Université de Versailles-Saint Quentin
Licence de Mathématiques - Cours d'Analyse Numérique - Année - 2016/2017
Christine Poirier

Dans tout ces rappels, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

1 Rappels sur les Systèmes et Applications Linéaires :

Propriété 1.1. *Soient E et F deux espaces vectoriels sur $\mathbb{K}$ de dimension finie. Si $E \subset F$ et $\dim E = \dim F$, alors les deux espaces E et F sont égaux.*

Définition 1.1. Soient $E_1, .., E_p$ des sous-espaces vectoriels de E. On désigne par $E_1 + \cdots + E_p$ le sous-espace vectoriel de E tel que

$$E_1 + \cdots + E_p \;=\; \{x \in E \;\text{ tq. }\; x \;=\; x_1 + \cdots + x_p \;\text{ avec }\; x_1 \in E_1, \cdots, x_p \in E_p.\}$$

On dit que la somme $E_1 \oplus \cdots \oplus E_p$ est directe si

$$\forall (x_1, \cdots, x_p) \in E_1 \times \cdots \times E_p, \quad x_1 + \cdots + x_p \;=\; 0 \quad \Rightarrow \quad x_1 = 0 \;\text{ et }\; \cdots \;\text{ et }\; x_p = 0.$$

Propriété 1.2. *Soient $E_1, \cdots E_p$ des sous-espaces vectoriels de E, E étant un espace vectoriel sur $\mathbb{K}$ de dimension finie. On a toujours*

$$\dim(E_1 + \cdots + E_p) \;\leq\; \sum_{i=1}^{p} \dim(E_i).$$

De plus la somme $E_1 \oplus \cdots \oplus E_p$ est directe si et seulement si $\dim(E_1 + \cdots + E_p) = \sum_{i=1}^{p} \dim(E_i)$.

Preuve. Pour chaque $i \in \{1, \cdots, p\}$, on note $(e_1^i, \cdots, e_{n_i}^i)$ une base de E_i et on a alors $\dim(E_i) = n_i$. On pose $F = E_1 + \cdots + E_p$. La famille $(e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p)$ est une famille génératrice de F donc $\dim(E_1 + \cdots + E_p) \leq \sum_{i=1}^{p} \dim(E_i)^{(*)}$.

1. Si la somme $F = E_1 \oplus \cdots \oplus E_p$ est directe alors $(e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p)$ est une famille libre . D'où $\dim(F) \geq \sum_{i=1}^{p} \dim(E_i)^{(**)}$. Donc d'après $(*)$ et $(**)$ on en déduit que $\dim(E_1 \oplus \cdots \oplus$

$E_p) \ = \ \sum_{i=1}^{p} \dim(E_i)$.

2. Réciproquement : Posons $n = \sum_{i=1}^{p} \dim(E_i)$. Si $\dim(F) \ = \ n$ alors la famille $(e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p)$ est une famille génératrice de F qui contient n éléments et $\dim(F) = n$ donc on en déduit que $(e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p)$ est une base de F donc en particulier est libre. Donc la somme $F = E_1 + \cdots + E_p$ est directe. □

Propriété 1.3. *Soient $E_1, \cdots E_p$ des sous-espaces vectoriels de E, E étant un espace vectoriel sur $\mathbb{K}$ de dimension finie. Pour chaque $i \in \{1, \cdots, p\}$, on note $(e_1^i, \cdots, e_{n_i}^i)$ une base de E_i. Alors*

$$E \ = \ E_1 \oplus \cdots \oplus E_p \quad \Longleftrightarrow \quad \mathcal{F} = (e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p) \ \textit{ est une base de } \ E.$$

Preuve. La somme est directe si et seulement si la famille $\mathcal{F}$ est libre (cf. définition d'une somme directe et d'une famille libre).
$E \ = \ E_1 + \cdots + E_p$ si et seulement si $\mathcal{F}$ est une famille génératrice de E (cf. définition d'une famille génératrice et d'une somme de sous-espaces vectoriels).
On obtient ainsi le résultat de la propriété. □

Propriété 1.4. *Soient $E_1, \cdots E_p$ des sous-espaces vectoriels de E, E étant un espace vectoriel sur $\mathbb{K}$ de dimension finie.* **On suppose que la somme $E_1 \oplus \cdots \oplus E_p$ est directe** *alors*

$$E = E_1 \oplus \cdots \oplus E_p \quad \Longleftrightarrow \quad \dim(E) \ = \ \sum_{i=1}^{p} \dim(E_i).$$

Preuve. Si $E = E_1 \oplus \cdots \oplus E_p$ alors $(e_1^1, \cdots, e_{n_1}^1, \cdots, e_1^p, \cdots, e_{n_p}^p)$ est une base de E donc $\dim(E) = n_1 + \cdots + n_p = \sum_{i=1}^{p} \dim(E_i)$.
Réciproquement, posons $F = E_1 \ \oplus \ \cdots \ \oplus \ E_p$, on a alors $F \subset E$ et d'après la propriété **1.2** $\dim(F) = \sum_{i=1}^{p} \dim(E_i)$. Or $\sum_{i=1}^{p} \dim(E_i) = dim(E)$ donc d'après la propriété **1.1**, $E = F$. □

1.1 Quelques Rappels sur les Applications Linéaires

Définition 1.2. Soient E et F deux espaces vectoriels sur $\mathbb{K}$. On dit que T, application de E dans F est **linéaire** si

$$\begin{cases} \forall x, y \in E & f(x+y) = f(x) + f(y) \\ \forall \lambda \in \mathbb{K} \ \forall x \in E & f(\lambda \cdot x) = \lambda \cdot f(x) \end{cases}$$

Si T est une application linéaire de E dans E alors on dit que T est un **endomorphisme** de E.

Soit T une application linéaire de E dans F, on définit les deux espaces vectoriels suivants :

$\mathrm{Ker}(T) = \{u \in E \text{ tq. } T(u) = 0_F\}$	$\mathrm{Ker}(T)$ s.e.v. de E, est appelé **Noyau de** T
$\mathrm{Im}(T) = \{T(u) \text{ tq. } u \in E\}$	$\mathrm{Im}(T)$ s.e.v. de F, est appelé **Image de** T

Propriété 1.5. *Si E est de dimension finie, dont une base est $e_1, .., e_n$ et si T est une application linéaire de E dans F alors*

$$Im(T) = Vect\{T(e_1), .., T(e_n)\},$$

où $Vect\{a_1, .., a_n\}$ désigne le sous-espace vectoriel engendré par les vecteurs $a_1, .., a_n$.

Propriété 1.6. *Soit T une application linéaire de E dans F. On a les propriétés suivantes*

T est injective $\iff$ $Ker(T) = \{0_E\}$.
T est surjective $\iff$ $Im(T) = F$.
T est bijective $\iff$ T est injective et surjective.

Théorème 1.1. du rang *Soient E et F deux espaces vectoriels de dimension finie. Soit T une application linéaire de E dans F. Alors*

$$\dim Ker(T) + \dim Im(T) = \dim E.$$

Remarque 1.1.
- Rappelons que le rang d'une famille de vecteurs $\{u_1, ..u_p\}$ est égal à $\dim \text{Vect}\{u_1, ..u_p\}$, c'est-à-dire au nombre maximum de vecteurs linéairement indépendants que l'on peut extraire de la famille $\{u_1, ..u_p\}$.
- $\dim \text{Im}(T)$ est encore appelée rang de T et notée $\text{rg}(T)$. Elle correspond au rang de la famille $\{T(e_1), ..T(e_n)\}$ où $\{e_1, .., e_n\}$ est une base de E.

Corollaire 1.1. *Soient E et F deux espaces vectoriels de dimension finie. Soit T une application linéaire de E dans F.*
Si $\dim E \neq \dim F$, alors T n'est pas bijective.
Si $\dim E = \dim F$, alors les conditions suivantes sont équivalentes :
(i) T est bijective
(ii) T est injective (c'est-à-dire $Ker(T) = \{0_E\}$)
(iii) T est surjective (c'est-à-dire $rg(T) = \dim(F)$).

Preuve. Il suffit d'utiliser le théorème du rang (cf. théorème 1.1) et remarquer que d'une part $\dim \text{Ker}(T) = 0 \Rightarrow \text{Ker}(T) = \{0_E\}$ et que d'autre part, d'après la propriété **??**, $\dim(\text{Im}(T)) = \dim F \Rightarrow \text{Im}(T) = F$. $\square$

Remarque 1.2.
– Notons que l'on a toujours $\text{Im}(T) \subset F$ et que $\text{Im}(T)$ est un sous-espace vectoriel de F. Donc si F est de dimension finie, $\dim(\text{Im}(T)) = \dim(F) \Rightarrow \text{Im}(T) = F$.
Attention : Notons aussi que si A et B sont deux sous-espaces vectoriels d'un espace vectoriel sur $\mathbb{K}$ alors $\dim(A) = \dim(B) \not\Rightarrow A = B$.

1.2 Résolution de $Tx = b$

Soient E et F deux espaces vectoriels de dimension finie, T une application linéaire de E dans F et b un élément de F. On désire chercher les vecteurs x de E tels que $T(x) = b$. On note S l'ensemble de ces vecteurs.

(i) Si T est bijective alors

$$\forall b \in F, \text{ il existe une unique solution } x_0 \in E \text{ tq. } T(x_0) = b.$$

(ii) Si T n'est pas bijective :

– Si $\underline{b \in \mathrm{Im}(T)}$, alors il existe au moins un $x_0 \in E$ tel que $T(x_0) = b$, donc $T(x) = b \Leftrightarrow T(x) = T(x_0) \Leftrightarrow T(x - x_0) = 0_F \Leftrightarrow x - x_0 \in \mathrm{Ker}(T)$.

$$S = \{x_0 + u \text{ tq. } u \in \mathrm{Ker}(T)\}.$$

Donc si T est injective alors $Tx = b$ admet une unique solution.
Si T n'est pas injective, alors $Tx = b$ admet une infinité de solutions.

– Si $\underline{b \notin \mathrm{Im}(T)}$ alors $S = \emptyset$.

1.3 Application à la Résolution des Systèmes Linéaires

Soient $n, p \in \mathbb{N}^\star$. On s'intéresse au système linéaire suivant :

$$(S) \quad \left\{ \begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & = & \vdots \\ a_{p1}x_1 & + & \cdots & + & a_{pn}x_n & = & b_p \end{array} \right. .$$

Résoudre le système (S) revient à résoudre $Tx = b$ où

$$\left\{ \begin{array}{l} b = (b_1, .., b_p) \in \mathbb{R}^p \\ \text{et } T \text{ est l'application linéaire de } \mathbb{R}^n \text{ dans } \mathbb{R}^p \text{ telle que} \\ T(x_1, .., x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \cdots, a_{p1}x_1 + \cdots + a_{pn}x_n). \end{array} \right. .$$

Propriété 1.7. *Le système (S) admet une solution et une seule quelque soit le second membre $b \in \mathbb{R}^p$ si et seulement si $n = p$ et l'une des trois conditions est vérifiée :*

(i) **Le système homogène associé à** (S) **(c'est-à-dire lorsque** $b = 0$**) n'admet que la solution** $x = 0$.

(ii) Pour tout second membre $b \in \mathbb{R}^p$, **le système admet toujours au moins une solution**.

(iii) **Le rang de la matrice** $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix}$ **est égal à** p.

On notera de plus, que si l'une des trois conditions est réalisée alors les deux autres le sont automatiquement.

Preuve. Le système (S) admet une unique solution quelque soit le second membre b si et seulement si T est bijective. Or, d'après le corollaire (1.1), une condition nécessaire pour que T soit bijective est que $\dim(\mathbb{R}^n) = \dim(\mathbb{R}^p)$ c'est-à-dire $n = p$. **On dit que le nombre d'inconnues doit être égal au nombre d'équations**.
Supposons maintenant que $n = p$. Toujours d'après le corollaire (1.1), T est alors bijective si et seulement si T est injective (c'est-à-dire $\mathrm{Ker}(T) = \{0\}$) si et seulement si T est surjective (c'est-à-dire $\mathrm{Im}(T) = \mathbb{R}^p$). Il suffit donc de vérifier l'une des deux propriétés, à savoir $\mathrm{Ker}(T) = \{0\}$ ou $\mathrm{Im}(T) = \mathbb{R}^p$, pour que l'autre soit alors automatiquement vérifiée.
<u>Comment interpréter $\mathrm{Ker}(T) = \{0\}$?</u>

$$\begin{array}{rcl} \mathrm{Ker}(T) = \{0\} & \Leftrightarrow & \text{“}T(x) = 0 \Rightarrow x = 0\text{”} \\ & \Leftrightarrow & \text{“} \underbrace{\left. \begin{array}{l} a_{11}x_1 + .. + a_{1n}x_n = 0 \\ \vdots \\ a_{p1}x_1 + .. + a_{pn}x_n = 0 \end{array} \right\}}_{\text{Système homogène associé à } (S)} \Rightarrow (x_1, .., x_n) = (0, .., 0).\text{”} \\ & \Leftrightarrow & \textbf{Système homogène associé à } (S) \textbf{ n'admet que la solution nulle.} \end{array}$$

Comment interpréter $\mathrm{Im}(T) = \mathbb{R}^p$?
$\mathrm{Im}(T) = \mathbb{R}^p$ signifie que $\forall b \in \mathbb{R}^p,\ b \in \mathrm{Im}(T)$ c'est-à-dire que $\forall b \in \mathbb{R}^p,\ \exists x \in \mathbb{R}^n,\ \text{tq. } T(x) = b$. Autrement dit : **Le système (S) admet au moins une solution quelque soit le second membre.**

D'après la remarque (1.2), $\mathrm{Im}(T) = \mathbb{R}^p \Leftrightarrow \dim(\mathrm{Im}(T)) = p \Leftrightarrow \mathrm{rg}(T) = p \Leftrightarrow \mathrm{rg}(\mathrm{Vect}\{T(e_1), T(e_2), .., T(e_n)\}) = p$ où $(e_1, .., e_n)$ désigne la base canonique de $\mathbb{R}^n$ ($e_1 = (1, 0, .., 0)$, $e_2 = (0, 1, 0, .., 0)$,..,$e_n = (0, .., 0, 1)$).
Or $T(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{p1} \end{pmatrix}, .., T(e_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{pn} \end{pmatrix}$.

Donc le $\mathrm{rg}(\mathrm{Vect}\{T(e_1), T(e_2), .., T(e_n)\}) = \mathrm{rg}(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix})$. $\square$

2 Réductions de Matrices :

$\mathbb{K}$ désigne le corps $\mathbb{C}$ ou $\mathbb{R}$.

2.1 Rappels et Notations

$\mathbb{R}^n$ désigne l'espace des vecteurs colonnes à n composantes réelles ($\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, où $x_i \in \mathbb{R}$).
$\mathbb{C}^n$ désigne l'espace des vecteurs colonnes à n composantes complexes.
x_{ij} désigne la i^{ieme} composante du vecteur x_j.
$(e_1, .., e_n)$ désigne la base canonique de $\mathbb{R}^n$ ou $\mathbb{C}^n$: $e_{ij} = \delta_{ij}$, $\forall i, j \in \{1, .., n\}$.
$\mathbb{R}^{n,p}$ désigne l'espace des matrices à n lignes et p colonnes à coefficients réels.
$\mathbb{C}^{n,p}$ désigne l'espace des matrices à n lignes et p colonnes à coefficients complexes.
Si $A \in \mathbb{K}^{n,p}$ alors A_{ij} désigne le coefficient de A qui est sur la i^{ime} ligne et la j^{ime} colonne.
I_n désigne la matrice **identité** d'ordre n : $I_n \in \mathbb{R}^{n,n}$ telle que $(I_n)_{ij} = \delta_{ij}$.
Si $A \in \mathbb{K}^{n,p}$ et $x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p$ alors $Ax \in \mathbb{K}^n$ et la $i^{\text{ième}}$ composante de Ax est

$$(Ax)_i \;=\; \sum_{k=1}^{n} A_{ik} x_k, \quad 1 \leq i \leq n.$$

Enfin si $A \in \mathbb{K}^{n,p}$ et $B \in \mathbb{K}^{p,r}$ alors $AB \in \mathbb{K}^{n,r}$ et

$$\forall i \in [|1:n|],\ \forall j \in [|1:r|], \qquad (AB)_{ij} \;=\; \sum_{k=1}^{p} A_{ik} B_{kj}.$$

On peut aussi définir A comme

$$\boxed{A = [A(e_1), \cdots, A(e_i), \cdots, A(e_p)], \quad \text{où } (e_1, \cdots, e_p) \text{ est la base canonique de } \mathbb{K}.}$$

Définition 2.1. Norme d'un espace vectoriel
Soit E un $\mathbb{K}$ espace vectoriel. Une norme N sur E est une application de E dans $\mathbb{R}^+$ vérifiant les 3 propositions suivantes :

- **i)** $N(\lambda x) = |\lambda| N(x) \quad \forall x \in E, \quad \lambda \in \mathbb{K}$
- **ii)** $N(x+y) \;\leq\; N(x) + N(y) \quad \forall x, y \in E$
- **iii)** $\forall x \in E,\ (N(x) = 0 \;\Leftrightarrow\; x = 0)$.

(E, N) est alors appelé **espace vectoriel normé**.

Propriété 2.1. *Si E et de dimension finie, alors toutes les normes de E sont équivalentes : c'est-à-dire que pour toutes normes N et N', il existe deux constantes réelles positives C_1 et C_2 telles que :*

$$\forall x \in E \quad C_1 N(x) \;\leq\; N'(x) \;\leq\; C_2 N(x).$$

Définition 2.2. Soit E un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On dit que l'application $(.,.) : E \times E \to \mathbb{K}$ est un **produit scalaire** si les propriétés suivantes sont vérifiées :

$$\begin{array}{lll}
(a) & \forall(\alpha,\beta) \in \mathbb{K}^2,\ \forall(x,y,z) \in E^3, & \\
 & (\alpha \cdot x + \beta \cdot y, z) = \alpha(x,z) + \beta(y,z) & \text{(on dit que } (.,.) \text{ est linéaire à gauche)} \\
 & (z, \alpha \cdot x + \beta \cdot y) = \overline{\alpha}(z,x) + \overline{\beta}(z,y) & \text{(on dit que } (.,.) \text{ est semi-linéaire à droite)} \\
(b) & \forall(x,y) \in E^2,\ (x,y) = \overline{(y,x)} & \text{(on dit que } (.,.) \text{ est hermitienne)} \\
(c) & \forall x \in E,\ (x,x) \geq 0 & \text{(on dit que } (.,.) \text{ est positive)} \\
(d) & \forall x \in E,\ (x,x) = 0 \Leftrightarrow x = 0 & \text{(on dit que } (.,.) \text{ est définie)}
\end{array}$$

• Si E est muni d'un produit scalaire $(.,.)$, on dit que $(E,(.,.))$ est un **espace préhilbertien**.
• Si $\dim(E)$ est finie, et $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$), on parle alors d'**espace euclidien** (resp. hermitien).

Remarque 2.1. Les physiciens remplacent la définition (a) par $(.,.)$ est linéaire à droite et semi-linéaire à gauche.

Propriété 2.2. *Un espace préhilbertien $(E,(.,.))$ est un espace normé, muni de la norme $\|.\|$ suivante :*

$$\forall x \in \mathbb{K} \quad \|x\| \ = \ \sqrt{(x,x)}.$$

Définition 2.3. On appelle **famille orthonormée** d'un espace pré-hilbertien E (de produit scalaire noté $(.,.)$) toute famille $\mathcal{B}$ de E telle que :

$$(x,y) = 0\ \forall x,y \in \mathcal{B} \ \text{ tq. } x \neq y, \ \text{ et } \|x\| = 1\ \forall x \in \mathcal{B}. \tag{1}$$

Propriété 2.3. *Toute famille finie orthonormée (cad vérifiant (1)) d'un espace préhilbertien E est une famille libre.*

Preuve. Si $\sum_{i=1}^{n} \alpha_i u_i = 0$ alors $\forall j \in [|1 : n|]$, $(\sum_{i=1}^{n} \alpha_i u_i, u_j) = 0$. Or $(\sum_{i=1}^{n} \alpha_i u_i, u_j) = \sum_{i=1}^{n} \alpha_i (u_i, u_j) = \alpha_j \times 1$. Donc $\forall j \in [|1 : n|],\ \alpha_j = 0$. □

Remarque 2.2. Si $(e_1,..,e_n)$ est une base orthonormée de E alors $\forall x \in E,\ (x,e_i)$ est la i ième coordonnée de x dans cette base orthonormée.

Théorème 2.1. Inégalité de Cauchy-Schwartz $(\star)$ *:*
$(E,(.,.))$ *est un espace préhilbertien et* $\|.\|$ *la norme associée au produit scalaire* $(.,.)$. *Alors*

$$\forall x,y \in E \quad |(x,y)| \ \leq \ \|x\|\ \|y\|.$$

Preuve. On suppose d'abord que (x,y) est réel et on considère pour $t \in \mathbb{R}$, $P(t) = \|x+ty\|^2$. Le discriminant étant toujours positif, on obtient le résultat.
Si (x,y) n'est pas réel, le produit scalaire $(x,(x,y)y) = \overline{(x,y)}(x,y)$ est réel positif et on utilise le cas précédent. □

Exemple 2.1. du produit scalaire usuel sur $\mathbb{K}^n$:

Soient x et $y \in \mathbb{C}^n$: $\quad (x,y) \ = \ \sum_{i=1}^{n} x_i \overline{y_i}$

Si $x,y \in \mathbb{R}^n$, $\quad (x,y) \ = \ \sum_{i=1}^{n} x_i y_i$

On notera $\|\cdot\|_2$ la norme associée au produit scalaire usuel : $\|x\|_2 = (x,x)^{1/2} \quad \forall x \in \mathbb{C}^n$.

Si $A \in \mathbb{C}^{n,p}$, on note $A^\star$, la **matrice adjointe** de A :

$$A^\star \in \mathbb{C}^{p,n}, \quad A^\star_{ij} = \overline{A_{ji}}, \quad 1 \leq i \leq p,\ 1 \leq j \leq n. \tag{2}$$

Propriété 2.4. *$\forall A \in \mathbb{C}^{n,p}$, $\quad (A^\star)^\star = A$ et si $A, B \in \mathbb{C}^{n,n} \quad (AB)^\star = B^\star A^\star$.*
D'autre part, si $A \in \mathbb{C}^{n,p}$, $A^\star$ est l'unique matrice $\in \mathbb{C}^{p,n}$, telle que

$$(Ax, y)_n = (x, A^\star y)_p,\ \forall x \in \mathbb{C}^p,\ \forall y \in \mathbb{C}^n. \tag{3}$$

L'égalité (3) généralise la définition de l'adjoint d'une matrice à tout opérateur linéaire continu sur un espace préhilbertien complet (: appelé espace de Hilbert).

Preuve.
$$(Ax, y)_n = \sum_{i=1}^{n}(Ax)_i\overline{y_i} = \sum_{i=1}^{n}\sum_{k=1}^{p} A_{ik}x_k\overline{y_i} = \sum_{k=1}^{p} x_k(\sum_{i=1}^{n}\overline{A^\star_{ki}}\overline{y_i}) = \sum_{k=1}^{p} x_k\overline{(A^\star y)_k} = (x, A^\star y)_p.$$
D'où l'égalité (3).
D'autre part, soit $B \in \mathbb{C}^{p,n}$ telle que $(x, By) = (Ax, y)$, $\forall x \in \mathbb{C}^p$, $\forall y \in \mathbb{C}^n$ alors $(x, (B - A^\star)y) = 0$, $\forall x \in \mathbb{C}^p$, $\forall y \in \mathbb{C}^n$. Alors si on choisit $x = (B - A^\star)y$ on obtient $\|(B - A^\star)y\|_2^2 = 0$ d'où $(B - A^\star)y = 0, \forall y \in \mathbb{C}^n$, ce qui signifie que $B = A^\star$. Donc la matrice $A^\star$ qui vérifie (3) est unique. $\square$

Si $A \in \mathbb{C}^{n,p}$, on note A^t, la **matrice transposée** de A :

$$A^t \in \mathbb{C}^{p,n}, \quad A^t_{ij} = A_{ji}, \quad 1 \leq i \leq p,\ 1 \leq j \leq n.$$
$$(Ax, y) = (x, A^t y) \quad \forall A \in \mathbb{R}^{n,p}, \quad \forall x \in \mathbb{R}^p, \quad \forall y \in \mathbb{R}^n.$$
$$\text{Si } A \in \mathbb{R}^{n,p}, \quad A^t = A^\star.$$

Si $x \in \mathbb{C}^n$, $x^\star \in \mathbb{C}^{1,n}$ est un vecteur ligne et $\quad (x, y) = x^t\, \overline{y} = y^\star x, \quad \forall x, y \in \mathbb{C}^n$.
Si $x \in \mathbb{R}^n$, $x^t \in \mathbb{R}^{1,n}$ est un vecteur ligne et $\quad (x, y) = x^t\, y, \quad \forall x, y \in \mathbb{R}^n$.

Définition 2.4.
Une matrice carrée $A \in \mathbb{C}^{n,n}$ est dite **hermitienne** si $A^\star = A$.
Une matrice carrée $A \in \mathbb{C}^{n,n}$ est dite **symétrique** si $A^t = A$.
Une matrice carrée $A \in \mathbb{C}^{n,n}$ est dite **normale** si $AA^\star = A^\star A$.
Une matrice carrée $U \in \mathbb{C}^{n,n}$ est dite **unitaire** si $UU^\star = U^\star U = I_n$.
Une matrice carrée $Q \in \mathbb{R}^{n,n}$ est dite **orthogonale** si $QQ^t = Q^t Q = I_n$.

Propriété 2.5. des matrice triangulaire
On rappelle qu'une matrice carrée T est **triangulaire supérieure** *si sa partie inférieure est nulle, c'est-à-dire :*

$$T \text{ est triangulaire supérieure} \iff T_{ij} = 0 \qquad \forall i, j \ \text{ tq. } \ i > j.$$

et qu'une matrice carrée T est **triangulaire inférieure** *si sa partie supérieure est nulle, c'est-à-dire*

$$T \text{ est triangulaire inférieure} \iff T_{ij} = 0 \qquad \forall i, j \ \text{ tq. } \ i < j.$$

i) *Une matrice T carrée triangulaire supérieure (resp. inférieure) est inversible si et seulement si ses éléments diagonaux sont non-nuls et alors T^{-1} est encore triangulaire supérieure (resp. inférieure) dont les éléments diagonaux sont les inverses des éléments diagonaux de T.*

ii) *Les valeurs propres d'une matrice triangulaire sont les coefficients diagonaux de cette matrice.*

iii) *Une matrice triangulaire hermitienne est diagonale, à diagonale réelle.*

iv) *Le déterminant d'une matrice triangulaire par blocs est égal au produit des déterminants des blocs diagonaux.*

Preuve. **i)** et **ii)** sont laissés en exercice.
iii) Si T est triangulaire supérieure alors T^* est triangulaire inférieure. Si de plus $T = T^*$, T est triangulaire inférieure et supérieure donc T est diagonale. $(T^*)_{ii} = T_{ii} \Rightarrow \overline{T_{ii}} = T_{ii}$. Donc T_{ii} est réel.
iv)

a) on traite d'abord le cas de 2 blocs :

Soit $T = = \left(\begin{array}{c|c} A_1 & B_1 \\ \hline 0 & A_2 \end{array}\right)$ où $A_1 \in \mathbb{C}^{p,p}$, $A_2 \in \mathbb{C}^{n,n}$, $B_1 \in \mathbb{C}^{p,n}$.

i) Le cas $p = 1$ est immédiat, car il suffit de développer le déterminant de T par rapport à la 1ière colonne.

ii) Le cas $p > 1$ peut se traiter par récurrence toujours en développant la matrice T de taille $p+1+n$ par rapport à la 1ière colonne et en appliquant l'hypothèse de récurrence à chaque sous matrice de taille $(p+n)$.

b) Le cas général se traite facilement par récurrence sur le nombre de blocs.

$\square$

Propriété 2.6. *Soit $Q \in \mathbb{R}^{n,n}$. On a les 3 équivalences suivantes :*
i) *Q est orthogonale si et seulement si*
ii) $\forall x \in \mathbb{R}^n$, $\|Q(x)\|_2 = \|x\|_2$ *si et seulement si*
iii) *les vecteurs colonnes de Q forment une base orthonormée de $\mathbb{R}^n$.*

Preuve.
Montrons : (i) $\Rightarrow$ (ii)$\Rightarrow$ (iii) $\Rightarrow$ (i).
(i) $\Rightarrow$ $(Q^tQx, x) = (x, x)\ \forall x \in \mathbb{R}^n \Rightarrow (Qx, Qx) = (x, x) \Rightarrow$ (ii).
Montrons que (ii) $\Rightarrow$ (iii) : Pour cela, remarquons que :

$$\forall x, y \in \mathbb{R}^n\ (x, y) = \frac{1}{2}((x+y, x+y) - (x, x) - (y, y)).$$

Remarquons aussi que les vecteurs colonnes de Q sont $Q(e_i)$ où e_i désigne le i ième vecteur de la base canonique de $\mathbb{R}^n$.

Donc $\forall i, j \in \{1, .., n\}\ (Qe_i, Qe_j) = \frac{1}{2}(\|Q(e_i + e_j)\|_2^2 - \|Q(e_i)\|_2^2 - \|Q(e_j)\|_2^2) =$
$\frac{1}{2}(\|e_i + e_j\|_2^2 - \|e_i\|_2^2 - \|e_j\|_2^2) = 0$ si $i \neq j$ et $= 1$ si $i = j$.

(iii) $\Rightarrow$ $(Q(e_i), Q(e_j)) = \delta_{ij}$, $1 \leq i, j \leq n$. Or $(Q(e_i), Q(e_j)) = (Q^tQ(e_i), e_j) =$ la j ième composante du vecteur $Q^tQ(e_i)$, c'est-à-dire $(Q^tQ)_{ji} = \delta_{ij} = \delta_{ji}$ ce qui signifie que $Q^tQ = I_n$. On en déduit donc (i). $\square$

2.2 Valeurs et Vecteurs Propres d'Endomorphismes ou de Matrices Carrées :

Dans la suite du cours, on suppose que E est un espace vectoriel sur $\mathbb{K}$ de **dimension finie** et on note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E.

Définition 2.5.
• Soit f un endomorphisme de E. L'élément $\lambda \in \mathbb{K}$ est appelé **valeur propre** de f si il existe un vecteur $x \in E$ non nul tel que $f(x) = \lambda x$.
• Soit $A \in \mathbb{K}^{n,n}$. $\lambda \in \mathbb{K}$ est appelé **valeur propre** de A si il existe un vecteur $x \in \mathbb{K}^n$ non nul tel que $Ax = \lambda x$.
x est alors appelé **vecteur propre** associé à la valeur propre λ.

Propriété 2.7. *λ est valeur propre de A $\Leftrightarrow$ $A-\lambda I_n$ n'est pas inversible $\Leftrightarrow$ $\det(A-\lambda I_n) = 0 \iff Ker(A - \lambda I_n) \neq \{0\}$.*

Définition 2.6. • On appelle polynôme caractéristique de A le polynôme noté P_A tel que

$$P_A(X) \;=\; \det(A - XI_n).$$

• Si λ est une valeur propre de A, on note m_λ la **multiplicité algébrique** de la racine λ de P_A dans $\mathbb{K}$:

$$P_A(X) = (X - \lambda)^{m_\lambda} Q(X) \quad \text{avec} \;\; Q(\lambda) \neq 0.$$

• $E_\lambda \;=\; \{x \in \mathbb{K}^n \;\; \text{tq.} \;\; Ax = \lambda x\}$ est appelé **sous-espace propre** associé à la valeur propre λ. La dimension de E_λ est appelé **multiplicité géométrique** de λ.
• L'ensemble des valeurs propres de A, noté $\mathrm{Sp}_{\mathbb{K}}(A)$, est appelé **spectre de** A.

Propriété 2.8.
Si $\mathbb{K} = \mathbb{R}$, alors $\displaystyle\sum_{\lambda \in Sp_{\mathbb{R}}(A)} m_\lambda \;\leq\; n.$
Si $\mathbb{K} = \mathbb{C}$, alors $\displaystyle\sum_{\lambda \in Sp_{\mathbb{C}}(A)} m_\lambda \;=\; n.$

Preuve. Dans $\mathbb{K}$, le polynôme caractéristique de $A \in \mathbb{K}^{n,n}$ est de degré n donc admet au plus n zéros, d'où le premier résultat. Dans $\mathbb{C}$, tout polynôme non-constant se décompose en produits de facteurs de degré 1 (On dit que tout polynôme est scindé), d'où le second résultat. □

Propriété 2.9. *admise*

$$\forall A \in \mathbb{K}^{n,n} \; \forall \lambda \in Sp_{\mathbb{K}}(A), \quad 1 \;\leq\; \dim(E_\lambda) \;\leq\; m_\lambda.$$

Remarque 2.3. Si λ est valeur propre simple alors $1 = \dim(E_\lambda) = m_\lambda$

Définition 2.7. ($\star$) Deux matrices carrées A et $B \in \mathbb{R}^{n,n}$ (resp. $\mathbb{C}^{n,n}$) sont dites **semblables** si il existe une matrice carrée $S \in \mathbb{R}^{n,n}$ (resp. $\mathbb{C}^{n,n}$) inversible telle que

$$A = S^{-1}BS.$$

(Deux matrices semblables représentent le même endomorphisme dans deux bases différentes)

Propriété 2.10. ($\star$)
Deux matrices semblables ont les mêmes valeurs propres mais la réciproque est fausse.

Preuve. $P_A(\lambda) = \det(A - \lambda I_n) = \det(S^{-1}BS - \lambda S^{-1}S) = \det(S^{-1}(B - \lambda I_n)S) = \det(S^{-1}\det(B - \lambda I_n)\det(S) = P_B(\lambda)$. A et B ont les mêmes polynômes caractéristiques donc ont les mêmes valeurs propres.
Les matrices I_2 et $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ont les mêmes valeurs propres. Si elles étaient semblables, il existerait S inversible telle que $A = S^{-1}I_2S$ d'où $A = I_2$ ce qui est faux. Donc elles ne sont pas semblables. $\square$

Propriété 2.11. *Soit $A \in \mathbb{K}^{n,n}$, soit $Sp(A) = \{\lambda_1, ..\lambda_p\}$ alors $\sum_{i=1}^{p} E_{\lambda_i}$ est directe.*

Preuve. • Cas $p = 2$ Soient $x_1 \in E_{\lambda_1} \cap E_{\lambda_2}$ alors $Ax_1 = \lambda_1 x_1 = \lambda_2 x_1$ avec $\lambda_1 \neq \lambda_2$ alors $x_1 = 0$. Donc $E_{\lambda_1} + E_{\lambda_2}$ est directe.
• Cas général par récurrence. Supposons qu'il existe k tel que $2 \leq k \leq p$ tel que la somme $\sum_{i=1}^{k-1} E_{\lambda_i}$ soit directe.

Soit $(x_1, .., x_k) \in E_{\lambda_1} \times .. \times E_{\lambda_k}$ tq. $\sum_{i=1}^{k} x_i = 0_E$. Alors $\sum_{i=1}^{k} Ax_i = A0_E = 0_E \Rightarrow \sum_{i=1}^{k} \lambda_i x_i = 0_E \Rightarrow \sum_{i=1}^{k-1} \lambda_i x_i + \lambda_k(-x_1 - \cdots - x_{k-1}) = 0_E$. D'où $\sum_{i=1}^{k-1}(\lambda_i - \lambda_k)x_i = 0_E$. La somme $\sum_{i=1}^{k-1} E_{\lambda_i}$ étant directe, pour tout $i \in \{1, \cdots, k-1\}(\lambda_i - \lambda_k)x_i = 0$, d'où $x_i = 0$ pour $i \leq k-1$ car les valeurs propres sont distinctes. $x_k = -x_1 - \cdots - x_{k-1} = 0$.

Donc la somme $\sum_{i=1}^{k} E_{\lambda_i}$ est directe. $\square$

2.2.1 Polynôme annulateur - Polynôme minimal

Définition 2.8. Si $P = \sum_{i=0}^{p} a_i X^i$ est un polynôme d'une variable à coefficients dans $\mathbb{K}$, la matrice carré d'ordre n, $P(A)$, est la matrice

$$P(A) = \sum_{i=0}^{p} a_i A^i, \quad \text{où} \quad A^0 = I_n.$$

De même, si $f \in \mathcal{L}(E)$, alors $P(f)$ est l'endomorphisme de E tel que

$$P(f) = \sum_{i=0}^{n} a_i f^i, \quad \text{où} \quad f^i = \underbrace{f \circ f .. \circ f}_{i \text{ fois}} \quad \text{et} \quad f^0 = Id.$$

Propriété 2.12. *Soient f un endomorphisme de E, P, Q deux polynômes à coefficients dans $\mathbb{K}$. Alors* **(1)** $P(f) + Q(f) = (P+Q)(f)$.
(2) $P(f) \circ Q(f) = (PQ)(f) = Q(f) \circ P(f)$.
(3) *$Ker(P(f))$ et $Im(P(f)$ sont des sous-espaces de E stables par f : $f(Ker(P(f))) \subset Ker(P(f))$ et $f(Im(P(f))) \subset Im(P(f))$.*

Lemme 2.1. des noyaux

Soient $f \in \mathcal{L}(E)$ et $P_1, \cdots, P_k$ des polynômes premiers entre eux. On note $P = \prod_{i=1}^{k} P_i$, alors

$$Ker(P(f) \quad = \quad Ker(P_1(f)) \oplus \cdots \oplus Ker(P_k(f)).$$

A tout $f \in \mathcal{L}(E)$, on associe l'application linéaire : $ev_f : \mathbb{K}[X] \to \mathcal{L}(E)$ telle que $ev_f(P) = P(f)$. Comme $\mathbb{K}[X]$ est de dimension infinie alors que $\mathcal{L}(E)$ est de dimension n^2, ev_f n'est pas injective, donc $\mathrm{Ker}(ev_f) \neq 0$ donc il existe au moins un polynôme P non nul tel que $P(f) = 0$.

Définition 2.9. Soit $f \in \mathcal{L}(E)$.
• On appelle **polynôme annulateur** de f tout polynôme P tel que $P(f) = 0$.
• On appelle **polynôme minimal** de f (ou A), que l'on note μ_f (ou μ_A) le polynôme annulateur de f (ou de A) unitaire et de plus bas degré.

Exemple 2.2. Un endomorphisme distinct de l'application nulle et de l'identité est un projecteur si et seulement si son polynôme minimal est $x^2 - X$.
Un endomorphisme nilpotent est un endomorphisme dont le polynôme minimal est de la forme X^d où d est appelé indice de nilpotence.

On admet le théorème fondamental suivant :

Théorème 2.2. Cayley-Hamilton :

$$P_A(A) \;=\; 0, \; \textit{où } P_A \textit{ est le polynôme caractéristique de } A.$$

Corollaire 2.1. $\deg(\mu_A) \;\leq\; n.$

Preuve. Comme P_A est un polynôme de degré n qui annule A, μ_A divise P_A donc le degré de μ_A est inférieur ou égal au degré de P_A. □

2.3 Réduction des Matrices Carrées :

L'idée est de trouver pour toute matrice $A \in \mathbb{K}^{n,n}$ une matrice B semblable à A, la plus simple possible. En général, on cherche B sous la forme diagonale ou triangulaire.

2.3.1 Diagonalisation de Matrices :

Définition 2.10. ($\star$)
Une matrice carrée $A \in \mathbb{C}^{n,n}$ est dite **diagonalisable** si elle est semblable à une matrice diagonale.

Propriété 2.13. ($\star$)
Si $A \in \mathbb{R}^{n,n}$ (resp. $\mathbb{C}^{n,n}$) est diagonalisable, il existe $S \in \mathbb{R}^{n,n}$ (resp. $\mathbb{C}^{n,n}$) matrice inversible et $\lambda_1, ..\lambda_n \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) tels que $S^{-1}AS \;=\; D := diag(\lambda_1, .., \lambda_n)$, où $diag(\lambda_1, .., \lambda_n) =$ $\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$. Alors les $(\lambda_i)_{1 \leq i \leq n}$ sont les valeurs propres de A et la $i^{ième}$ colonne de S est un vecteur propre de A associé à λ_i.

Preuve.
$S^{-1}AS(e_i) = \lambda_i e_i \Rightarrow AS(e_i) = \lambda_i S(e_i)$. $S(e_i)$, qui est la i ième colonne de S est non-nul car S est inversible et est donc un vecteur propre de A associé à λ_i.
D'autre part, $\lambda \in Sp(A) \Leftrightarrow \exists x \neq 0$ tel que $Ax = \lambda x \Leftrightarrow \exists x \neq 0,\ y = S^{-1}x\ ASy = \lambda Sy \Leftrightarrow \exists y \neq 0,\ S^{-1}ASy = \lambda y \Leftrightarrow \exists y \neq 0,\ (D - \lambda I)y = 0 \Leftrightarrow \lambda$ est l'un des λ_i. □

Corollaire 2.2. *A est diagonalisable si et seulement si $\mathbb{K}^n$ est la somme directe de ses sous-espaces propres.*

Preuve. Si A est diagonalisable alors d'après la propriété précédente, S étant inversible, $(S(e_1), ..S(e_n))$ est une base de $\mathbb{K}^n$ et donc $\mathbb{K}^n = \sum_{i=1}^{n} \mathrm{Vect}\{S(e_i)\}$, qui est une somme directe de sous-espaces propres de A. Donc $\mathbb{K}^n$ est égal à la somme de ses sous-espaces propres.
Réciproquement, si $\mathbb{K}^n$ est la somme de ses sous-espaces propres alors il existe $p \geq 1$ valeurs propres distinctes $\{\lambda_1, .., \lambda_p\}$ tels que $\mathbb{K}^n = \sum_{i=1}^{p} E_{\lambda_i}$ (somme directe). Si $(e_1^i, ..e_{n_i}^i)$ est une base de E_{λ_i} alors $S = (e_1^1, .., e_{n_1}^1, .., .., e_1^p, .., e_{n_p}^p)$ est une matrice carrée inversible (puisque la somme est directe et égale à $\mathbb{K}^n$). De plus $A(e_j^i) = \lambda_i e_j^i$, ce qui se traduit par $AS = S\mathrm{diag}(\underbrace{\lambda_1, .., \lambda_1}_{n_1 \text{ fois}}, .., \underbrace{\lambda_p, .., \lambda_p}_{n_p \text{ fois}})$. Donc $S^{-1}AS$ est une matrice diagonale. □

Théorème 2.3. *admis* ($\star$)
La matrice $A \in \mathbb{K}^{n,n}$ est diagonalisable si et seulement si les deux conditions suivantes sont réalisées :

- **i)** *Le polynôme caractéristique de A, P_A, est scindé sur $\mathbb{K}$, c'est-à-dire que P_A est le produit de polynômes du premier degré.*
- **ii)** $\forall \lambda \in Sp(A) \quad \dim(E_\lambda) = m_\lambda$.

Corollaire 2.3. *Si $A \in \mathbb{K}^{n,n}$ a n valeurs propres distinctes, alors A est diagonalisable.*

Preuve. On sait que si A a n valeurs propres distinctes alors $P_A(X) = \prod_{i=1}^{n}(\lambda_i - X)$ donc P_A est scindé et $\forall i \in [|1:n|]$, $1 \leq \dim(E_{\lambda_i}) \leq 1$ donc $\forall \lambda \in Sp(A) \quad \dim(E_\lambda) = m_\lambda = 1$. □

Remarque 2.4. Si $A \in \mathbb{C}^{n,n}$, alors la propriété **i)** du théorème 2.3 est toujours vérifiée car dans $\mathbb{C}$, tout polynôme non constant est scindé.

Remarque 2.5. La matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est diagonalisable dans $\mathbb{C}$ car elle admet i et $-i$ comme valeurs propres, mais n'est pas diagonalisable dans $\mathbb{R}$, car elle n'admet pas de valeurs propres réelles.

Corollaire 2.4. *La matrice $A \in \mathbb{R}^{n,n}$ est diagonalisable dans $\mathbb{R}$ si et seulement si les deux conditions suivantes sont réalisées :*

- **i)** *A est diagonalisable dans $\mathbb{C}$,*
- **ii)** *Toutes les valeurs propres de A sont réelles.*

Preuve. • Toute famille libre de $\mathbb{R}^n$ est une famille libre de $\mathbb{C}^n$. Donc toute base de $\mathbb{R}^n$ est une base de $\mathbb{C}^n$ pris comme $\mathbb{C}$-espace vectoriel (la réciproque est fausse).
Donc si A est diagonalisable dans $\mathbb{R}$ alors elle admet une base de vecteurs propres de $\mathbb{R}^n$ et donc aussi de $\mathbb{C}^n$. De plus elle admet n valeurs propres réelles (comptées avec leur multiplicité) donc elle est diagonalisable dans $\mathbb{C}$ et ses valeurs propres (dans $\mathbb{C}$) sont toutes réelles.
• Réciproquement, supposons que A soit diagonalisable dans $\mathbb{C}$ et que toutes ses valeurs propres soient réelles.
Alors d'une part le polynôme caractéristique P_A est scindé dans $\mathbb{C}$ et comme toutes les valeurs propres sont réelles, il est scindé dans $\mathbb{R}$.
D'autre part, notons pour $\lambda \in Sp_{\mathbb{C}}(A) \subset \mathbb{R}$:

$$E_\lambda \;=\; \{x \in \mathbb{C}^n,\ \text{tq. } \ Ax = \lambda x\} \quad \text{ et } \quad F_\lambda \;=\; \{x \in \mathbb{R}^n,\ \text{tq. } \ Ax = \lambda x\}.$$

Si on montre que $\dim_{\mathbb{C}}(E_\lambda) = \dim_{\mathbb{R}}(F_\lambda)$ alors on aura montré que A est diagonalisable dans $\mathbb{R}$ puisque l'on a déjà : $\dim_{\mathbb{C}}(E_\lambda) = m_\lambda$ (cf. théorème 2.3).
Soit $\lambda \in Sp_{\mathbb{C}}(A)$ et notons $(v_1, .., v_p)$ une base de F_λ. Montrons que E_λ est engendré par $(v_1, .., v_p)$:
Soit $x \in E_\lambda$ alors $A(\Re(x) + i\Im(x)) = \lambda(\Re(x) + i\Im(x))$ d'où A et λ étant réels, on a $A(\Re(x)) = \lambda\Re(x)$ et $A(\Im(x)) = \lambda\Im(x)$ donc $\Re(x) \in F_\lambda$ et $\Im(x) \in F_\lambda$ donc $\Re(x)$ et $\Im(x)$ sont combinaisons linéaires de $v_1, .., v_p$ donc x est aussi combinaison linéaire de $v_1, .., v_p$ (dans $\mathbb{C}$). De plus toute famille libre de $\mathbb{R}^n$ est une famille libre de $\mathbb{C}^n$, donc $(v_1, .., v_p)$ est une base de E_λ d'où $\dim_{\mathbb{R}} F_\lambda = \dim_{\mathbb{C}} E_\lambda$. □

2.3.2 Quelques compléments sur les polynômes annulateurs

Vous trouverez les démonstrations de ce complément dans l'article de Matthieu Romagny.

Théorème 2.4. *Soit A une matrice carrée d'ordre n. Les conditions suivantes sont équivalentes :*
(1) *A est diagonalisable*
(2) *Le polynôme minimal de A est scindé à racines simples.*
(3) *Il existe un polynôme scindé à racines simples qui annule A.*

Théorème 2.5. *Soit A une matrice carrée d'ordre n. Les conditions suivantes sont équivalentes :*
(1) *A est trigonalisable*
(2) *Le polynôme caractéristique de A est scindé.*

On rappelle qu'un polynôme non constant est **scindé** s'il est le produit de polynômes de degré 1.

Définition 2.11. Soit $\lambda \in \mathbb{K}$ une valeur propre de A et α la multiplicité de λ dans le polynôme caractéristique de A. On appelle **sous-espace caractéristique de A associé à** λ le sous-espace, noté $F_\lambda \;=\; \operatorname{Ker}((A - \lambda I_n)^\alpha)$.

Théorème 2.6. Sous-espaces caractéristiques de A

On suppose que A a un polynôme caractéristique scindé : $P_A(X) = \prod_{i=1}^{p} (X - \lambda_i)^{\alpha_i}$. On note

μ_A le polynôme minimal de A. Alors
(1) *$E = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_p}$.*
(2) *Pour tout i tel que $1 \leq i \leq p$, on pose $f_i : F_{\lambda_i} \to F_{\lambda_i}$ la restriction de f à F_{λ_i}. L'endomorphisme $f_i - \lambda_i Id$ est nilpotent et la seule valeur propre de f_i est λ_i. Soit μ_i le polynôme minimal de f_i, alors $\mu_i(X) = (X - \lambda_i)^{\beta_i}$ avec $\beta_i \leq \alpha_i$ est l'indice de nilpotence de $f_i - \lambda_i Id$.*
(3) *$\mu_A = \mu_1 \cdots \mu_p$.*
(4) *Pour tout i tel que $1 \leq i \leq p$, $\dim(F_{\lambda_i}) = \alpha_i$.*
(3) *Soit $\mathcal{B}_i$ une base qui trigonalise f_i (une telle base existe d'après le théorème 2.5) et $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_p$ (qui est une base de E). Alors la matrice M de f dans la base $\mathcal{B}$ est triangulaire, diagonale par blocs (triangulaires) de tailles α_i.*

De plus f est diagonalisable si et seulement si $\mu_A = (X - \lambda_1) \cdots (X - \lambda_p)$ (cad : $\beta_i = 1$) si et seulement si M est diagonale.

2.3.3 Trigonalisation de Matrices Carrées :

Lemme 2.2. Matrice élémentaire dite de Householder
Toute matrice $H \in \mathbb{C}^{n,n}$ de la forme $H = I_n - 2uu^*$, où u est un vecteur unitaire ($\|u\|_2 = \sqrt{u^*u} = 1$ norme issue du produit scalaire usuel $(u, v) = v^*u$) est hermitienne et unitaire et est appelée **matrice unitaire élémentaire***.*
De plus, H est la symétrie orthogonale par rapport à l'hyperplan orthogonal à u.

Preuve. • Montrons que H est hermitienne :
$H^\star = (I_n^\star - 2uu^\star)^\star = I_n^\star - 2(uu^\star)^\star = I_n - 2((u^\star)^\star u^\star) = H$ donc H est hermitienne.
• $HH = (I_n - 2uu^\star)(I_n - 2uu^\star) = I_n - 4uu^\star + 4(uu^\star)(uu^\star)$. Or $(uu^\star)(uu^\star) = u(u^\star u)u^\star = (\|u\|^2)uu^\star = uu^\star$. Donc $HH = I_n$. H est donc orthogonale.
• Pour montrer que H est la symétrie orthogonale par rapport à $\{u\}^\perp$, il suffit de montrer que $Hu = -u$ et $Hv = v \quad \forall v \in \{v \text{ tq. } (u, v) = 0\} = \{u\}\perp$.

- $H(u) = u - 2uu^\star u = u - 2u\|u\|^2 = -u$.
- $H(v) = v - 2uu^\star v = v - 2u(v, u) = v$ pour tout $v \in \{u\}^\perp$.

□

Lemme 2.3. *On note $(e_1, .., e_n)$ la base canonique de $\mathbb{C}^n$. Soit $a \in \mathbb{C}^n$, tel que $a \neq 0$.*

$$\exists H \text{ une matrice élémentaire unitaire et } \alpha \in \mathbb{C}^\star \text{ tels que } H(a) = \alpha e_1. \tag{4}$$

(Ce résultat permet ensuite de démontrer le théorème de Schur, vu en cours.)

Preuve. Soit $a \in \mathbb{C}^n$, tel que $a \neq 0$.

i) Analyse : On montre que si $H = I - 2uu^\star$ et α vérifient (4), alors

$$\begin{cases} |\alpha| = \|a\|_2 \\ \text{si on pose } \mu = 2u^\star a \text{ alors } \mu u = a - \alpha e_1 \\ |\mu|^2 = 2a^\star \mu u = 2\alpha(\overline{\alpha} - \overline{a_1}) = 2\overline{\alpha}(\alpha - a_1) \end{cases}$$

ii) Synthèse : Si on pose $\alpha = \begin{cases} \dfrac{-a_1}{|a_1|}\|a\|_2 & \text{si } a_1 \neq 0 \\ \|a\|_2 & \text{si } a_1 = 0 \end{cases}$, $\mu = \sqrt{2\overline{\alpha}(\alpha - a_1)}$ et $u = \dfrac{1}{\mu}(a - \alpha e_1)$, on montre que

$$\mu = \sqrt{2\|a\|_2(\|a\|_2 + |a_1|)}, \quad 2u^\star a = \mu, \quad \|u\|_2 = 1, \quad H(a) = \alpha e_1.$$

□

Remarque 2.6. Dans le cas réel, si $u \in \mathbb{R}^n$ est un vecteur unitaire, la matrice $H = I_n - 2uu^t$ est appelée matrice orthogonale élémentaire.
De plus, si $a \in \mathbb{R}^n$ est un vecteur non-nul, alors il existe une matrice élémentaire orthogonale H et un réel α non-nul tels que $H(a) = \alpha e_1$, e_1 étant le premier vecteur de la base canonique de $\mathbb{R}^n$ (considéré comme un $\mathbb{R}$ espace vectoriel). La preuve est analogue à celle du lemme précédent.

Théorème 2.7. de Schur($\star\star$) *:*
Soit $A \in \mathbb{C}^{n,n}$, il existe une matrice $U \in \mathbb{C}^{n,n}$ unitaire et une matrice T triangulaire supérieure telle que

$$T = U^\star A\, U, \tag{5}$$

les éléments diagonaux de T étant les valeurs propres de A.

Preuve. On montre le résultat du théorème par récurrence sur n, c'est-à-dire :
i) On vérifie que (5) est vraie pour $n = 1$ ($u_{11} = 1$ et $T_{11} = A_{11}$).
ii) Soit $n \geq 2$ tel que (5) est vraie à l'ordre $n-1$. On désire alors montrer (5) à l'ordre n. Soit $A \in \mathbb{C}^{n,n}$. Dans $\mathbb{C}$, A admet nécessairement n valeurs propres (comptées avec leur ordre de multiplicité).
Soient $\lambda \in Sp(A)$ et x un vecteur propre associé à λ. Alors il existe H une matrice élémentaire unitaire et $\alpha \in \mathbb{C}^\star$ tels que $Hx = \alpha e_1$.
$HAH(e_1) = \dfrac{1}{\alpha}HAHH(x) = \dfrac{1}{\alpha}HA(x) = \dfrac{\lambda}{\alpha}H(x) = \lambda e_1$. Donc HAH est de la forme suivante :

$$HAH = \left(\begin{array}{c|c} \lambda & b^t \\ \hline 0 & \\ \vdots & A_{n-1} \\ 0 & \end{array}\right) \text{ où } b \in \mathbb{C}^{n-1},\ A_{n-1} \in \mathbb{C}^{n-1,n-1}.$$

Or $A_{n-1} \in \mathbb{C}^{n-1,n-1}$ donc d'après l'hypothèse de récurrence, il existe U_{n-1} unitaire et T_{n-1} triangulaire supérieure telles que $T_{n-1} = U_{n-1}^\star A_{n-1} U_{n-1}$. Posons $B_n = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_{n-1} \\ 0 & \end{array}\right)$ alors

$$B_n^\star = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_{n-1}^\star \\ 0 & \end{array}\right) \text{ et } B_n B_n^\star = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_n U_{n-1}^\star \\ 0 & \end{array}\right) = I_n.$$

Calculons $B_n^\star HAHB_n$:

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_{n-1}^\star \\ 0 & \end{array}\right) \left(\begin{array}{c|c} \lambda & b^T \\ \hline 0 & \\ \vdots & A_{n-1} \\ 0 & \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_{n-1} \\ 0 & \end{array}\right) =$$

$$\left(\begin{array}{c|c} \lambda & b^T \\ \hline 0 & \\ \vdots & U_{n-1}^\star A_{n-1} \\ 0 & \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & U_{n-1} \\ 0 & \end{array}\right) = \left(\begin{array}{c|c} \lambda & b^T U_{n-1} \\ \hline 0 & \\ \vdots & U_{n-1}^\star A_{n-1} U_{n-1} \\ 0 & \end{array}\right) =$$

$$\left(\begin{array}{c|c} \lambda & b^T U_{n-1} \\ \hline 0 & \\ \vdots & T_{n-1} \\ 0 & \end{array}\right).$$

La matrice $T_n = \left(\begin{array}{c|c} \lambda & b^T U_{n-1} \\ \hline 0 & \\ \vdots & T_{n-1} \\ 0 & \end{array}\right)$ est triangulaire supérieure et

$$T_n = B_n^\star HAHB_n = (HB_n)^\star A(HB_n).$$

Posons $U_n = HB_n$ alors U_n est unitaire en tant que produit de deux matrices unitaires. D'où $T_n = U_n^\star AU_n$. □

Corollaire 2.5. $(\star)$
Une condition nécessaire et suffisante pour qu'une matrice $A \in \mathbb{C}^{n,n}$ soit **hermitienne** *est qu'elle soit* **unitairement semblable à une matrice diagonale réelle** *: $\Lambda = U^\star AU$ (Λ : matrice diagonale, U matrice unitaire).*
Les colonnes de U sont des vecteurs propres de A. Ils forment une base orthonormée de $\mathbb{C}^n$. En particulier, toute matrice $A \in \mathbb{C}^{n,n}$ hermitienne admet n valeurs propres réelles (comptées avec leur ordre de multiplicité) et une base orthonormée de vecteurs propres.

Preuve.
• $A \in \mathbb{C}^{n,n}$ est hermitienne. Il existe U unitaire et T triangulaire supérieure telles que $T = U^\star AU$. Or $U^\star AU$ est hermitienne, donc T aussi. Une matrice triangulaire hermitienne est diagonale à diagonale réelle.
• Réciproque : Si $D = U^\star AU$ où D est diagonale réelle alors $A = UDU^\star$ est hermitienne.
□

Corollaire 2.6. *Soit $A \in \mathbb{C}^{n,n}$. Une condition nécessaire et suffisante pour que A soit* **normale** *est qu'elle soit unitairement semblable à une matrice diagonale* **(la matrice diagonale n'est pas en général réelle)**.

Preuve. • Montrons qu'une matrice T triangulaire supérieure normale est diagonale :
$(TT^\star)_{ii} = \sum_{j=1}^{n} T_{ij}\overline{T_{ji}^T} = \sum_{j=1}^{n} T_{ij}\overline{T_{ij}} = \sum_{j=1}^{n} |T_{ij}|^2 = \sum_{j=i}^{n} |T_{ij}|^2$

$((T^\star)T)_{ii} = \sum_{j=1}^{n} \overline{T^T_{ij}} T_{ji} = \sum_{j=1}^{n} \overline{T_{ji}} T_{ji} = \sum_{j=1}^{n} |T_{ji}|^2 = \sum_{j=1}^{i} |T_{ji}|^2$. Donc

$$\text{pour } 1 \leq i \leq n \quad \sum_{j=i}^{n} |T_{ij}|^2 \;=\; \sum_{j=1}^{i} |T_{ji}|^2.$$

- Pour $i = 1$: $\sum_{j=1}^{n} |T_{1j}|^2 = |T_{11}|^2$, d'où $T_{1j} = 0$ pour $2 \leq j \leq n$, c'est-à-dire que la 1ière ligne de T est nulle excepté T_{11}.
- En procédant par récurrence, on suppose que les k premières lignes de T sont nulles excepté les éléments diagonaux de T. Alors on a $\sum_{j=k+1}^{n} |T_{k+1,j}|^2 \;=\; \sum_{j=1}^{k} \underbrace{|T_{j,k+1}|^2}_{=0} + |T_{k+1,k+1}|^2$ (les $T_{j,k+1}$ appartiennent aux k premières lignes) donc $\sum_{j=k+2}^{n} |T_{k+1,j}|^2 = 0$, c'est-à-dire $T_{k+1,j} = 0$ pour $j > k+1$. Donc la $k+1$ ième ligne de T est nulle, excepté son coefficient diagonal.

• On en déduit alors le résultat du théorème, grâce au théorème de Schur :
$T^\star T = (U^\star AU)^\star (U^\star AU) = U^\star A^\star U U^\star AU = U^\star A^\star I_n AU$ et
$TT^\star = (U^\star AU)(U^\star AU)^\star = U^\star AU(U^\star A^\star U) = U^\star A I_n A^\star U = T^\star T$ car $A^\star A = AA^\star$. Donc T est normale. Etant triangulaire, elle est diagonale. □

Théorème 2.8. ($\star$)
Soit $A \in \mathbb{R}^{n,n}$ une matrice carrée ayant toutes ses valeurs propres réelles, alors il existe une matrice $Q \in \mathbb{R}^{n,n}$ orthogonale telle que la matrice $T = Q^t AQ$ soit triangulaire supérieure réelle, les éléments diagonaux de T étant les valeurs propres de A, T étant semblable à A.

Preuve. La démonstration est l'analogue de celle du théorème de Schur. Utiliser la remarque (2.6). □

Corollaire 2.7. ($\star$)
Une condition nécessaire et suffisante pour que $A \in \mathbb{R}^{n,n}$ soit symétrique est qu'elle soit semblable à une matrice diagonale réelle par une transformation orthogonale :

$$Q^t AQ = \mathit{diag}(\lambda_i), \quad \text{avec } Q \text{ orthogonale et } \lambda_i \text{ réels.}$$

Les colonnes de Q sont des vecteurs propres de A. Ils forment une base orthonormée de $\mathbb{R}^n$.

Preuve. Il suffit d'utiliser le théorème 2.8 et le fait qu'une matrice triangulaire supérieure et symétrique est diagonale. □

Corollaire 2.8. ($\star$)
Soit $A \in \mathbb{C}^{n,n}$. La trace de A ($= \sum_{i=1}^{n} A_{ii}$) est la somme des valeurs propres de A dans $\mathbb{C}$, même si A est à coefficients réels.
Le déterminant de A est égale au produit de ses valeurs propres dans $\mathbb{C}$, même si A est une matrice carrée réelle.

Propriété 2.14. *Quelques Rappels sur la trace : Pour toute matrice $A, B, P \in \mathbb{K}^{n,n}$, telles que P est inversible, on a*

$$tr(AB) = tr(BA) \qquad tr(A) = tr(P^{-1}AP).$$

Preuve. La deuxième égalité est une conséquence de la première : en effet $\mathrm{tr}((P^{-1}A)P) = \mathrm{tr}(P \cdot (P^{-1}A)) = \mathrm{tr}(A)$.
$\mathrm{tr}(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}(\sum_{k=1}^{n} A_{ik}B_{ki})$ (en intervertissant les deux sommes) $= \sum_{k=1}^{n}(\sum_{i=1}^{n} B_{ki}A_{ik}) = \sum_{k=1}^{n}(BA)_{kk} \ = \ \mathrm{tr}(BA)$. □

Preuve. **du corollaire** : D'après le théorème de Schur (2.7), il existe T une matrice triangulaire et U une matrice unitaire $U^\star$ telle que $T = U^\star AU$. D'où $\det(T) = \det(U\star)\det(A)\det(U)$. Or $U^\star = U^{-1}$ et $\det(U^{-1}) = 1/\det(U)$. Donc $\det(T) = \det(A)$. Or $\det(T)$ est le produit de ses coefficients diagonaux, lesquels sont les valeurs propres de A. Donc $\det(A)$ est égal au produit de ses valeurs propres comptées avec leur multiplicité.

D'autre part, $\mathrm{tr}(T) = \mathrm{tr}((U^\star A)U) = \mathrm{tr}(U(U^\star)A)) = \mathrm{tr}(A)$ car $UU^\star = I_n$. Or la trace de T est la somme de ses coefficients diagonaux donc aussi la somme des valeurs propres de A comptées avec leur multiplicité. □

2.4 Matrices définies positives :

Définition 2.12. $(\star)$
On considère une matrice A hermitienne si $A \in \mathbb{C}^{n,n}$ ou symétrique si $A \in \mathbb{R}^{n,n}$. On dit que A est **définie positive** (resp. **semi-définie positive**) si

$$\forall x \in \mathbb{K}^n \backslash \{0\} \quad (Ax, x) > 0 \quad (\text{resp. } (Ax, x) \geq 0),$$

où $(\cdot, \cdot)$ est le produit scalaire usuel sur $\mathbb{K}^n$.

Remarque 2.7. Soit $A \in \mathbb{C}^{n,n}$ alors $\forall x \in \mathbb{C}$ $(Ax, x) \in \mathbb{C}$ donc n'est pas en général un réel. Toutefois, si A est hermitienne alors $\forall x \in \mathbb{C}$ $\overline{(Ax, x)} = (x, Ax) = (A^\star x, x) = (Ax, x)$. Donc (Ax, x) est un réel.

Théorème 2.9. $(\star)$
Une condition nécessaire et suffisante pour qu'une matrice $A \in \mathbb{C}^{n,n}$ hermitienne ou $A \in \mathbb{R}^{n,n}$ symétrique soit définie positive (resp. semi-définie positive) est que toutes ses valeurs propres $\lambda_1, .., \lambda_n$ soient > 0(resp. ≥ 0). On a alors :

$$\forall x \in \mathbb{C}^n, \quad (\min_{1 \leq i \leq n} \lambda_i) \; \|x\|_2^2 \ \leq \ (Ax, x) \ \leq \ (\max_{1 \leq i \leq n} \lambda_i) \; \|x\|_2^2. \tag{6}$$

Preuve.
Si A est hermitienne, ses valeurs propres sont toutes réelles : Soient $\lambda_1 \leq .. \leq \lambda_n$ ses valeurs propres et $(u_1.., u_n)$ une base orthonormée de $\mathbb{C}^n$, formée de vecteurs propres associés aux valeurs propres $\lambda_1, .., \lambda_n$.
Condition nécessaire : Si A est définie positive alors $0 < (Au_i, u_i) = \lambda_i\|u_i\|_2^2$. Donc $\lambda_i > 0$ (et $\lambda_i \geq 0$ si A est semi-définie positive).

Condition suffisante : Soit A hermitienne et $\lambda_1 > 0$. Soit $x \in \mathbb{C}^n$, alors $x = \sum_{i=1}^{n} \alpha_i u_i$ et

$\|x\|_2^2 = (\sum_{i=1}^{n} \alpha_i u_i, \sum_{k=1}^{n} \alpha_k u_k) = \sum_{i=1}^{n}\sum_{k=1}^{n} \alpha_i \overline{\alpha_k} \underbrace{(u_i, u_k)}_{=\delta_{ik}} = \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} = \sum_{i=1}^{n} |\alpha_i|^2$ (car la base des vecteurs propres $(u_1, .., u_n)$ est orthonormée).
$(Ax, x) = (\sum_{i=1}^{n} \alpha_i A u_i, \sum_{j=1}^{n} \alpha_j u_j) = (\sum_{i=1}^{n} \alpha_i \lambda_i u_i, \sum_{j=1}^{n} \alpha_j u_j) = \sum_{i=1}^{n} |\alpha_i|^2 \lambda_i \geq (\min_{1 \leq i \leq n} \lambda_i) \|x\|_2^2 = \lambda_1 \|x\|^2 > 0,\ \forall x \in \mathbb{C}^n \backslash \{0\}$.
On en déduit que A est définie positive.
De même, on a aussi : $(Ax, x) = \sum_{i=1}^{n} |\alpha_i|^2 \lambda_i \leq (\max_{1 \leq i \leq n} \lambda_i) \|x\|_2^2$. □

Propriété 2.15. et Définition *Soit $A \in \mathbb{C}^{n,p}$. Alors $(A^\star)A$ est hermitienne et semi-définie positive. De plus, $A^\star A$ est définie positive si et seulement si A est injective.*
On appelle **valeurs singulières de** *A les racines carrées des valeurs propres de $A^\star A$.*

Preuve. vu en td.
$A^\star A \in \mathbb{C}^{p,p}$. De plus $\forall x \in \mathbb{C}^p ((A^\star)Ax, x) = (Ax, Ax) \geq 0$. Donc $(A^\star)A$ est semi-définie positive. De plus $((A^\star)Ax, x) = 0 \iff Ax = 0$. Donc $(A^\star)A$ est définie positive si et seulement si $Ax = 0 \Rightarrow x = 0$, ce qui n'est rien d'autre que le fait que A est injective. □

Définition 2.13. $(\star\star)$
Soit $A \in \mathbb{C}^{n,n}$, si on note $\lambda_1, ..\lambda_n$ les valeurs propres de A, on appelle **rayon spectral** de A le réel noté $\rho(A)$ tel que

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \qquad (|\cdot| \text{ désigne le module}).$$