tout $k \in \mathbb{N}$, X_k existe et est dans $B(X^*, \rho)$ et que $\|X_k - X^*\| \leqslant \dfrac{(\rho C)^{2^k}}{C}$ en choisissant correctement ρ.

On choisit déjà $\rho \leqslant r$ de sorte que si $X_0 \in B(X^*, \rho) \subset B(X^*, r)$, alors X_1 existe. On choisit aussi ρ tel que $\rho C < 1$. On prend donc $\rho = \text{Min}\left\{r, \dfrac{1}{2C}\right\} > 0$. Dans ce cas,

$$\frac{(\rho C)^{2^0}}{C} = \rho \leqslant \rho$$

puis, la suite $\left(\dfrac{(\rho C)^{2^k}}{C}\right)_{k \in \mathbb{N}}$ étant décroissante (car $0 < \rho C < 1$), pour tout $k \in \mathbb{N}$, $\dfrac{(\rho C)^{2^k}}{C} \leqslant \rho$.

Montrons alors par récurrence que, si $X_0 \in B(X^*, \rho)$,

$$\text{pour tout } k \in \mathbb{N},\ X_k \text{ existe et est dans } B(X^*, \rho) \text{ et que } \|X_k - X^*\| \leqslant \frac{(\rho C)^{2^k}}{C} \quad (\mathscr{P}_k).$$

- X_0 est dans $B(X^*, \rho) \subset B(X^*, r)$. De plus,

$$\|X_0 - X^*\| \leqslant \rho = \frac{(\rho C)^{2^0}}{C}.$$

- Soit $k \geqslant 0$. Supposons $(\mathscr{P}_k)$. Alors, X_{k+1} existe

$$\|X_{k+1}-X^*\| = \|G(X_k)-X^*\| \leqslant C\|X_k-X^*\|^2 \leqslant C\left(\frac{(\rho C)^{2^k}}{C}\right)^2 = C\frac{(\rho C)^{2^{k+1}}}{C^2} = \frac{(\rho C)^{2^{k+1}}}{C}.$$

En particulier, $\|X_{k+1}-X^*\| \leqslant \rho$.

Le résultat est démontré par récurrence. Puisque qu'on a choisit ρ tel que $0 < \rho C < 1$, $\lim\limits_{k\to+\infty} \dfrac{\left(\rho\sqrt{C}\right)^{2^k}}{C} = 0$ puis la suite $(X_k)_{k\in\mathbb{N}}$ converge et $\lim\limits_{k\to+\infty} X_k = X^*$.

D. Forme équivalente

14) Supposons la suite $(X_k)_{k\in\mathbb{N}}$ bien définie par (N) et que $U_0 = X_0$. Montrons par récurrence que pour tout $k\in\mathbb{N}$, U_k est défini et que $U_k = X_k$.

- C'est vrai quand $k = 0$.
- Soit $k \geqslant 0$. Supposons que U_k existe et que $U_k = X_k$. La matrice H_k vérifie alors $X_kH_k + H_kX_k = A - X_k^2$ ou encore $dF_{X_k}(H_k) = -F(X_k)$. X_k est dans $\overline{B}(X^*, r)$ et donc dF_{X_k} est inversible. L'équation $dF_{X_k}(H_k) = -F(X_k)$ a une solution et une seule à savoir $H_k = -\left(dF_{X_k}\right)^{-1}(F(X_k))$. On en déduit que U_{k+1} existe et que

$$U_{k+1} = U_k + H_k = X_k - \left(dF_{X_k}\right)^{-1}(F(X_k)) = X_{k+1}.$$

Le résultat est démontré par récurrence.

Réciproquement, supposons la suite $(U_k)_{k\in\mathbb{N}}$ bien définie par (I) et que $X_0 = U_0$. Montrons par récurrence que pour tout $k\in\mathbb{N}$, X_k est défini et que $K_k = U_k$.

- C'est vrai quand $k = 0$.
- Soit $k \geqslant 0$. Supposons que X_k existe et que $X_k = U_k$. L'équation $U_kH_k + H_kU_k = A - U_k^2$ s'écrit encore $dF_{U_k}(H_k) = -F(U_k)$. Le fait que la suite U soit bien définie sous-entend probablement le fait que cette équation, d'inconnue H_k, a une solution et une seule. Si dF_{U_k} n'était pas inversible, on sait que l'ensemble des solutions de l'équation $dF_{U_k}(M) = -F(U_k)$ est soit vide, soit de la forme $\{M_0\} + \mathrm{Ker}\left(dF_{U_k}\right) \neq \{M_0\}$ et, en aucun cas, l'équation considérée a une et une seule solution. Donc, $dF_{U_k} = dF_{X_k}$ est inversible puis X_{k+1} existe et

$$X_{k+1} = X_k - \left(dF_{X_k}\right)^{-1}(F(X_k)) = U_k - \left(dF_{U_k}\right)^{-1}(F(U_k)) = U_{k+1}.$$

Le résultat est démontré par récurrence.

15) Par hypothèse, la suite $(X_k)_{n\in\mathbb{N}}$ est bien définie. En particulier, pour tout $k\in\mathbb{N}$, dF_{X_k} est inversible. D'après la question 9), pour tout $k\in\mathbb{N}$, $U_k = X_k$ est une matrice inversible.

Supposons que les conditions (I) sont vérifiées et que $U_0 = V_0$ commute avec A. Montrons par récurrence que pour tout $k\in\mathbb{N}$, V_k existe et $V_k = U_k$ commute avec A.

- $V_0 = U_0$ existe et commute avec A.
- Soit $k \geqslant 0$. Supposons le résultat pour k. Puisque U_k est inversible, on peut poser $G_k = \dfrac{1}{2}\left(U_k^{-1}A - U_k\right)$. Puisque U_k commute avec A, il en est de même de U_k^{-1} car $U_kA = AU_k \Rightarrow AU_k^{-1} = U_k^{-1}A$.

$$U_kG_k + G_kU_k = \frac{1}{2}\left(U_k\left(U_k^{-1}A - U_k\right) + \left(AU_k^{-1} - U_k\right)U_k\right) = A - U_k^2.$$

Par unicité, on a donc $H_k = G_k = \dfrac{1}{2}\left(U_k^{-1}A - U_k\right)$ puis

$$U_{k+1} = U_k + \frac{1}{2}\left(U_k^{-1}A - U_k\right) = \frac{1}{2}\left(U_k^{-1}A + U_k\right) = \frac{1}{2}\left(V_k + V_k^{-1}A\right) = V_{k+1}.$$

Enfin, $U_{k+1} = \dfrac{1}{2}\left(U_k^{-1}A - U_k\right) \in C(A)$ car U_k^{-1}, A et U_k sont dans $C(A)$ et car $C(A)$ est une sous-algèbre de $(\mathscr{M}_n(\mathbb{C}), +, ., \times)$.

Le résultat est démontré par récurrence.

16) La matrice $V_0 = \mu I_n$ commute avec A. Il s'agit de démontrer que pour tout $k\in\mathbb{N}$, $V_k = PD_kP^T$ où D_k est une matrice diagonale à coefficients strictement positifs : $D_k = \mathrm{diag}(\lambda_{k,1}, \ldots, \lambda_{k,n})$ où $\forall \ell \in [\![1, n]\!]$, $\lambda_{k,\ell} > 0$.

On montre ce résultat par récurrence. Le résultat à démontrer est vrai quand $k = 0$: $V_0 = \mu I_n = PD_0P^T$ où $D_0 = \text{diag}(\mu, \ldots, \mu)$. On a donc $\lambda_{0,1} = \ldots = \lambda_{0,n} = \mu > 0$.

Soit $k \geqslant 0$. Supposons que $V_k = PD_kP^T$ où $D_k = \text{diag}(\lambda_{k,1}, \ldots, \lambda_{k,n})$ avec $\forall \ell \in [\![1, n]\!]$, $\lambda_{k,\ell} > 0$. Alors, 0 n'est pas valeur propre de V_k et donc V_k est inversible puis

$$\begin{aligned} V_{k+1} &= \frac{1}{2}\left(V_k + V_k^{-1}A\right) = \frac{1}{2}\left(PD_kP^T + PD_k^{-1}P^TPDP^T\right) = P\left(\frac{1}{2}\left(D_k + D_k^{-1}D\right)\right)P^T \\ &= PD_{k+1}P^T \end{aligned}$$

où $D_{k+1} = \text{diag}\left(\frac{1}{2}\left(\lambda_{k,\ell} + \frac{1}{\lambda_{k,\ell}}\lambda_\ell\right)\right)_{1\leqslant\ell\leqslant n}$. Enfin, les coefficients $\lambda_{k+1,\ell} = \frac{1}{2}\left(\lambda_{k,\ell} + \frac{1}{\lambda_{k,\ell}}\lambda_\ell\right)$, $1 \leqslant \ell \leqslant n$, sont strictement positifs.

Le résultat est démontré par récurrence.

17) Soient $k \in \mathbb{N}$ et $\ell \in [\![1, n]\!]$.

$$\lambda_{k+1,\ell} - \sqrt{\lambda_\ell} = \frac{1}{2}\left(\lambda_{k,\ell} + \frac{1}{\lambda_{k,\ell}}\lambda_\ell\right) - \sqrt{\lambda_\ell} = \frac{1}{2\lambda_{k,\ell}}\left(\lambda_{k+1,\ell}^2 - 2\lambda_{k+1,\ell}\sqrt{\lambda_\ell} + \lambda_\ell\right) = \frac{1}{2\lambda_{k,\ell}}\left(\lambda_{k+1,\ell} - \sqrt{\lambda_\ell}\right)^2$$

et de même, $\lambda_{k+1,\ell} + \sqrt{\lambda_\ell} = \frac{1}{2\lambda_{k,\ell}}\left(\lambda_{k+1,\ell} + \sqrt{\lambda_\ell}\right)^2$ puis, puisque $\lambda_{k+1,\ell} + \sqrt{\lambda_\ell} \neq 0$,

$$\frac{\lambda_{k+1,\ell} - \sqrt{\lambda_\ell}}{\lambda_{k+1,\ell} + \sqrt{\lambda_\ell}} = \left(\frac{\lambda_{k,\ell} - \sqrt{\lambda_\ell}}{\lambda_{k,\ell} + \sqrt{\lambda_\ell}}\right)^2.$$

Mais alors, par récurrence, pour tout $k \in \mathbb{N}$, $\frac{\lambda_{k+1,\ell} - \sqrt{\lambda_\ell}}{\lambda_{k+1,\ell} + \sqrt{\lambda_\ell}} = \left(\frac{\lambda_{0,\ell} - \sqrt{\lambda_\ell}}{\lambda_{0,\ell} + \sqrt{\lambda_\ell}}\right)^{2^{k+1}} = \left(\frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}}\right)^{2^{k+1}}$ (ou aussi, pour tout $k \in \mathbb{N}$, $\frac{\lambda_{k,\ell} - \sqrt{\lambda_\ell}}{\lambda_{k,\ell} + \sqrt{\lambda_\ell}} = \left(\frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}}\right)^{2^k}$ qui paraissait plus naturel).

18) On en déduit encore que pour $k \in \mathbb{N}$ et $\ell \in [\![1, n]\!]$,

$$\lambda_{k,\ell}\left(1 - \left(\frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}}\right)^{2^k}\right) = \sqrt{\lambda_\ell}\left(1 + \left(\frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}}\right)^{2^k}\right).$$

On choisit alors $\mu = \sqrt{\lambda_n} > 0$ (où λ_n est la plus grande valeur propre de A).

Pour tout $\ell \in [\![1, n]\!]$, on a $0 \leqslant \frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}} < \frac{\mu + \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}} = 1$ (car $\lambda_\ell \neq 0$) et donc $\lim\limits_{k\to+\infty}\left(\frac{\mu - \sqrt{\lambda_\ell}}{\mu + \sqrt{\lambda_\ell}}\right)^{2^k} = 0$ puis

$$\forall \ell \in [\![1, n]\!],\ \lim_{k\to+\infty} \lambda_{k+1,\ell} = \sqrt{\lambda_\ell}.$$

On en déduit encore que

$$\lim_{k\to+\infty} V_k = \lim_{k\to+\infty} PD_kP^T = P\text{diag}\left(\sqrt{\lambda_\ell}\right)_{1\leqslant\ell\leqslant n} P^T = \sqrt{A}.$$

E. Stabilité

19) Puisque $V_0 = \sqrt{A}$, il est clair par récurrence que pour tout $k \in \mathbb{N}$, $V_k = \sqrt{A}$.

$(V_0 + \Delta)\left(V_0^{-1} - V_0^{-1}\Delta V_0^{-1}\right) = I_n - \Delta V_0^{-1} + \Delta V_0^{-1} - \left(\Delta V_0^{-1}\right)^2 = I_n - \left(\Delta V_0^{-1}\right)^2$. Vérifions que $\left(\Delta V_0^{-1}\right)^2 = 0$.

$\left(\Delta V_0^{-1}\right)^2 = \varepsilon^2 C_iC_j^TV_0^{-1}C_iC_j^TV_0^{-1}$. Puisque $V_0 = \sqrt{A} = P\text{diag}\left(\sqrt{\lambda_\ell}\right)_{1\leqslant\ell\leqslant n} P^T$, C_i est un vecteur propre de V_0 associé à la valeur propre λ_i puis C_i est un vecteur propre de V_0^{-1} associé à la valeur propre $\frac{1}{\sqrt{\lambda_i}}$. Donc, $V_0^{-1}C_i = \frac{1}{\sqrt{\lambda_i}}C_i$ puis

$$\left(\Delta V_0^{-1}\right)^2 = \frac{\varepsilon^2}{\sqrt{\lambda_i}}C_iC_j^TC_iC_j^TV_0^{-1}.$$

Maintenant, $C_j^TC_i$ est le produit scalaire usuel des colonnes C_i et C_j de la matrice orthogonale P. Puisque $i \neq j$, $C_j^TC_i = 0$ et donc $\left(\Delta V_0^{-1}\right)^2 = 0$. On a montré que $(V_0 + \Delta)\left(V_0^{-1} - V_0^{-1}\Delta V_0^{-1}\right) = I_n$. Donc, la matrice $\widehat{V_0} = V_0 + \Delta$ est inversible et $\widehat{V_0}^{-1} = (V_0 + \Delta)^{-1} = V_0^{-1} - V_0^{-1}\Delta V_0^{-1}$.

Ensuite, puisque $V_1 = V_0 = \sqrt{A}$,

$$\begin{aligned}\Delta_1 &= \widehat{V_1} - V_1 = \frac{1}{2}\left(\widehat{V_0} + \widehat{V_0}^{-1}A\right) - V_0 = \frac{1}{2}\left(V_0 + \Delta + \left(V_0^{-1} - V_0^{-1}\Delta V_0^{-1}\right)V_0^2 - 2V_0\right)\\ &= \frac{1}{2}\left(\Delta - V_0^{-1}\Delta V_0^{-1}V_0^2\right) = \frac{1}{2}\left(\Delta - V_0^{-1}\Delta V_0^{-1}A\right).\end{aligned}$$

20) Maintenant, $\sqrt{A}$ est symétrique, et donc

$$V_0^{-1}\Delta V_0^{-1}A = \varepsilon\left(\sqrt{A}\right)^{-1}C_iC_j^T\sqrt{A} = \varepsilon\left(\left(\sqrt{A}\right)^{-1}C_i\right)\left(\sqrt{A}C_j\right)^T = \varepsilon\sqrt{\frac{\lambda_j}{\lambda_i}}C_iC_j^T = \sqrt{\frac{\lambda_j}{\lambda_i}}\Delta.$$

Finalement, $\Delta_1 = \frac{1}{2}\left(1 - \sqrt{\frac{\lambda_j}{\lambda_i}}\right)\Delta$ puis $\widehat{V_1} = \sqrt{A} + \frac{1}{2}\left(1 - \sqrt{\frac{\lambda_j}{\lambda_i}}\right)\Delta$.

Mais alors, par récurrence, pour tout $k \in \mathbb{N}$, $\Delta_k = \left(\frac{1}{2}\left(1 - \sqrt{\frac{\lambda_j}{\lambda_i}}\right)\right)^k \Delta$ puis

$$\widehat{V_k} = \sqrt{A} + \left(\frac{1}{2}\left(1 - \sqrt{\frac{\lambda_j}{\lambda_i}}\right)\right)^k\Delta.$$

$\eta = \frac{1}{2}\left(1 - \sqrt{\frac{\lambda_j}{\lambda_i}}\right)$ convient.

21) On prend en particulier $i = 1$ et $j = n$ et on note $c = \frac{\lambda_n}{\lambda_1}$ le conditionnement de A. D'après la question précédente,

$$\forall k \in \mathbb{N},\ \widehat{V_k} = \sqrt{A} + \left(\frac{1}{2}\left(1 - \sqrt{c}\right)\right)^k\Delta.$$

La suite $\left(\widehat{V_k}\right)_{k\in\mathbb{N}}$ converge si et seulement si $-1 < \frac{1}{2}\left(1 - \sqrt{c}\right) \leqslant 1$ ou encore $-3 < -\sqrt{c} \leqslant 1$ ou enfin $c < 9$.

