Université de Versailles Saint-Quentin-en-Yvelines **Année Universitaire 2020-2021**
Master 1 CHPS **Cours de Calcul Numérique**
Tahar Z. Boulmezaoud

Rappels en algèbre linéaire matriciel

Le présent document ne remplace pas le cours effectué en amphi. Ce document contient en particulier des détails ou des approfondissements qui n'ont pas été abordés en cours. Chaque étudiant est libre de lire ou non ces approfondissements (bien qu'il est fortement conseillé de les lire). Seuls les notions et les développements abordés en cours sont au programme.

1 Généralités

Dans toute la suite n désignera un entier naturel non nul et K l'un de deux corps scalaires $\mathbb{R}$ ou $\mathbb{C}$. Parfois nous serons amenés à spécifier lequel des deux corps est considéré, et cela essentiellement pour la raison suivante ; contrairement à $\mathbb{R}$, $\mathbb{C}$ est un corps algébriquement clos, c'est-à-dire que tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine dans $\mathbb{C}$. Ainsi, tout polynôme non constant $P \in \mathbb{C}[X]$ est **scindé** sur $\mathbb{C}$, c'est-à-dire qu'il se factorise en produit de polynômes de degré 1.

Soit E un K-espace vectoriel. On rappelle qu'une **norme** $N(.)$ (ou $\|.\|$) sur un E est une application de E dans $\mathbb{R}_+$ vérifiant les trois propriétés suivantes

1. $N(\lambda v) = |\lambda| N(v)$ pour tous $\lambda \in K$ et $v \in E$,
2. $N(u + v) \leqslant N(u) + N(u)$ pour tous $u \in E$ et $v \in E$ (c'est l'inégalité triangulaire),
3. Pour tout $v \in E$, $N(v) = 0 \Longrightarrow v = 0$.

Dans ce cas, on dit que le couple (E, N) est un **espace normé**. Quand N vérifie uniquement les deux premiers critères, on dit que c'est une **semi-norme** sur E.

Notons par ailleurs qu'à toute norme N définie sur E, on peut associer une distance $d(.,.)$ de E définie par

$$\forall (u, v) \in E^2,\ d(u, v) = N(u - v).$$

Deux normes N et N' sont dites équivalentes s'il existe deux constantes $c_1 > 0$ et $c_2 > 0$ telles que

$$\forall u \in E,\ c_1 N(u) \leqslant N'(u) \leqslant c_2 N(u).$$

On rappelle l'important résultat

Théorème 1 *Si E est de dimension finie, alors toutes les normes sur E sont équivalentes.*

Définition 1 *Soit* $\langle .,. \rangle$ *une application de* E^2 *dans* K. *On dit que* $\langle .,. \rangle$ *est un produit scalaire sur* E *si elle est*

1. *sesquilinéaire : c'est-à-dire linéaire par rapport à la première variable et antilinéaire par rapport à la seconde. Autrement dit, pour tout* $\lambda \in K$ *et pour tous* u, v *et* w *dans* E, *on a*

$$\langle u + \lambda w, v \rangle = \langle u, v \rangle + \lambda \langle w, v \rangle \text{ et } \langle u, v + \lambda w \rangle = \langle u, v \rangle + \overline{\lambda} \langle u, w \rangle,$$

2. *symétrique hermitienne :* $\forall u,\ v \in E,\ \langle u, v \rangle = \overline{\langle v, u \rangle}$,
3. *positive :* $\forall u \in E,\ \langle u, u \rangle \geqslant 0$.
4. *définie :* $\forall u \in E,\ \ \langle u, u \rangle = 0 \Longleftrightarrow u = 0$.

Dans ce cas, on dit que le couple $(E, \langle .,. \rangle)$ *est un espace préhilbertien. Quand* E *est de dimension finie et* $K = \mathbb{R}$ *(resp.* $K = \mathbb{C}$*), on dit que c'est un espace* ***euclidien*** *(resp.* ***hermitien****).*

Un espace préhilbertien est un espace normé car on peut associer au produit scalaire $\langle .,. \rangle$ la norme

$$\|u\| = \sqrt{\langle u, u \rangle}. \tag{1}$$

On a alors l'inégalité de Cauchy-Schwarz

$$\forall u, v \in E, \ \ |\langle u, v \rangle| \leqslant \|u\|.\|v\|. \tag{2}$$

Deux vecteurs u et v sont dits **orthogonaux** lorsque $\langle u, v \rangle = 0$. On dit qu'ils sont **perpendiculaires** lorsque $Re(\langle u, v \rangle) = 0$ (Re désigne la partie réelle).
Si on se place dans K^n, le produit scalaire usuel

$$\langle X, Y \rangle = \sum_{i=1}^{n} x_i \overline{y_i},$$

est appelé le produit scalaire **canonique** (dit aussi **euclidien** quand $K = \mathbb{R}$ et **hermitien** quand $K = \mathbb{C}$). On notera $\|.\|_2$ la norme associée

$$\|X\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}.$$

Soient maintenant F_1,...,F_p des sous-espaces de E. La somme $F_1 + ... + F_p$ est dite **directe** si l'application

$$\begin{array}{rcl} \varphi & : & F_1 \times ... \times F_p \longrightarrow E \\ & & (x_1, ..., x_p) \mapsto x_1 + ... + x_p, \end{array}$$

est injective (autrement dit, si la décompostion de tout vecteur $x \in F_1 + ... + F_p$ sous forme $x_1 + ... + x_p$, avec $x_i \in F_i$ pour tout $i \leqslant p$, est unique). On écrit alors $F_1 \oplus ... \oplus F_p$ ou $\oplus_{i=1}^{p} F_i$.
On montre qu'une condition nécessaire et suffisante pour qu'une somme $F_1 + ... + F_p$ soit directe est que

$$F_i \cap (F_1 + ... + F_{i-1}) = \{0\} \text{ pour tout } 2 \leqslant i \leqslant p.$$

(ce qui se réduit à $F_1 \cap F_2 = \{0\}$ quand $p = 2$).

Soient désormais n et m deux entiers. On note $\mathscr{M}_{n,m}(K)$ l'espace des matrices à coefficients dans K comportant n lignes et m colonnes. L'espace des matrice carrées de taille $n \times n$ est notée $\mathscr{M}_n(K)$.
Dans toute la suite, les vecteurs colonnes de $\mathscr{M}_{n,1}(K)$ seront identifiés aux éléments de K^n, quand il n'y a pas de confusion.
Soit A une matrice de $\mathscr{M}_{n,m}(K)$. Pour tous entiers i et j avec $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant m$, on note $(A)_{ij}$ ou $(a_{i,j})$ l'élement de A qui est à la i-ème ligne et la j-ème colonne. On note A^T et A^* les deux matrices de $\mathscr{M}_{m,n}(K)$ définies respectivement par

$$(A^T)_{i,j} = A_{j,i}, \ (A^*)_{i,j} = \overline{A_{j,i}} \text{ pour tous } i \leqslant m \text{ et } j \leqslant n.$$

A^T est appelée la matrice **transposée** de A, tandis que A^* est appelée la matrice **adjointe** de A. Ces deux matrices sont identiques si et seulement si la matrice A est réelle (c'est-à-dire quand $\overline{A} = A$).
Supposons maintenant que la matrice A est carrée, $A \in \mathscr{M}_n(K)$. Alors, on dit que

- A est **singulière** si A n'est pas inversible, c'est-à-dire si $\det(A) = 0$.
- A est **réelle** si $\overline{A} = A$.
- A est **diagonale** si $(A)_{i,j} = 0$ pour tous $i, \ j \leqslant n$ tels que $i \neq j$.
- A est **triangulaire supérieure** (resp. inférieure) si $(A)_{i,j} = 0$ pour tous $i, \ j \leqslant n$ tels que $i > j$ (resp. $i < j$).
- A est **tridiagonale** si $(A)_{i,j} = 0$ pour tous $i, \ j \leqslant n$ tels que $|i - j| > 1$.
- A est **hermitienne** si $A^* = A$.
- A est **symétrique** si A est hermitienne et réelle (donc $A^T = A$).
- A est **unitaire** si $AA^* = A^*A = I$ (I est la matrice identité). Il en résulte que les vecteurs colonnes de A constituent une base orthonormale de K^n et que $X \in K^n$, $\|AX\|_2 = \|X\|_2$.
- A est **orthogonale** si A est réelle et unitaire (i. e. si $A^TA = AA^T = I$).
- A est **normale** si $A^*A = AA^*$.
- A est **nilpotente** s'il existe un entier $k \geqslant 1$ tel que $A^k = 0$,
- A est **semi-définie positive (SDP)** si A est hermitienne et

$$\forall x \in \mathbb{C}^n, \ x^*Ax \geqslant 0 \tag{3}$$

  (on note que x^*Ax est toujours réel car A est hermitienne).
  Si de plus l'inégalité (**??**) estv stricte pour tout $x \neq 0$, on dit que A est **définie positive (DP).**
- A est **semi-définie négative (SDN)** si $-A$ est SDP.
- A est **définie négative (DN)** si $-A$ est DP.

Rappelons les deux critères suivants concernant les matrices SDP et DP

Proposition 1 *Une matrice hermitienne A est semi-définie positive (resp. définie positive) si et seulement si toutes ses valeurs propres sont positives ou nulles (resp. strictement positives).*

Proposition 2 (Règle de Sylvestre) *Une matrice hermitienne A de taille n est définie positive si et seulement si les n déterminants Δ_k des matrices $k \times k$ obtenues en supprimant de A les $n-k$ dernières lignes et colonnes ($k = 1, ..., n$) sont strictements positifs.*

Exemple 1

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \ \Delta_1 = 2, \ \Delta_2 = 3, \ \Delta_3 = 4.$$

Attention, la règle de Sylvestre ne fonctionne pas pour les matrices semi-définies positives : c'est-à-dire si les Δ_k sont positifs ou nuls, A n'est pas nécessairement semi-définie positive comme le montre l'exemple qui suit

Exemple 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \ \Delta_1 = 1, \ \Delta_2 = 0, \ \Delta_3 = 0.$$

Pourtant avec $x = (0, -1, 2)$ on a $x^T Ax = -3 < 0$.

Soient maintenant deux bases $\mathcal{B} = (e_1 \dots e_n)$ (dite ancienne ici) et $\mathcal{B}' = (e'_1 \dots e'_n)$ (dite nouvelle) d'un espace de dimension finie E. Les éléments de la nouvelle base se décomposent dans l'ancienne sous la forme

$$e'_j = \sum_{i=1}^{n} p_{i,j} e_i, \ 1 \leqslant j \leqslant n.$$

La matrice $P = ((p_{i,j}))_{1 \leqslant i,j \leqslant n}$ est alors appelée la *matrice de passage de la base* $\mathcal{B}$ à la base $\mathcal{B}'$. On peut observer que les colonnes de la matrice P sont les vecteurs coordonnées des éléments de la nouvelles base. On peut aussi interpréter P comme la matrice représentative de l'application identité, de E muni de la base $\mathcal{B}'$ dans E muni de la base $\mathcal{B}$ (la base départ est la nouvelle base et non l'ancienne).
On rappelle que deux matrices carrées A et B sont **semblables** s'il existe une matrice S inversible telle que

$$A = SA'S^{-1}$$

La relation "semblabe à" est une relation d'équivalence dans $\mathscr{M}_n(K)$. *Deux matrices de $\mathscr{M}_n(K)$ sont semblables si et seulement si elles représentent le même endomorphisme d'un K-espace vectoriel de dimenion égale à n, dans deux bases différentes* $\mathcal{B}$ (pour A) et $\mathcal{B}'$ (pour A') (S est dans ce cas est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$).
Montrons par exemple que *toute matrice triangulaire supérieure est semblable à une matrice triangulaire inférieure.* Soit T une matrice triangulaire supérieure. Soit u l'endormorphisme de K^n de matrice T dans la base canonique $e_1, ..., e_n$. Dans la base $e_n, ..., e_1$, la matrice de u est la matrice triangulaire inférieure

$$T' = ((t'_{i,j}))_{1 \leqslant i, \ j \leqslant n}, \quad \text{avec } t'_{i,j} = t_{n+1-i, n+1-j}.$$

T est donc semblable à T'.

2 Réduction des matrices

La réduction d'une matrice A consiste à trouver une matrice B semblable à A et ayant une forme simple, de préférence la plus simple que possible. Nous allons voir qu'en général cette matrice B peut être choisie triangulaire (on parle alors de la triangulation de matrice A), voire même diagonale (on parle de diangonalisation). On verra aussi que quand B ne peut pas être choisie diagonale (c'est-à-dire quand A n'est pas diagonalisable), on peut néanmoins la choisir comme étant une matrice triangulaire supérieure comportant quelques termes égales à 1 au dessus de la diagonale et 0 partout ailleurs (on parle de réduction de Jordan ou de Jordanisation).
La réduction des matrices est un outil très utile dans la pratique et ses applications sont très nombreuses.

2.1 Valeurs propres et vecteurs propres

Définition 2 *Soit $A \in \mathscr{M}_n(K)$ une matrice carrée. Un scalaire $\lambda \in K$ est valeur propre de A s'il existe un vecteur non nul X de K^n tel que*

$$AX = \lambda X.$$

Cela est équivalent à dire que la matrice $A - \lambda I$ n'est pas inversible, c'est-à-dire que λ est racine du polynôme caractéristique $P_A \in K[X]$ défini par

$$\forall \lambda \in K,\ P_A(\lambda) = \det(A - \lambda I).$$

Dans ce cas,

- *La multiplicité m_λ de λ en tant que racine de P_A est appelé multiplicité* **algébrique**. *Si $m_\lambda = 1$, on dit que λ est valeur propre* **simple**. *Sinon, on dit qu'elle est* **multiple.**
- *L'espace $E_\lambda = \{X \in K^n \mid (A - \lambda I)X = 0\}$ est le* **sous-espace propre** *associé à λ. Tout élément de E_λ est appelé vecteur propre associé à λ.*
- *La dimension de E_λ, notée g_λ, est la* **multiplicité géométrique** *de λ.*

L'ensemble des valeurs propres de A, noté $\mathrm{sp}_K(A)$, est appelé le **spectre** de A. Ainsi

$$\mathrm{sp}_K(A) = \{\lambda \in K \mid P_A(\lambda) = 0\}. \tag{4}$$

Observons que quand $K = \mathbb{C}$, le polynôme caractéristique P_A admet exactement n racines et on a

$$\sum_{\lambda \in \mathrm{sp}_{\mathbb{C}}(A)} m_\lambda = n.$$

(ainsi $\mathrm{sp}_{\mathbb{C}}(A) \neq \emptyset$). Quand $K = \mathbb{R}$, on a

$$\sum_{\lambda \in \mathrm{sp}_{\mathbb{R}}(A)} m_\lambda \leqslant n. \tag{5}$$

Lemme 1 *Soit $A \in \mathscr{M}_n(K)$. Pour tout $\lambda \in \mathrm{sp}_K(A)$, on a*

$$1 \leqslant g_\lambda \leqslant m_\lambda. \tag{6}$$

Démonstration – Soit u l'endomorphisme de K^n dans lui même de matrice A dans la base canonique. Soit $(f_1, ..., f_{g_\lambda})$ une base de E_λ. En complétant cette la famille en une base $(f_1, ..., f_n)$ de K^n, la matrice de l'endomorphisme u dans cette base s'écrit

$$A' = \begin{pmatrix} \lambda I_g & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

où I_g est la matrice identité de taille g_λ, A_{12} une matrice de taille $g_\lambda \times (n - g_\lambda)$, tandis que A_{22} est une matrice carrée de taille $(n - g_\lambda)^2$. Il en résulte que λ est valeur propre de A' de multiplicité algébrique m'_λ au moins égale à g_λ. Puisque A et A' sont semblables (elles représentent le même endomorphisme u dans deux bases différentes), on conclut que $m_\lambda = m'_\lambda \geqslant g_\lambda$. ■

Quand $g_\lambda < m_\lambda$, on dit que la valeur propre λ est **défectueuse**. Une valeur propre simple ne peut être défectueuse car nécessairement $1 = g_\lambda = m_\lambda$.

Proposition 3 *Deux matrices semblables dans $\mathscr{M}_n(K)$ ont forcément les mêmes valeurs propres avec les mêmes multiplicités algébriques et géométriques. Elles ont aussi le même polynôme caractéristique.*

Proposition 4 *Soit $A \in \mathscr{M}_n(K)$ et soit $\mathrm{sp}_K(A) = \{\lambda_1, ..., \lambda_p\}$. Alors, la somme des sous-espaces propres $E_{\lambda_1} + ... + E_{\lambda_p}$ est directe.*

Démonstration – On a besoin du lemme général suivant

Lemme 2 *[Lemme des noyaux] Soit u un endomorphisme d'un K-espace vectoriel de dimension finie E. Soit $P(X)$ un polynôme de $K[X]$ de la forme*

$$P(X) = \Pi_{i=1}^{p} (X - k_i)^{\alpha_i},$$

où les k_i sont des éléments de K deux à deux différents, tandis que les α_i sont des entiers naturels non nuls. Si $P(u) = 0$ alors,

$$E = \oplus_{i=1}^{p} \mathrm{Ker}\,[(u - k_i I)^{\alpha_i}] \text{ (la somme est directe).}$$

Démonstration – Montrons d'abord que la somme des sous-espace $\mathrm{Ker}\,[(u - k_i I)^{\alpha_i}]$ est directe. Commençons par observer que

$$\forall i,\ j \leqslant p,\ i \neq j \Longrightarrow \mathrm{Ker}\,[(u - k_i I)^{\alpha_i}] \cap \mathrm{Ker}\,[(A - k_j I)^{\alpha_j}] = \{0\}.$$

En effet, étant donné que les deux polynômes $(X - k_i)^{\alpha_i}$ et $(X - k_j)^{\alpha_j}$ sont premiers entre eux, il existe deux polynômes Q_1 et Q_2 tels que

$$Q_1(X)(X - k_i)^{\alpha_i} + Q_2(X)(X - k_j)^{\alpha_j} = 1.$$

On a alors pour tout $x \in E$

$$Q_1(u)(u - k_i I)^{\alpha_i}(x) + Q_2(u)(u - k_j I)^{\alpha_j}(x) = x.$$

Ceci implique que $(u - k_i I)^{\alpha_i}(x)$ et $(u - k_j I)^{\alpha_j}(x)$ ne peuvent s'annuler au même temps si $x \neq 0$.
Supposons maintenant par l'absurde qu'il existe au moins un sous-ensemble non vide $S \subset \{1, ..., p\}$ et des vecteurs non nuls $(x_j)_{j \in S}$ tels que

$$\forall j \in S, x_j \in \mathrm{Ker}\,(u - k_j I)^{\alpha_j} \text{ et } \sum_{j \in S} x_j = 0$$

On peut choisir - sans perdre de généralité - S de cardinal minimal. Soit ℓ un élément particulier de S et posons $S' = S \backslash \{\ell\}$. On a $\# S' = \# S - 1$. En appliquant $(u - k_\ell I)^{\alpha_\ell}$ à l'identité précédente, on obtient

$$\sum_{j \in S'} x'_j = 0.$$

où, pour tout $j \in S'$, on a posé $x'_j = (u - k_\ell I)^{\alpha_\ell} x_j \in \mathrm{Ker}\,(u - k_j I)^{\alpha_j}$. De plus, pour tout $j \in S'$, $x'_j \neq 0$ (sinon $x_j \in \mathrm{Ker}\,(u - k_j I)^{\alpha_j} \cap \mathrm{Ker}\,(u - k_\ell I)^{\alpha_\ell} = \{0\}$). Cela contredit le fait que S est choisi de cardinal minimal.
Montrons maintenant que cette somme directe est égale à E. Rappelons tout d'abord l'identité de Bézout généralisée

Lemme 3 *Les P_1,...,P_ℓ des polynômes de $K[X]$. Ces polynômes sont premiers entre eux si et seulement s'il existe des polynômes Q_1,...,Q_ℓ tels que*

$$Q_1(X)P_1(X) + ... + Q_\ell(X)P_\ell(X) = 1.$$

On considère maintenant les polynômes

$$P_i(X) = \Pi_{j \neq i}(X - k_j)^{\alpha_j}, \ i = 1, ..., p.$$

Ces polynômes sont premiers entre eux. D'après l'identité de Bézout, il existe une famille de polynômes $(Q_i)_{i \leqslant p}$ telle que

$$\sum_{i=1}^{p} Q_i(X)P_i(X) = 1.$$

Donc

$$\sum_{i=1}^{p} Q_i(u)P_i(u) = I.$$

En remarquant que $(u - k_i I)^{\alpha_i} P_i(u) = P(u) = 0$, on en déduit que pour tout $x \in K^n$, le vecteur $x_i = Q_i(u)P_i(u)(x)$ appartient à $\mathrm{Ker}\,(u - k_i I)^{\alpha_i}$ et

$$x = \sum_{i=1}^{p} x_i.$$

On conclut que $\sum_{i=1}^{p} \mathrm{Ker}\,(u - k_i I)^{\alpha_i} = E$.
■

On complète maintenant la démonstration de la Proposition **??**. Soit H le sous-espace de K^n

$$H = E_{\lambda_1} + ... + E_{\lambda_p}.$$

Ponsons

$$P(X) = \Pi_{i=1}^{p}(X - \lambda_i)$$

et considérons l'endomorphisme u défini de H dans H par $u(X) = AX$. Tout vecteur $X \in H$ se décompose sous la forme

$$X = X_1 + ... + X_p,$$

avec $X_j \in E_{\lambda_j}$ pour $j \leqslant p$. Il en résulte que

$$P(u)(X) = \sum_{j=1}^{p} \Pi_{i=1}^{p}(A - \lambda_i I)X_j = 0.$$

Donc $P(u) = 0$. En appliquant le Lemme **??**, on en déduit que

$$H = \oplus_{i=1}^{p} \text{Ker}\,[u - k_i I] = \oplus_{i=1}^{p} E_{\lambda_i}.$$

■

2.2 Polynôme minimal, polynôme caractéristique et sous-espaces caractéristiques

Nous allons introduire maintenant la notion de polynôme minimal associé à une matrice. Considérons donc une matrice A de $\mathscr{M}_n(K)$. L'ensemble des polynômes

$$\mathscr{A} = \{p \in K[X] \mid p(A) = 0\}.$$

n'est pas réduit au singleton $\{0\}$. En effet, la famille de matrices $\{I, A, A^2, ..., A^{n^2}\}$ est forcément liée car son cardinal $(n^2 + 1)$ est strictement plus grand que la dimension de $\mathscr{M}_n(K)$ (qui égale à n^2). Il existe ainsi des scalaires $\alpha_0, ..., \alpha_{n^2}$, non tous nuls, tels que

$$\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + ... + \alpha_{n^2} A^{n^2} = 0.$$

Le polynôme $\alpha_0 + \alpha_1 X + \alpha_2 X^2 + ... + \alpha_{n^2} X^{n^2}$ appartient donc à $\mathscr{A}$.
Soit maintenant p_A le polynôme normalisé (i. e. ayant 1 comme coefficient dominant) de degré minimum de $\mathscr{A}$. On montre facilement, en utilisant la division euclidienne, qu'un tel polynôme est bien défini et qu'il vérifie la propriété suivante

$$\forall p \in K[X]\backslash\{0\},\ p(A) = 0 \Longrightarrow p_A \text{ divise } p.$$

Le polynôme p_A est appelé le **polynôme minimal** de A.

Théorème 2 (Cayley-Hamilton) *Soit $A \in \mathscr{M}_n(K)$. Le polynôme minimal p_A divise le polynôme caractéristique P_A, c'est-à-dire*

$$P_A(A) = 0.$$

Démonstration – Etant donné que toute matrice de $\mathscr{M}_n(\mathbb{R})$ peut être considérée comme une matrice de $\mathscr{M}_n(\mathbb{C})$ ayant le même polynôme caractéristique, on va démontrer ce résultat uniquement quand $K = \mathbb{C}$.
Commençons d'abord par le cas d'une matrice $A = ((a_{i,j}))_{1 \leqslant i,j \leqslant n}$ triangulaire supérieure. On a $P_A(X) = \Pi_{i=1}^{n}(a_{i,i} - X)^n$. Si on note u l'endormorphisme de K^n dont la matrice est A dans la base canonique $\{e_1, ..., e_n\}$, on a

$$u(e_1) - a_{1,1}e_1 = 0, u(e_2) - a_{2,2}e_2 = a_{1,2}e_1, ...,$$

$$u(e_j) - a_{j,j} = \sum_{i=1}^{j-1} a_{i,j}e_i, ...$$

On montre facilement par recurrence que

$$\forall j \leqslant n, (u - a_{1,1}I)...(u - a_{j,j}I)(e_j) = 0.$$

Donc

$$\forall j \leqslant n, (u - a_{1,1}I)...(u - a_{n,n}I)(e_j) = 0.$$

C'est-à-dire $P_A(u) = 0$, ce qui s'écrit matriciellement $P_A(A) = 0$.
Considérons maintenant le cas général d'une matrice quelconque $A \in \mathscr{M}_n(\mathbb{C})$. On a besoin du lemme suivant dont la démonstration est la même que celle du théorème **??** relatif à la factorisation de Schur.

Lemme 4 *Toute matrice carrée de $\mathscr{M}_n(\mathbb{C})$ est semblable à une matrice triangulaire.*

On en déduit qu'il existe une matrice triangulaire T et une matrice inversible P telles que $A = PTP^{-1}$. On a $P_A(X) = P_T(X)$ et $P_A(A) = PP_A(T)P^{-1} = PP_T(T)P^{-1} = 0$. ■

Proposition 5 *Soit $A \in \mathscr{M}_n(K)$. Toute valeur propre de A dans K est racine du polynôme minimal p_A.*

Démonstration – Soit λ une valeur propre de A et $X \neq 0$ un vecteur propre associé. On a d'une part $p_A(A) = 0$ et d'autre part $p_A(A)X = p_A(\lambda)X$. Donc $p_A(\lambda) = 0$. ■

Corollaire 1 *On suppose $\mathrm{sp}_K(A) \neq \emptyset$. Alors, le polynôme $\Pi_{\lambda \in \mathrm{sp}_K(A)}(X - \lambda)$ divise le polynôme minimal.*

On sait que le polynôme caractéristique se factorise sous la forme

$$P_A(X) = \Pi_{\lambda \in \mathrm{sp}_K(A)}(\lambda - X)^{m_\lambda} Q(X),$$

avec Q un polynôme sans racines dans K.
Le polynôme minimal s'écrit

$$p_A(X) = \Pi_{\lambda \in \mathrm{sp}_K(A)}(\lambda - X)^{\beta_\lambda} R(X),$$

avec R un polynôme sans aucune racine appartenant à $\mathrm{sp}_K(A)$. Le théorème de Cayley-Hamilton et la Proposition **??** impliquent que

$$\forall \lambda \in \mathrm{sp}_K(A), \ \ 1 \leqslant \beta_\lambda \leqslant m_\lambda.$$

De plus, le théorème de Cayley-Hamilton implique aussi que le polynôme $R(X)$ divise Q. Il est donc sans racines dans K.
Quand $K = \mathbb{C}$, le polynôme caractéristique admet exactement n racines (comptées avec leur ordres de multiplicité) et les deux polynômes P_A et p_A s'écrivent forcément sous la forme

$$P_A(X) = (-1)^n \Pi_{\lambda \in \mathrm{sp}_{\mathbb{C}}(A)}(X-\lambda)^{m_\lambda},\ p_A(X) = \Pi_{\lambda \in \mathrm{sp}_{\mathbb{C}}(A)}(X-\lambda)^{\beta_\lambda}. \qquad (7)$$

Considérons maintenant pour tout $\lambda \in K$ les sous-espaces

$$E_\lambda^k = \mathrm{Ker}\,(A-\lambda I)^k,\ k = 0, 1, 2, ...$$

On a

$$E_\lambda^0 = \{0\} \subset E_\lambda^1 = E_\lambda \subset E_\lambda^2 \subset ... \subset E_\lambda^k \subset E_\lambda^{k+1} \subset ....$$

On pose

$$F_\lambda = U_{k \geqslant 0} E_\lambda^k.$$

Proposition 6 *Soient $A \in \mathscr{M}_n(K)$ et $\lambda \in K$. Alors*

1. $E_\lambda \subset F_\lambda$.
2. *Il existe $k_\lambda \in \mathbb{N}$ tel que $F_\lambda = E_\lambda^k$ pour tout $k \geqslant k_\lambda$.*
3. *$F_\lambda \neq \{0\}$ si et seulement si λ est une valeur propre de A.*
4. *F_λ est un sous-espace vectoriel de K^n stable par A.*
5. *Pour tout $\mu \neq \lambda$, $F_\mu \cap F_\lambda = \{0\}$.*
6. $F_\lambda \oplus \mathrm{Im}\,(A-\lambda I)^{k_\lambda} = K^n$.
7. *Si λ est valeur propre, $1 \leqslant \dim F_\lambda \leqslant m_\lambda$.*
8. *Si λ est valeur propre simple $F_\lambda = E_\lambda$.*

Pour tout $\lambda \in \mathrm{sp}_K(A)$, F_λ est appelé *le sous-espace caractéristique (ou spectral)* associé à λ.

Démonstration –

1. L'inclusion $E_\lambda \subset F_\lambda$ est évidente car $E_\lambda = E_\lambda^1$.
2. On a clairement

$$\{0\} \subsetneq E^1 = E_\lambda \subset E_\lambda^2 \subset ... \subset E_\lambda^k \subset E^{k+1} \subset ... \subset K^n.$$

La suite des dimensions $(\dim E_\lambda^k)_{k \geqslant 0}$ de ces sous-espaces est croissante et majorée par n (la dimension de K^n). Il existe $k_\lambda \geqslant \{0, 1, ..., n\}$ tel que

$$\forall k \geqslant k_\lambda,\ \dim E_\lambda^k = \dim E_\lambda^{k_\lambda}$$

D'après les propriétés des sous-espaces vectoriels, on en déduit que

$$\forall k \geqslant k_\lambda,\ E_\lambda^k = E_\lambda^{k_\lambda}, \quad \text{et } F_\lambda = E_\lambda^{k_\lambda}.$$

3. Pour tout $k \in \mathbb{N}$, on a

$$\det[(A-\lambda I)^k] = [\det(A-\lambda I)]^k.$$

Il en résulte que $E_\lambda^k = \{0\}$ si et seulement si $A - \lambda I$ est inversible, c'est-à-dire ssi $\lambda \notin \mathrm{sp}_K(A)$.

4. Le fait que F_λ soit un sous-espace vectoriel résulte directement de la question précédente. Il est stable par A car

$$\forall X \in F_\lambda, \forall k \in \mathbb{N}, \ \ (A-\lambda I)^k AX = A(A-\lambda I)^k X = 0.$$

5. Soient $\lambda,\ \mu \in K$. Soit u l'endormorphisme de K^n défini par $u(X) = AX$ (la matrice de u dans la base canonique est A). F_λ est stable par u. On note u_λ la restriction de u à F_λ ; c'est un endomorphisme de F_λ dans lui-même. On a $(u_\lambda - \lambda I)^k = 0$ pour $k \geqslant 1$ assez grand.
   Supposons qu'il existe un vecteur non nul X apparetenant à $F_\lambda \cap F_\mu$. Nécessairement $(u_\lambda - \mu Id)^{k'} X = 0$ pour un entier k' assez grand (ici Id désigne l'application identité de F_λ). D'où $\det(u_\lambda - \mu Id)^{k'} = 0$, c'est-à-dire $\det(u_\lambda - \mu Id) = 0$. Il existe donc $X_0 \neq 0$ tel que $(u_\lambda - \mu Id)X_0 = 0$. Or $(u_\lambda - \lambda I)^k X_0 = 0$, d'où $(\mu-\lambda)^k X_0 = 0$. Nécessairement $\mu = \lambda$.
6. Soit $G_\lambda = \operatorname{Im}(A-\lambda I)^{k_\lambda}$. Montrons que $F_\lambda \cap G_\lambda = 0$. Soit $Y = (A-\lambda I)^{k_\lambda} X$ un élément de G_λ. Si de plus $Y \in F_\lambda$, alors $(A-\lambda I)^{k_\lambda} Y = (A-\lambda I)^{2k_\lambda} X = 0$. Donc $X \in \operatorname{Ker}(A-\lambda I)^{2k_\lambda} \subset F_\lambda$. D'où $Y = (A-\lambda I)^{k_\lambda} X = 0$.
   Par ailleurs, d'après le théorème du rang, on sait que $\dim F_\lambda + \dim G_\lambda = n$. On conclut que $F_\lambda \oplus \dim G_\lambda = K^n$.
   Les autres affirmations qui suivent sont admises.
7. On se place dans le cas où λ est valeur propre. Soit u l'endomorphisme de K^n de matrice A dans la base canonique. Soit $\ell = \dim F_\lambda$. On considère une base $(f_1, ..., f_\ell)$ de F_λ et $(g_{\ell+1}, ..., g_n)$ une base de G_λ. Étant donné que F_λ et G_λ sont stables par A, on en déduit que la matrice de l'endormorphisme u dans la base $(f_1, ..., f_\ell, g_{\ell+1}, ..., g_n)$ est de la forme

$$A' = \begin{pmatrix} A_F & 0 \\ 0 & A_G \end{pmatrix}.$$

   où les A_F et A_G sont les matrices de restrictions de u aux sous-espaces F_λ et G_λ. Il en résulte que

$$P_A(X) = P_u(X) = P_{A_F}(X) P_{A_G}(X).$$

   Par ailleurs, soit $k \geqslant 1$ tel que $F_\lambda = \operatorname{Ker}(A-\lambda)^k$. Il est clair que $(A_F - \lambda I)^k = 0$. Il en résulte que la seule valeur propre possible de A_F dans $\mathbb{C}$ est λ (en effet, si considère un vecteur complexe $X_0 \neq 0$ de A_F, vérifiant $A_F X_0 = \mu X_0$. On a $0 = (A_F - \lambda I)^k X_0 = (\mu - \lambda)^k X_0$. D'ù $\mu = \lambda$). On en déduit que, factorisé dans $\mathbb{C}$, le polynôme caractéristique de A_F est forcément de la forme $P_{A_F}(X) = (\lambda - X)^{d_\lambda}$ où $d_\lambda = \dim F_\lambda$. Puisque $P_A(X) = P_{A_F}(X) P_{A_G}(X)$, on en déduit que $d_\lambda \leqslant m_\lambda$.
8. Si $m_\lambda = 1$, on a forcément $\dim F_\lambda = 1$ car $1 \leqslant \dim F_\lambda \leqslant m_\lambda$.

■

Proposition 7 *On suppose que le polynôme minimal p_A scindé et s'écrit sous la forme*

$$p_A(X) = \Pi_{\lambda \in \mathrm{sp}_K(A)} (\lambda - X)^{\beta_\lambda},$$

avec $1 \leqslant \beta_\lambda \leqslant m_\lambda$ pour tout $\lambda \in \mathrm{sp}_K(A)$ (cette hypothèse est automatiquement satisfaite si $K = \mathbb{C}$ ou si $K = \mathbb{R}$ et P_A scindé). Alors,

1. *$K^n = \oplus_{\lambda \in \mathrm{sp}_{\mathbb{C}}(A)} F_\lambda$.*
2. *Pour tout $\lambda \in \mathrm{sp}_K(A)$, $F_\lambda = \mathrm{Ker}\,(A - \lambda I)^{\beta_\lambda} = \mathrm{Ker}\,(A - \lambda I)^{m_\lambda}$ et $\dim F_\lambda = m_\lambda$.*
3. *Le polynôme caractéristique de A est scindé aussi.*

Démonstration –

1. D'après la proposition **??**, pour tout $\lambda \in \mathbb{C}$, $F_\lambda = \mathrm{Ker}\,((A - \lambda I)^{k_\lambda})$ pour un certain entier k_λ, qui peut être choisi plus grand que β_λ. Le polynôme

$$P(X) = \Pi_{\lambda \in \mathrm{sp}_K(A)} (X - \lambda I)^{k_\lambda} = (-1)^n p_A(X) \Pi_{\lambda \in \mathrm{sp}_K(A)} (X - \lambda I)^{k_\lambda - \beta_\lambda}$$

vérifie $P(A) = 0$ (car $p_A(A) = 0$). Il en résulte d'après le Lemme **??**, que

$$K^n = \oplus_{\lambda \in \mathrm{sp}_K(A)} \mathrm{Ker}\,[(A - \lambda I)^{k_\lambda}] = \oplus_{\lambda \in \mathrm{sp}_K(A)} F_\lambda.$$

De plus, puisque $p_A(A) = 0$, on a aussi

$$K^n = \oplus_{\lambda \in \mathrm{sp}_K(A)} \mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}].$$

D'où

$$n = \sum_{\lambda \in \mathrm{sp}_K(A)} \dim F_\lambda = \sum_{\lambda \in \mathrm{sp}_K(A)} \dim \mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}].$$

De plus $\dim \mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}] \leqslant \dim F_\lambda$ pour tout $\lambda \in \mathrm{sp}_K(A)$, car $\mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}] \subset F_\lambda$. Forcément $\dim \mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}] = \dim F_\lambda$ et donc $F_\lambda = \mathrm{Ker}\,[(A - \lambda I)^{\beta_\lambda}]$ pour tout $\lambda \in \mathrm{sp}_K(A)$.
Nous allons montrer maintenant que $\dim F_\lambda = m_\lambda$ pour tous les λ. Soient $\lambda_1, ..., \lambda_p$ les éléments de $\mathrm{sp}_K(A)$. Soit u l'endomorpshime de K^n de matrice A dans la base canonique. Notons d'abord que les sous-espaces F_{λ_i} sont stables par u. Pour tout $i \leqslant p$, posons $\nu_i = \dim F_{\lambda_i}$ et considérons une base $\{f_{i,1}, ..., f_{i,\nu_i}\}$ de F_{λ_i}. La famille $(f_{i,k})_{1 \leqslant i \leqslant p,\ 1 \leqslant k \leqslant \nu_i}$ est une base de $K^n = \oplus_{i=1}^p F_{\lambda_i}$. Dans cette base, la matrice de u est de la forme

$$\begin{pmatrix} A_1 & & & \\ & A_2 & 0 & \\ & 0 & \ddots & \\ & & & A_p \end{pmatrix}.$$

où les A_i sont les matrices des restrictions de u aux sous-espaces F_{λ_i}. Il en résulte que

$$P_A(X) = \Pi_{i=1}^p P_{A_i}(X).$$

Mais puisque dans F_{λ_i} la seule valeur propre possible de u est λ_i, il s'en suit que $P_{A_i}(X) = (-1)^{\nu_i}(X-\lambda)^{\nu_i}$ pour $i \leqslant p$. Par comparaison des puissances dans la factorisation de $P_A(X)$, on en déduit que $\nu_i = m_{\lambda_i}$ pour tout $i \leqslant p$.

2. D'après ce qui précède, le polynôme caractéristique de A est scindé.

■

2.3 Trigonalisation. Factorisation de Schur

Théorème 3 (factorisation de Schur) *Soit A une matrice carrée de $\mathscr{M}_n(K)$ dont le polynôme caractéristique est scindé (ce qui est vrai si $K = \mathbb{C}$). Alors il existe une matrice unitaire U (orthogonale quand A est réelle) telle que*

$$A = UTU^{-1},$$

où T est une matrice triangulaire supérieure (réelle quand A est réelle).

Corollaire 2 *Une matrice A est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

DÉMONSTRATION – On a besoin du lemme suivant

Lemme 5 *Soit X un vecteur de K^n tel que $\|X\|_2 = 1$. Alors il existe des vecteurs $X_2, ..., X_n$ de K^n tels que la matrice de U dont les colonnes sont $X_1, ..., X_n$ soit unitaire.*

DÉMONSTRATION – Il suffit de compléter le vecteur X en une base de K^n et utiliser ensuite le procédé d'orthonomalisation de Gram-Schmidt. ■

Montrons maintenant le Théorème **??** par récurrence sur la dimension de n. Pour $n = 1$ le théorème est évident. Supposons que l'énoncé est vrai pour $n = 1, .., k-1$, $k \geqslant 2$. Montrons qu'il est vrai pour $n = k$. Soit $A \in \mathscr{M}_k(K)$ et soit C_1 un vecteur propre unitaire de A, associé à la valeur propre λ_1. D'après le lemme **??**, on peut construire une matrice unitaire U_1 ayant C_1 comme première colonne AU_1 aura comme première colonne $\lambda_1 U_1$ et la matrice $M = U_1^{-1}AU_1 = U_1^* AU_1$ est de la forme

$$M = \begin{pmatrix} \lambda_1 & \times ... \times \\ 0 & \boxed{M_1} \end{pmatrix}.$$

On $P_A(X) = P_M(X) = (\lambda_1 - X)P_{M_1}(X)$. Le polynôme caractéristique P_{M_1} est donc sincdé aussi. En appliquant l'hypothèse de récurrence à la matrice M_1, on en déduit qu'il existe une matrice unitaire G_1 telle que la matrice $T_1 = G_1^{-1}M_1G_1$ est triangulaire. La matrice

$$U_2 = \begin{pmatrix} 1 & 0...0 \\ 0 & \boxed{G_1} \end{pmatrix}.$$

est unitaire aussi. La matrice $U_2^{-1}MU_2 = U^{-1}AU$ avec $U = U_1U_2$ (unitaire) est triangulaire supérieure. Ce qui achève la démonstration. ■

2.4 Diagonalisation

Définition 3 *On dit qu'une matrice $A \in \mathscr{M}_n(K)$ est diagonalisable si et seulement si elle est semblable à une matrice diagonale.*

Considérons maintenant la somme des sous-espaces propres d'une matrice A

$$H = \sum_{\lambda \in \mathrm{sp}_K(A)} E_\lambda.$$

Cette somme est directe car $E_\lambda \subset F_\lambda$ pour tout λ et la somme des sous-espaces F_λ est directe. Il est de plus évident que tout vecteur propre de A appartient à H.

Proposition 8 *La matrice $A \in \mathscr{M}_n(K)$ est diagonalisable si et seulement si*

$$\oplus_{\lambda \in \mathrm{sp}_K(A)} E_\lambda = K^n.$$

Dans ce cas $F_\lambda = E_\lambda$ pour tout $\lambda \in \mathrm{sp}_K(A)$.

Démonstration – Posons encore

$$H = \oplus_{\lambda \in \mathrm{sp}_K(A)} E_\lambda.$$

Ainsi, si $H = K^n$, on peut construire une base $\mathcal{B}$ de K^n composée de vecteurs propres de A. Soit u l'endomorphisme de K^n dont la matrice est égale à A dans la base canonique. La matrice D de u dans la base $\mathcal{B}$ des vecteurs propres est diagnale et on a

$$A = PDP^{-1}.$$

La matrice A est donc diagonalisable.
Réciproquement, supposons que A est diagonalisable, s'écrivant sous la forme $A = PDP^{-1}$ avec D diagonale. Les vecteurs colonnes de P sont, d'une part, des vecteurs propres de A, donc appartiennent tous à H, et d'autre part ils forment une base de K^n puisque P est inversible. Il en résulte que $H = K^n$. ■

Proposition 9 *La matrice $A \in \mathscr{M}_n(K)$ est diagonalisable si et seulement si les deux conditions suivantes sont réalisées*

1. *le polynôme caractéristique de A est scindé sur K,*
2. *pour tout $\lambda \in \mathrm{sp}(A)$, $\dim E_\lambda = m_\lambda$ (ou $g_\lambda = m_\lambda$).*

Dans ce cas $F_\lambda = E_\lambda$ pour tout $\lambda \in \mathrm{sp}_K(A)$.

Démonstration – La somme directe

$$\oplus_{\lambda \in \mathrm{sp}_K(A)} E_\lambda,$$

ne peut être égale à K^n que si, et seulement si,

$$\sum_{\lambda \in \mathrm{sp}_K(A)} g_\lambda = n. \tag{8}$$

où $g_\lambda = \dim E_\lambda$ pour tout λ. Or on a vu que

$$\forall \lambda \in \mathrm{sp}_K(A), \quad g_\lambda \leqslant m_\lambda, \quad \sum_{\lambda \in \mathrm{sp}_K(A)} m_\lambda \leqslant n.$$

L'égalité (**??**) ne peut se produire que ssi $\sum_{\lambda \in \mathrm{sp}_K(A)} m_\lambda = n$, ce qui est à la condition 1 du théorème, et $g_\lambda = m_\lambda$ pour tout $\lambda \in \mathrm{sp}_K(A)$, ce qui est la condition 2. ■

Corollaire 3 *Soit $A \in \mathscr{M}_n(K)$. Si A admet n valeurs propres distinctes, alors A est diagonalisable.*

Dans le cas d'une matrice complexe, la première condition de la proposition **??** est automatiquement vérifiée et on a le corollaire

Corollaire 4 *Soit A une matrice de $\mathscr{M}_n(\mathbb{C})$. A est diagonalisable (dans $\mathscr{M}_n(\mathbb{C})$) si et seulement si*

$$\forall \lambda \in \mathrm{sp}_{\mathbb{C}}(A), g_\lambda = m_\lambda.$$

Le polynôme minimal fournit une caractérisation remarquable de la diagonalisation

Théorème 4 *Soit $A \in \mathscr{M}_n(K)$. A est diagonalisable si et seulement si le polynôme minimal de A est scindé à racines simples, c'est-à-dire si et seulement si*

$$p_A(X) = \Pi_{\lambda \in \mathrm{sp}_K(A)}(X - \lambda).$$

Démonstration – Supposons que A est diagonalisable et considérons le polynôme scindé à racines simples

$$P(u) = \Pi_{\lambda \in \mathrm{sp}(A)}(X - \lambda),$$

On sait que P divise P_A (Corollaire **??**). Par ailleurs, pour tout $\mu \in \mathrm{sp}(A)$ et tout $X \in E_\mu$, on a

$$P(A)(X) = \left(\Pi_{\lambda \in \mathrm{sp}(A) \setminus \{\mu\}}(A - \lambda I)\right)(A - \mu I)(X) = 0$$

(les matrices $A - \lambda I$, $\lambda \in \mathrm{sp}(A)$) commutent). Puisque $K^n = \oplus_{\lambda \in \mathrm{sp}(A)} E_\lambda$, on en déduit que $P(A)(X) = 0$ pour tout $X \in K^n$. Donc $P(A) = 0$ et p_A divise P. Puisque ces deux polynômes sont normalisés, on a $p_A = P$.
Réciproquement, supposons que p_A est scindé à racines simples a_1,...,a_k. Puisque p_A divise P_A et qu'il s'annule en tous les valeurs propres de A, les a_i sont exactement les valeurs propres de A. Puisque $p_A(A) = 0$, on en déduit d'après le lemme **??** que $K^n = \oplus_{i=1}^n E_{a_i}$. A est donc diagonalisable. ■

Considérons maintenant le cas où la matrice A est réelle. Une matrice réelle peut être diagonalisable dans $\mathscr{M}_n(\mathbb{C})$ mais pas dans $\mathscr{M}_n(\mathbb{R})$. Par exemple, la matrice

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

est diagonalisable dans $\mathscr{M}_n(\mathbb{C})$ car elle admet deux valeurs propres distinctes (i et $-i$), mais elle n'est pas diagonalisable dans $\mathscr{M}_n(\mathbb{R})$ (elle n'a pas de valeurs propres réelles).

Proposition 10 *Soit A une matrice réelle. Alors A est diagonalisable dans $\mathscr{M}_n(\mathbb{R})$ si et seulement si les deux conditions suivantes sont réalisées*

1. *Toutes les valeurs propres de A sont réelles.*
2. *A est diagonalisable dans $\mathscr{M}_n(\mathbb{C})$.*

DÉMONSTRATION – Il suffit de montrer que si les deux conditions sont vérifiées, alors A est diagonalisable dans $\mathscr{M}_n(\mathbb{R})$ (la réciproque étant évidente). Soit λ une valeur propre de A ($\lambda \in \mathbb{R}$). On pose

$$E_\lambda = \{X \in \mathbb{C}^n \mid AX = \lambda X\}, \hat{E}_\lambda = \{X \in \mathbb{R}^n \mid AX = \lambda X\}$$

Soit $(V_1, ..., V_g)$ une base de $\hat{E}_\lambda$. On montre facilement que c'est une famille libre de $\mathbb{C}^n$. Montrons qu'elle engendre E_λ dans $\mathbb{C}^n$. Soit W un éléments de E_λ. Les deux vecteurs réels $\Re(W)$ et $\Im(W)$ sont des vecteurs propres de A associés à λ (car $\lambda \in \mathbb{R}$). Ils appartiennent donc à $\hat{E}_\lambda$, et s'expriment en conséquence comme combinaisons linéaires (réelles) des vecteurs $V_1, ..., V_g$. Puisque $W = \Re(W) + i\Im(W)$, W s'exprime aussi comme combinaison linéaire (complexe) des vecteurs $V_1, ..., V_g$. On en déduit que $V_1, ..., V_g$ est une base du sous-espace E_λ (dans $\mathbb{C}^n$).
Il en résulte aussi que $\dim_{\mathbb{R}} \hat{E}_\lambda = \dim_{\mathbb{C}} E_\lambda$. Puisque A est diagonalisable dans $\mathscr{M}_n(\mathbb{C})$, on a

$$\oplus_{\lambda \in \mathrm{sp}(A)} E_\lambda = \mathbb{C}^n,$$

D'où

$$\sum_{\lambda \in \mathrm{sp}(A)} \dim_{\mathbb{R}} \hat{E}_\lambda = \sum_{\lambda \in \mathrm{sp}(A)} \dim_{\mathbb{C}} E_\lambda = n.$$

Il en résulte que $\oplus \hat{E}_\lambda = \mathbb{R}^n$ et A est diagonalisable dans $\mathscr{M}_n(\mathbb{R})$. ■

2.5 Deux familles de matrices diagonalisables

2.5.1 Matrices hermitiennes

Proposition 11 *Soit $A \in \mathscr{M}_n(\mathbb{C})$ une matrice hermitienne. Alors*

1. *Toutes les valeurs propres de A sont réelles.*
2. *Si X et Y sont deux vecteurs propres associés à deux valeurs propres distinctes, alors ils sont orthogonaux pour le produit scalaire hermitien. Autrement dit $X^T\overline{Y} = 0$.*

DÉMONSTRATION – Soit A une matrice hermitienne, λ une valeur propre de A et X un vecteur propre non nul associé à λ. On a d'une part

$$X^T\overline{AX} = \overline{\lambda}X^T\overline{X} = \overline{\lambda}\|X\|^2.$$

D'autre part,

$$X^T\overline{AX} = (X^*A^*X)^T = (X^*AX)^T = (\lambda\|X\|^2)^T = \lambda\|X\|^2.$$

Par comparaison, on en déduit que $\overline{\lambda} = \lambda$.
De même, si Y_1 et Y_2 sont deux vecteurs propres associés à deux valeurs propres différentes λ_1 et λ_2, on a

$$Y_1^T\overline{AY_2} = \overline{\lambda}_2 Y_1^T\overline{Y}_2 = (Y_2^*A^*Y_1)^T = \lambda_1 Y_1^T\overline{Y}_2.$$

D'où $Y_1^T \overline{Y}_2 = 0$. ■

Théorème 5 *Soit A une matrice hermitienne. Alors il existe une matrice réelle diagonale D et une matrice unitaire U (orthogonale quand A est réelle) telles que*

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

En d'autres termes, A est diagonalisable dans une base orthonormée de vecteurs propres.

Démonstration – Il s'agit d'un cas particulier du théorème **??** plus général démontré dans la suite. ■

Les valeurs propres d'une matrice hermitienne ont d'autres propriétés particulières. Nous allons énumerer certaines. Soit A une matrice de $\mathscr{M}_n(K)$ hermitienne quand $K = \mathbb{C}$ ou symétrique quand $K = \mathbb{R}$. Le **quotient de Rayleigh** associé à la matrice A est l'application de $K^n \backslash \{0\}$ dans $\mathbb{R}$ définie par

$$\forall V \in K^n \backslash \{0\}, \;\; R_A(V) = \frac{V^*AV}{V^*V} = \frac{V^*AV}{\|V\|_2^2}.$$

Ce quotient est bien un réel, même quand la matrice A est complexe (mais hermitienne tout de même).
Par ailleurs, soient $\lambda_1 \leqslant \lambda_2 \leqslant .... \leqslant \lambda_n$ les valeurs propres de la matrice A (on sait qu'elles sont réelles). Soit $(V_i)_{1 \leqslant i \leqslant n}$ une base **orthonormale** de vecteurs propres associés respectivement à ces valeurs propres. On a

$$\forall k \leqslant n, \; R_A(V_k) = \lambda_k.$$

Pour tout entier $k \leqslant n$, on note $\mathscr{V}_k$ l'ensemble de tous les sous-espace de K^n dont la dimension est égale à k et on pose, quand $k \geqslant 1$,

$$G_k = \mathrm{vect}\{V_1, ..., V_k\}.$$

Théorème 6 *Pour tout $1 \leqslant k \leqslant n$, on a*

$$\lambda_k(A) = \max_{V \in G_k, \; V \neq 0} R_A(V), \tag{9}$$

$$\lambda_k(A) = \min_{V \in G_{k-1}^{\perp}, \; V \neq 0} R_A(V), \tag{10}$$

$$\lambda_k(A) = \min_{W \in \mathscr{V}_k} \max_{V \in W, \; V \neq 0} R_A(V), \tag{11}$$

$$\lambda_k(A) = \max_{W \in \mathscr{V}_{k-1}} \min_{V \in W^{\perp}, \; V \neq 0} R_A(V). \tag{12}$$

Démonstration – Soit ℓ et m deux entiers tels que $1 \leqslant \ell \leqslant m \leqslant n$. Soit

$$H = \mathrm{vect}\{V_\ell, ..., V_m\}.$$

Pour tout $V = \sum_{i=\ell}^m \alpha_i V_i \in H \backslash \{0\}$, on a

$$R_A(V) = \frac{\sum_{i=\ell}^m \lambda_i |\alpha_i|^2}{\sum_{i=\ell}^m |\alpha_i|^2}.$$

D'où clairement

$$\lambda_\ell \leqslant R_A(V) \leqslant \lambda_m,$$

avec égalités quand $V = V_\ell$ et $V = V_\ell$ respectivement. Il en résulte que

$$\min_{V \in H \backslash \{0\}} R_A(V) = \lambda_\ell, \max_{V \in H \backslash \{0\}} R_A(V) = \lambda_m.$$

On obtient (**??**) quand $\ell = 1$ et $m = k$ et (**??**) quand $\ell = k$ et $m = n$ (on note que $G_{k-1}^\perp = \text{vect}\{V_k, ..., V_n\}$).
Puisque G_k est un sous-espace de dimension k, il en résulte en particulier que

$$\lambda_k(A) \geqslant \inf_{W \in \mathscr{V}_k} \max_{V \in W,\, V \neq 0} R_A(V),$$

Inversement, soit W un sous-espace de dimension k. Puisque $\dim W + \dim G_{k-1}^\perp = k + (n+1-k) = n+1$, la somme $W + G_{k-1}^\perp$ n'est pas directe et il existe forcément un vecteur non nul $V_0 \in W \cap G_{k-1}^\perp$. Pour ce vecteur, on a les deux inégalités

$$\min_{V \in G_{k-1}^\perp,\, V \neq 0} R_A(V) \leqslant R_A(V_0) \leqslant \max_{V \in W,\, V \neq 0} R_A(V),$$

Il en résulte que

$$\lambda_k(A) \leqslant \max_{V \in W,\, V \neq 0} R_A(V).$$

Cette égalité est vraie pour tout $W \in \mathscr{V}_k$. On conclut que

$$\lambda_k(A) \leqslant \inf_{W \in \mathscr{V}_k} \max_{V \in W,\, V \neq 0} R_A(V),$$

Il s'agit de plus d'une égalité (car on a déjà montré l'inégalité inverse). Le "inf" à droite est atteint quand $W = G_k$. C'est donc un "min".
La preuve de la dernière égalité (**??**) est analogue. ■

Corollaire 5 *Soit A une matrice hermitienne et M une matrice semi-définie positive. On note $\lambda_1(A) \leqslant ... \leqslant \lambda_n(A)$ (resp. $\lambda_1(A+M) \leqslant ... \leqslant \lambda_n(A+M)$) les valeurs propres de la matrice A (resp. $A+M$). Alors*

$$\lambda_k(A) \leqslant \lambda_k(A+M). \tag{13}$$

Démonstration – Il suffit d'utiliser la caractérisation précédente des valeurs propres d'une matrice hermitienne en remarquant que pour tout $V \in \mathbb{C}^n$ on a

$$R_{A+M}(V) = R_A(V) + R_M(V) \geqslant R_A(V).$$

■

2.5.2 Matrices normales

On rappelle qu'une matrice A est dite normale si elle vérifie $A^*A = AA^*$. Ainsi, toutes les matrices hermitiennes sont normales. De même, toute matrice unitaire (vérifiant $A^*A = AA^* = I$) est normale. Avant d'énoncer le théorème principal concernant la réduction des matrices normales, commençons par le lemme suivant qui caractérise les matrices triangulaires normales

Proposition 12 *Une matrice triangulaire est normale si et seulement si elle est diagonale.*

DÉMONSTRATION – Soit $T = ((t_{i,j}))_{1 \leqslant i,j \leqslant n}$ une matrice triangulaire supérieure et normale. Montrons qu'elle est diagonale. Posons $A = TT^* = T^*T$ avec $A = ((a_{i,j}))_{1 \leqslant i,j \leqslant n}$. On a d'une part

$$a_{i,j} = \sum_{k=1}^{n} t_{i,k}\bar{t}_{j,k} = \sum_{k=\max(i,j)}^{n} t_{i,k}\bar{t}_{j,k}.$$

D'autre part,

$$a_{i,j} = \sum_{k=1}^{n} \bar{t}_{k,j} t_{k,i} = \sum_{k=1}^{\min(i,j)} \bar{t}_{k,j} t_{k,i}.$$

En prenant $j = i$ dans ces deux égalités, on obtient

$$\sum_{k=i}^{n} |t_{i,k}|^2 = \sum_{k=1}^{i} |t_{k,i}|^2. \qquad (14)$$

Avec $i = 1$, on a

$$\sum_{k=2}^{n} |t_{1,k}|^2 = 0.$$

ce qui implique

$$\forall k \geqslant 2,\ t_{1,k} = 0.$$

Supposons maintenant par récurrence que

$$\forall i \leqslant m,\ \forall k > i,\ t_{i,k} = 0,$$

pour un certain $m < n$. Montrons que

$$\forall k > m+1,\ t_{m+1,k} = 0, \qquad (15)$$

En prenant $i = m+1$ dans (**??**), on a

$$\sum_{k=m+1}^{n} |t_{m+1,k}|^2 = \sum_{k=1}^{m+1} |t_{k,m+1}|^2 = |t_{m+1,m+1}|^2,$$

D'où

$$\sum_{k=m+2}^{n} |t_{m+1,k}|^2 = 0,$$

et on obtient (**??**) . On conclut que T est diagonale. ■

Théorème 7 *Une matrice N est normale si et seulement si il existe une matrice unitaire U et une matrice diagonale D telle que*

$$N = U^*DU. \tag{16}$$

Démonstration – Soit N une matrice de la forme $N = U^*DU$ avec D diagonale et U unitaire. On a

$$N^*N = U^*DUU^*D^*U = U^*DD^*U = U^*D^*DU,$$

$$NN^* = U^*DUU^*D^*U = U^*DD^*U = U^*D^*DU = N^*N.$$

Donc N est normale.
Inversement, soit N une matrice normale. En utilisant le théorème de factorisation de Schur on peut écrire N sous la forme $N = U^*TU$, avec T triangulaire et U normale, on a

$$T^*T = (UNU^*)^*(UNU^*) = UN^*U^*UNU^* = UN^*NU^* = UNN^*U^* = TT^*.$$

Il en résulte que T est triangulaire et normale. T est donc diagonale, ce qui achève la démonstration. ■

Corollaire 6 *Une matrice normale est diagonalisable dans une base orthonormale de vecteurs propres.*

Remarque – Si A est une matrice unitaire alors pour tout $\lambda \in \mathrm{sp}_{\mathbb{C}}(A)$, $|\lambda| = 1$. Attention, si de plus A est réelle (donc orthogonale), on ne peut dire que $\lambda \in \{-1, 1\}$. En effet, il se peut que λ soit complexe non réel de module 1. Par exemple la matrice

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

est orthogonale mais n'a aucune valeur propre réelle si $\sin\theta \neq 0$.

2.6 Une famille de matrices non diagonalisables : les matrices nilpotentes

Théorème 8 *Soit $A \in \mathscr{M}_n(K)$ une matrice nilpotente mais non nulle. Alors*

1. *A est singulière.*
2. $\mathrm{sp}_K(A) = \{0\}$.
3. $p_A(X) = X^m$ *avec* $2 \leqslant m \leqslant n$.
4. *A n'est pas diagonalisable.*
5. *toute matrice semblable à A est nilpotente aussi.*

Démonstration – Soit $k \geqslant 2$ tel que $A^k = 0$. Puisque $\det(A^k) = \det(A)^k$, on en déduit que $\det(A) = 0$. A est donc singulière et $0 \in \mathrm{sp}_K(A)$. De plus, si λ est valeur propre de A, λ^k est valeur propre de $A^k = 0$. Donc $\lambda = 0$. On en déduit que $\mathrm{sp}(A) = 0$.
Le polynôme minimal p_A divise le polynôme X^k, car ce dernier s'annule en A. Donc p_A est de la forme X^m avec $2 \leqslant m \leqslant k$ ($m \neq 1$, sinon $A = p_A(A) = 0$). Comme p_A est scindé mais pas à racines simples, A n'est pas diagonalisable.
Enfin, si $A' = PAP^{-1}$ est une matrice semblable à A, alors $(A')^k = PA^kP^{-1} = 0$. ■

Les matrices nilpotentes de Jordan sont un exemple classique. Elles s'écrivent sous la forme

$$U = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & ... & 0 & 0 & 1 \\ 0 & ... & ... & 0 & 0 \end{pmatrix}.$$

(la matrice nulle 1×1 sera considérée comme une matrice de Jordan nilpotente dans la suite).
Plus généralement, toute matrice de la forme suivante est nilpotente

$$A = \begin{pmatrix} \boxed{U_1} & 0 & 0 & ... & 0 \\ 0 & \boxed{U_2} & 0 & ... & 0 \\ ... & ... & .. & ... & ... \\ 0 & ... & 0 & \boxed{U_{m-1}} & 0 \\ 0 & ... & ... & 0 & \boxed{U_m} \end{pmatrix}. \tag{17}$$

où U_1,...,U_m sont des matrices niloptentes de Jordan. Bien qu'une matrice nilpotente n'est pas diagonalisable, le théorème suivant affirme qu'elle est toujours semblable à une matrice de type (**??**)

Théorème 9 *Toute matrice nilpotente est semblable à une matrice de la forme (**??**).*

DÉMONSTRATION – Soit $p_A(X) = X^k$, $m \geqslant 2$ le polynôme minimal de A. On considère les sous-espaces de K^n $N_i = \mathrm{Ker}\,(A^m)$, $0 \leqslant m \leqslant k$. On a $N_0 = \{0\} \subset N_1 \subset ... \subset N_{k-1} \subset N_k = K^n$. On notera l'implication

$$X \in N_m \Longrightarrow AX \in N_{m-1},$$

valable pour $2 \leqslant m \leqslant k$.
On construit par une récurrence descendante la suite $F_k, ..., F_0$ suivante

— F_k est un supplémentaire quelconque de N_{k-1} dans N_k.
— Supposons qu'on a construit pour tout m tel que $i+1 \leqslant m \leqslant k$ un sous-espace F_m supplémentaire de N_{m-1} dans N_m.On note G_i l'image de F_{i+1} par l'application $X \mapsto AX$. Il est clair que $G_i \subset N_i$ et que $G_i \cap N_{i-1} = \{0\}$ (sinon, il existe un vecteur non nul $X \in M_{i+1}$, $AX \in N_{i-1}$ et donc $X \in N_i$). On construit F_i comme étant un supplémentaire quelconque de N_{i-1} dans N_i qui contient G_i.

On a alors

$$K^n = F_1 \oplus ... \oplus F_k.$$

De plus, l'application $X \in F_k \mapsto AX \in F_{k-1}$ est injective pour tout $k \geqslant 2$ (en effet, si $AX = 0$ avec $X \in F_k$, alors $X \in N_1 \cap F_k \subset N_{k-1} \cap F_k = \{0\}$). Soit $(Y_{k,i})_{i \leqslant n_k}$ une base F_k. La famille $(Y_{k-1,i} = AY_{k,i})_{i \leqslant n_k}$ est dans F_{k-1} elle est libre. On peut la compléter en une base $(Y_{k-1,i} = AY_{k,i})_{i \leqslant n_{k-1}}$ de F_{k-1} ($n_{k-1} \geqslant n_k$). En recommençant cette opération $k-2$ fois, on construit pour tout $m \leqslant k$, une base

$(Y_{m,i})_{i\leqslant n_m}$ telle que pour tout $i \leqslant n_{m-1}$, $Y_{m,i} = AY_{m-1,i}$. On obtient ainsi une base $(Y_{m,i})_{1\leqslant m\leqslant k,\ i\leqslant n_m}$ de K^n. On peut réordonner cette base verticalement de la façon suivante

$$Z_1 = Y_{1,1},\ Z_2 = Y_{2,1}, .., Z_k = Y_{k,1},\ Z_{k+1} = Y_{1,2}, ..., Z_{2k} = Y_{k,2}, ...,$$

$$Z_{(i-1)k+j} = Y_{j,i} \text{ pour } 1 \leqslant i \leqslant n_k, 1 \leqslant j \leqslant k,$$

$$\begin{array}{rcl}
Z_{kn_k+1} & = & Y_{1,n_k+1},\ Z_{kn_k+2} = Y_{2,n_k+1}, \\
 & \vdots & \\
Z_{kn_k+k-1} & = & Y_{k-1,n_k+1}, \\
Z_{kn_k+k} & = & Y_{1,n_k+2}, \\
Z_{kn_k+k+1} & = & Y_{2,n_k+2}, \\
 & \vdots & \\
Z_{kn_k+2k-1} & = & Y_{k-1,n_k+2}, \\
Z_{kn_k+(i-1)(k-1)+j} & = & Y_{j,n_k+i} \text{ pour } 1 \leqslant i \leqslant n_{k-1} - n_k, 1 \leqslant j \leqslant k-1,
\end{array}$$

et ainsi de suite. Soit l'endormophisme de K^n dont la matrice est A dans la base canonique. La matrice de cette endormophisme dans la base $(Z_\ell)_{\ell\leqslant n}$ est de la forme (**??**). ■

2.7 Compléments

2.7.1 Théorème de Décomposition de Dunford pour les matrices

Théorème 10 *Soit $A \in \mathscr{M}_n(K)$ une matrice dont le polynôme minimal est scindé. Alors, il existe un couple de matrices (B, N) tel que*

1. *$A = B + N$,*
2. *B est diagonalisable et N est nilpotente,*
3. *B et N commutent.*

De plus, B et N sont données par

$$B = \lambda_1 P_1 + ... + \lambda_p P_p,\ N = (A - \lambda_1 I)P_1 + ... + (A - \lambda_p I)P_p,$$

où pour $\lambda_1, ..., \lambda_p$ sont les valeurs propres de A et $P_1, .., P_p$ sont les matrices des projections sur les sous-espaces caractéristiques $F_{\lambda_1}, ..., F_{\lambda_p}$.

Démonstration – Soit u l'endormorphisme associé à A dans la base canonique de K^n. Soient m_1,...,m_p les multiplicités algébriques des valeurs propres $\lambda_1, ..., \lambda_p$ respectivement.
On sait d'après la proposition (**??**) que $K^n = \oplus_{i=1}^p F_{\lambda_i}$. On sait que chacun des sous-espaces caractéristiques F_{λ_i} est stable par u et que pour tout $i \leqslant p$, $F_{\lambda_i} = \mathrm{Ker}\,(A - \lambda_i I)^{m_i}$ (proposition (**??**)) et $\dim F_{\lambda_i} = m_i$.
Soit v et n les deux endomorphismes définis par leurs restrictions v_i et n_i à F_{λ_i}, $i = 1, ..., p$, comme suit

$$\forall x \in F_{\lambda_i},\ v_i(x) = \lambda_i x,\ n_i(x) = (u_i - \lambda_i id)(x).$$

où u_i est la restriction de u à F_{λ_i}. Considérons une base $\{f_{i,1}, ..., f_{i,m_i}\}$ de F_{λ_i}. La famille $(f_{i,k})_{1\leqslant i\leqslant p,\ 1\leqslant k\leqslant m_i}$ est une base de $K^n = \oplus_{i=1}^p F_{\lambda_i}$.

Dans cette base, la matrice de v est clairement diagonale. On conclut que v est diagonalisable.
Par ailleurs, la restriction n_i de n à F_{λ_i}, $i = 1, ..., p$, est nilpotente car elle vérifie $n_i^{m_i} = 0$ ($F_{\lambda_i} = \mathrm{Ker}\,(A - \lambda_i I)^{m_i}$). Donc n est nilpotent ($n^m = 0$ où $m = \max(m_1, ..., m_p)$).
De plus, étant donné que pour tout $i \leqslant p$ n_i et v_i commutent au sein de F_{λ_i} (car ce sont tous les deux des polynômes en u_i), on en déduit que v et n commutent.
Les matrices B et N peuvent être choisies comme les matrices de v et n respectivement dans la base canonique de K^n.
L'unicité est admise. ■

Une application directe de cette décomposition est le calcul des puissances de la matrice A quelconque. En effet, pour tout $k \geqslant 0$, puisque B et N commutent, on a

$$A^k = (B+N)^k = \sum_{m=0}^{k} \left(\begin{array}{c} k \\ m \end{array} \right) B^{k-m} N^m.$$

Le calcul des puissances de la matrice diagonalisable B et de la matrice nilpotente N est facile.

2.7.2 Réduction de Jordan

On appelle matrice de Jordan toute matrice de la forme

$$J = \begin{pmatrix} \lambda & 1 & 0 & ... & 0 \\ 0 & \lambda & 1 & ... & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & ... & 0 & \lambda & 1 \\ 0 & ... & ... & 0 & \lambda \end{pmatrix}, \tag{18}$$

avec $\lambda \in \mathbb{C}$. On a le théorème suivant

Théorème 11 *Soit A une matrice de $\mathscr{M}_n(\mathbb{C})$. Alors A est semblable à une matrice de la forme*

$$\begin{pmatrix} \boxed{J_1} & 0 & 0 & ... & 0 \\ 0 & \boxed{J_2} & 0 & ... & 0 \\ ... & ... & .. & ... & ... \\ 0 & ... & 0 & \boxed{J_{m-1}} & 0 \\ 0 & ... & ... & 0 & \boxed{J_m} \end{pmatrix}. \tag{19}$$

où J_1,...,J_m sont des matrices de Jordan. De plus, A est diagonalisable si et seulement si les matrices J_k sont de taille 1×1.

Soulignons que les valeurs propres des matrices de Jordan J_1,...,J_m dans la factorisation (**??**) ne sont pas nécessairment deux à deux distinctes.
La réduction de Jordan peut aussi servir dans le calcul des puissances d'une matrice. En effet, on montre que si J est une matrice de Jordan J de la forme (**??**) et de taille m^2, alors pour tout $k \geqslant 1$, on a $J^k = ((d_{ij}^{(k)}))_{1 \leqslant i,j \leqslant m}$ avec

$$d_{ij}^{(k)} = \begin{cases} \left(\begin{array}{c} k \\ j-i \end{array} \right) \lambda^{k-j+i} & \text{si } 0 \leqslant j-i \leqslant k \\ 0 & \text{si } i > j \text{ ou } j-i \geqslant k+1 \end{cases}$$

Une autre application de la réduction de Jordan concerne les estimations sur le rayon spectral et la convergence des puissances d'une matrices.

2.7.3 Valeurs singulières

Définition 4 *Soit A une matrice quelconque de $\mathscr{M}_n(\mathbb{C})$. On appelle valeurs singulières de A les racines carrées positives des valeurs propres de la matrice A^*A (qui est hermitienne semi-définie positive).*

On les notera désormais $s_1(A), s_2(A), ..., s_n(A)$ avec la convention $0 \leqslant s_1(A) \leqslant ... \leqslant s_n(A)$ (ainsi $\mathrm{sp}(A^*A) = \{s_1^2(A), ..., s_n^2(A)\}$.
Quand la matrice A est hermitienne, ses valeurs singulières sont les valeurs absolues de ses valeurs propres.
Commençons par le résultat de réduction suivant

Proposition 13 *Soit $A \in \mathscr{M}_n(\mathbb{C})$ (resp. $A \in \mathscr{M}_n(\mathbb{R})$) et soient $s_1, ..., s_n$ ses valeurs singulières. Alors il existe deux matrices unitaires (resp. orthogonales) telles que*

$$UAV^* = \mathrm{diag(s_i)}.$$

DÉMONSTRATION – Soit $A \in \mathscr{M}_n(\mathbb{C})$. La matrice A^*A étant hemitienne, on peut introduire une matrice unitaire V^* telle que

$$A^*A = V\mathrm{diag(s_i^2)V^*}.$$

Posons $G = AV$ et soient C_i, $1 \leqslant i \leqslant n$, les vecteurs colonnes de G. On a

$$G^*G = V^*A^*AV = \mathrm{diag(s_i^2)}.$$

Il en résulte que pour tous $i,\ j \leqslant n$, on a

$$C_i^*C_j = s_i^2\delta_{i,j}.$$

Ainsi $C_i = 0$ si $s_i = 0$. Pour tout i tel que $s_i \neq 0$, on pose

$$Y_i = \frac{1}{s_i}C_i.$$

La famille $(Y_i)_{s_i\neq 0}$ est orthonormale ($Y_i^*Y_j = \delta_{i,j}$). On peut la compléter en une base orthonormale $(Y_i)_{1\leqslant i\leqslant n}$. La matrice U de vecteurs lignes Y_i^* est unitaire et vérifie $UAV^* = UG = \mathrm{diag(s_i)}$.
La preuve reste identique quand $A \in \mathscr{M}_n(\mathbb{R})$. ■

Proposition 14 *Soient $A,\ B \in \mathscr{M}_n(\mathbb{C})$. On a les propriétés suivantes*

$$\begin{aligned}
s_j(cA) &= |c|s_j(A),\ \ \forall j \in \{1, ..., n\},\ \ \forall c \in \mathbb{C}, && (20)\\
s_n(A) &= \|A\|_2, && (21)\\
s_j(A) &= s_j(A^*),\ \ \forall j \in \{1, ..., n\}, && (22)\\
s_j(BA) &\leqslant \|B\|_2 s_j(A), \forall j \in \{1, ..., n\}, \forall B \in \mathscr{M}_n(\mathbb{C}), && (23)\\
s_j(AB) &\leqslant \|B\|_2 s_j(A), \forall j \in \{1, ..., n\}, \forall B \in \mathscr{M}_n(\mathbb{C}), && (24)
\end{aligned}$$

DÉMONSTRATION –

1. Pour tout $c \in \mathbb{C}$, on a $(cA)^*(cA) = |c|^2 A^*A$ et $\mathrm{sp}_{\mathbb{C}}(|c|^2 A^*A) = \{|c|^2\lambda \mid \lambda \in \mathrm{sp}_{\mathbb{C}}(A^*A)\}$.
2. On sait d'après le théorème **??** appliqué à la matrice A^*A que

$$s_n^2(A) = \max_{X \in \mathbb{C}^n \setminus \{0\}} \frac{X^*A^*AX}{X^*X} = \max_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|AX\|_2^2}{\|X\|_2^2} = \|A\|_2^2.$$

   D'où $s_n(A) = \|A\|_2$.
3. D'après la Proposition **??**, il existe deux matrices unitaires U et V telles que $A = U^*\mathrm{diag(s_i)V}$. Ainsi $AA^* = U^*\mathrm{diag(s_i^2)U}$ et $A^*A = U^*\mathrm{diag(s_i^2)U}$. AA^* et A^*A ont donc les mêmes vecteurs propres.
4. Soit $M = \|B\|_2^2 A^*A - A^*B^*BA$. Pour tout $X\ in \mathbb{C}^n$, on a

$$X^*MX = \|B\|_2^2\|AX\|_2^2 - \|BAX\|_2^2 \geqslant 0.$$

   La matrice M est donc semi-définie positive et

$$\|B\|_2^2 A^*A = A^*B^*BA + M.$$

   D'après le corollaire **??**, on en déduit que

$$\forall j \leqslant n,\ \lambda_j(\|B\|_2^2 A^*A) \geqslant \lambda_j(A^*B^*BA),$$

   autrement dit

$$\|B\|_2^2 s_j(A)^2 \geqslant s_j(BA)^2,$$

   De même,

$$\|B\|_2^2 s_j(A^*)^2 = \|B^*\|_2^2 s_j(A^*)^2 \geqslant s_j(B^*A^*)^2 = s_j(AB)^2.$$

■

Théorème 12 *Soit $A \in \mathscr{M}_n(\mathbb{C})$. Alors A est normale si et seulement si*

$$s_j(A) = |\lambda_j(A)|,\ \forall j \in \{1, ..., n\}, \tag{25}$$

où $\lambda_1(A), ..., \lambda_n(A)$ sont les valeurs propres de A ordonnées par modules croissants.

Démonstration – Supposons que A est normale. A s'écrit alors sous la forme

$$A = U^*DU,$$

avec U une matrice unitaire et D une matrice diagonale. Il en résulte que

$$A^*A = U^*D^*UU^*DU = U^*D^*DU.$$

La matrice D^*D est une matrice diagonale dont les valeurs propres sont $|\lambda_1(A)|^2, ..., |\lambda_n(A)|^2$. D'où $s_j^2(A) = |\lambda_j(A)|^2$ pour tous $j \leqslant n$. La réciproque est admise. ■

3 Normes matricielles et rayon spectral

On s'intéresse ici aux normes des matrices, c'est-à-dire les normes définies sur les espaces $\mathscr{M}_{n,p}(K)$, $n \geqslant 1$, $p \geqslant 1$. Nous allons voir qu'il y a au moins deux façon de définir les normes sur cet espace.
La première façon consiste simplement à considérer une matrice $A = ((a_{i,j}))_{i \leqslant n, j \leqslant p}$ de $\mathscr{M}_{n,p}(K)$ comme un vecteur de K^{np}. Ainsi, on peut définir sur $\mathscr{M}_{n,p}(K)$ des normes analogues aux normes vectorielles. Considérons par exemple la norme dite de Frobenius (ou de Hilbert-Schmidt) définie par

$$\|A\|_F = \left(\sum_i^n \sum_{j=1}^p |a_{ij}|^2 \right)^{1/2} \quad (\text{où } a_{ij} = (A)_{i,j}).$$

Cette norme n'est autre que la norme hermitienne sur K^{np}. Elle est associée au produit scalaire

$$A : B = \sum_{i=1}^n \sum_{j=1}^p a_{ij} \overline{b_{ij}}. \tag{26}$$

On peut écrire aussi

$$A : B = \mathrm{Tr}(AB^*), \;\; \|A\|_F = \sqrt{\mathrm{Tr}(AA^*)} = \sqrt{\mathrm{Tr}(A^*A)}.$$

Notons que

$$\|I\|_F = \sqrt{n}.$$

On a la propriété suivante pour cette norme : pour tous $A \in \mathscr{M}_{n,p}(K)$ et $B \in \mathscr{M}_{p,m}(K)$

$$\|AB\|_F \leqslant \|A\|_F \|B\|_F. \tag{27}$$

Cette propriété sur le produit n'est pas satisfaite par toutes les normes définies sur la matrices. Si on se place dans l'espace $\mathscr{M}_n(K)$ des matrices carrées, on peut considérer la norme

$$\|A\|_{\max} = \max_{1 \leqslant i,\, j \leqslant n} |a_{ij}|.$$

Si on choisit

$$A = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

On a

$$\|A\|_{\max} = 1, \;\; \|A^2\|_{\max} = 2 > \|A\|_{\max}^2.$$

Dans toute la suite, on dira qu'une norme $\|.\|$ est matricielle si

$$\forall A, \;\; B \in \mathscr{M}_n(K), \;\; \|AB\| \leqslant \|A\| \|B\|.$$

Ainsi, la norme $\|.\|_F$ est matricielle, contrairement à la norme $\|.\|_{\max}$.

Un deuxième façon de construire des normes matricielles consiste à considérer toute matrice carrée $A = ((a_{i,j}))_{i,j\leqslant n}$ comme aussi une répésentation de l'endomorphisme $X \mapsto AX$ de K^n dans la base canonique. Ainsi, si $\|.\|$ est une norme *vectorielle* de K^n, on peut définir la norme matricielle suivante dite **subordonnée** à cette norme vectorielle (et notée identiquement)

$$\|A\| = \sup_{X\neq 0} \frac{\|AX\|}{\|X\|}.$$

En d'autres termes, $\|.\|$ est la plus petit constante $C > 0$ telle que l'inégalité

$$\|AX\| \leqslant C\|X\|,$$

soit vraie pour tout $X \in K^n$.
Dans la suite, on note $\|.\|_p$, $1 \leqslant p < +\infty$ la norme vectorielle

$$\forall X = (x_1, ..., x_n)^T,\ \|X\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

(le cas $p = 2$ correspond à la norme hermitienne).
On note $\|.\|_\infty$ la norme vectorielle définie par

$$\forall X = (x_1, ..., x_n)^T,\ \|X\|_\infty = \max_{1\leqslant i\leqslant n} |x_i|.$$

On note de même $\|.\|_p$ et $\|.\|_\infty$ les normes subordonnées à ces deux normes respectivement. La proposition suivante est facile à démontrer

Proposition 15 *Soit $\|.\|$ une norme de $\mathscr{M}_n(K)$ subordonnée à une norme vectorielle notée identiquement. Alors on a les propriétés suivantes*
- *Pour tout $A \in \mathscr{M}_n(K)$, $\|A\| = \sup_{X\in K^n,\ \|X\|=1} \|AX\|$.*
- *Pour tout $A \in \mathscr{M}_n(K)$, il existe au moins un élément $X_A \in K^n$ tel que $\|X_A\| = 1$ et*

$$\|A\| = \|AX_A\|.$$

- $\|I\| = 1$.

Proposition 16 *Soit $.\|$ une norme subordonnée à une norme vectorielle notée identiquement. Soit M une matrice carrée inversible. Alors l'application*

$$A \in \mathscr{M}_n(\mapsto)\|M^{-1}AM\|,$$

est une norme matricielle subordonnée à la norme vectorielle

$$\|X\|_M = \|M^{-1}X\|.$$

Démonstration – Il est évident que l'application $X \in K^n \mapsto \|X\|_M = \|M^{-1}X\|$ est une norme sur K^n. Calculons la norme matricielle subrodonnée à cette norme vectorielle. Soit $X \in K^n$ et posons $Y = M^{-1}X$. On a

$$\|X\|_M = \|Y\| \text{ et } \|AX\|_M = \|M^{-1}AX\| = \|M^{-1}AMM^{-1}X\| = \|M^{-1}AMY\|.$$

Puisque l'application $X \mapsto Y = M^{-1}X$ est une bijection de K^n dans lui même, on peut écrire

$$\sup_{X \in K^n \backslash 0} \frac{\|AX\|_M}{\|X\|_M} = \sup_{Y \in K^n \backslash 0} \frac{\|M^{-1}AMY\|}{\|Y\|} = \|M^{-1}AM\|.$$

Ce qui montre que l'application $A \in \mathscr{M}_n(K) \mapsto \|M^{-1}AM\|$ est la norme subordonnée à la norme $\|.\|_M$. ■

Définition 5 *Soit $A \in \mathscr{M}_n(\mathbb{C})$. On appelle* **rayon spectral** *de A le réel positif*

$$\rho(A) = \max_{\lambda \in \mathrm{sp}_{\mathbb{C}}(A)} |\lambda|. \tag{28}$$

Nous allons découvrir petit à petit l'importance du rayon spectral dans le calcul matriciel, notamment dans la convergence de la suite des puissances d'une matrice.
Commençons par souligner les propriétés suivantes, qui sont simples à démontrer

$$\forall A \in \mathscr{M}_n(\mathbb{C}),\ \forall \alpha \in \mathbb{C},\ \rho(\alpha A) = |\alpha|\rho(A).$$

$$\forall A \in \mathscr{M}_n(\mathbb{C}),\ \forall k \in \mathbb{N},\ \rho(A^k) = \rho(A)^k.$$

Cette dernière identité résulte du fait que $\mathrm{sp}_{\mathbb{C}}(A^k) = \{\lambda^k\ ;\ \lambda \in \mathrm{sp}_{\mathbb{C}}(A)\}$.
Par ailleurs, notons que deux matrices semblables ont le même rayon spectral.
Il est important de noter que le rayon spectral n'est pas une norme sur les matrices. Toutefois, si $\rho(A) = 0$, cela signifie que toutes la valeurs propres de A sont nulles. Le polynôme caractéristique de A est donc de la forme $(-1)^n X^n$. Il en résulte que $A^n = 0$ et A est forcément nilopotente. La réciproque est vraie aussi. D'où la proposition

Proposition 17 *Soit $A \in \mathscr{M}_n(\mathbb{C})$. Alors $\rho(A) = 0$ si et seulement si A est nilpotente.*

En particulier, on peut avoir $\rho(A) = 0$ sans que A soit nulle. Il suffit de considérer la matrice

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix},$$

pour laquelle on a $\rho(A) = 0$ pourtant $A \neq 0$.

Nous allons maintenant élucider le lien entre le rayon spectral et les normes de matrices. Commençons tout d'abord par la la caractérisation de la norme subordonnée $\|.\|_2$

Proposition 18 *Soit $A \in \mathscr{M}_n(\mathbb{C})$. Alors*

$$\|A\|_2 = \sqrt{\rho(A^*A)}.$$

C'est aussi la plus grande valeur singulière de A.

Démonstration – C'est une conséquence directe du Théorème **??**.
■

La proposition suivante donne une estimation des normes matricielles subordonnées en fonction du rayon spectrale

Proposition 19 *Soit $\|.\|$ une norme matricielle sur $\mathscr{M}_n(\mathbb{C})$. Alors, pour toute matrice $A \in \mathscr{M}_n(\mathbb{C})$ on a*

$$\rho(A) \leqslant \|A\|.$$

Réciproquement, pour toute matrice $A \in \mathscr{M}_n(\mathbb{C})$ et tout réel $\epsilon > 0$, il existe au moins une norme subordonnée $\|.\|$ dépendant de A et de ϵ telle que

$$\|A\| \leqslant \rho(A) + \epsilon.$$

Démonstration – Soit λ une valeur propre de A et $X \in \mathbb{C}^n$ tel que $AX = \lambda X$ et $X \neq 0$. La matrice carrée XX^* est non nulle car elle comporte au moins un élément non nul sur sa diagonale. On a

$$A.(XX^*) = \lambda(XX^*).$$

Donc

$$\|\lambda(XX^*)\| = \|A.(XX^*)\| \leqslant \|A\|\|XX^*\|.$$

Ce qui donne

$$|\lambda| \leqslant \|A\|,$$

et donc $\rho(A) \leqslant \|A\|$.
La réciproque est admise. ■

Théorème 13 *Soit $A \in \mathscr{M}_n(\mathbb{C})$. Alors*

$$\lim_{k\to+\infty} A^k = 0 \Longleftrightarrow \rho(A) < 1.$$

Démonstration – Soit $\|.\|$ une norme matricielle quelconque si $\rho(A) \geqslant 1$ et vérifiant

$$\|A\| \leqslant \rho(A) + \frac{1-\rho(A)}{2} < 1.$$

si $\rho(A) < 1$. Si $\lim_{k\to+\infty} A^k = 0$, alors $\lim_{k\to+\infty} \|A^k\| = 0$. Nécessairement $\lim_{k\to+\infty} \rho(A^k) = 0$ car $\rho(A^k) \leqslant \|A^k\|$ pour tout $k \geqslant 0$. D'où $\lim_{k\to+\infty} \rho(A)^k = 0$ (puisque $\rho(A^k) = \rho(A)^k$). Donc $\rho(A) < 1$.
Inversement, si $\rho(A) < 1$ alors

$$0 \leqslant \|A^k\| \leqslant \|A\|^k \leqslant \left(\rho(A) + \frac{1-\rho(A)}{2}\right)^k.$$

Par passage à la limite quand $k\to+\infty$ on en déduit que $\lim_{k\to+\infty} A^k = 0$.

■

Corollaire 7 (admis) *Pour tout matrice $A \in \mathscr{M}_n(\mathbb{C})$, on a*

$$\lim_{k\mapsto+\infty} \|A^k\|^{1/k} = \rho(A).$$

Théorème 14 *Soit B une matrice carrée telle que*

$$\rho(B) < 1.$$

Alors les matrices carrées $I - B$ et $I + B$ sont inversible. De plus

$$(I - B)^{-1} = I + B + B^2 + ... \tag{29}$$

De plus, pour toute norme subordonnée $\|.\|$ telle que $\|B\| < 1$, on a

$$\|(I - B)^{-1}\| \leqslant \frac{1}{1 - \|B\|}. \tag{30}$$

DÉMONSTRATION – Le spectre $\mathrm{sp}(I - B) = \{1 - \mu \mid \mu \in \mathrm{sp}(B)\}$ ne contient pas 0. Donc $I - B$ est inversible. De plus, pour tout $m \geqslant 0$, on a

$$(I - B)(I + B + ... + B^m) = I - B^{m+1}.$$

Il en résulte que

$$(I - B)^{-1} - (I + B + ... + B^m) = (I - B)^{-1} B^{m+1}.$$

On a alors pour toute norme matricielle subordonnée telle que $\|B\| <$ 1

$$\|(I-B)^{-1} - (I+B+...+B^m)\| = \|(I-B)^{-1}B^{m+1}\| \leqslant \|(I-B)^{-1}\|\|B\|^{m+1}.$$

En passant à la limite quand m tend vers $+\infty$, on en déduit que

$$\lim_{m \mapsto \infty} (I - B)^{-1} - (I + B + ... + B^m) = 0.$$

Autrement dit $(I - B)^{-1} = \sum_{k=1}^{+\infty} B^k$. De plus, pour tout $m \geqslant 0$,

$$\begin{aligned} \|(I - B)^{-1}\| &\leqslant \|(I - B)^{-1} - (I + B + ... + B^m)\| + \|(I + B + ... + B^m)\| \\ &\leqslant \|(I - B)^{-1}\|\|B\|^{m+1} + \|I\| + ...\|B\|^m, \\ &\leqslant \|(I - B)^{-1}\|\|B\|^{m+1} + \frac{1 - \|B\|^{m+1}}{1 - \|B\|}. \end{aligned}$$

En faisant tendre vers $+\infty$ on obtient (**??**). ■