Université Paris-Saclay - L3 EU

Analyse matricielle et optimisation

Notes de cours - 2023/2024

Filipa Caetano (filipa.caetano@universite-paris-saclay.fr)

Ce cours est divisé en deux parties.

L'objectif de la première est de présenter les outils nécessaires pour étudier les principaux problèmes de l'analyse numérique matricielle :

- la résolution de systèmes de n équations linéaires à n inconnues : étant donnés $A \in \mathcal{M}_n(\mathbb{K})$ une matrice inversible et $b \in \mathbb{K}^n$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, trouver $x \in \mathbb{K}^n$ tel que $Ax = b$;
- le calcul de valeurs et vecteurs propres d'une matrice : étant donnée $A \in \mathcal{M}_n(\mathbb{K})$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, trouver $\lambda \in \mathbb{C}$ et $u \in \mathbb{C}^n \backslash \{0\}$ tel que $Au = \lambda u$.

La deuxième partie du cours consiste en une introduction à l'optimisation numérique. L'objectif de cette deuxième partie est d'étudier des problèmes d'optimisation sur $\mathbb{R}^n$: étant donnée une fonction $f : \Omega \longrightarrow \mathbb{R}$ définie dans un ouvert Ω de $\mathbb{R}^n$,

- trouver $x_m \in \Omega$ tel que $f(x_m) = \min\limits_{x \in \Omega} f(x)$, si un tel x_m existe (si f est bornée inférieurement dans Ω et atteint sa borne inférieure sur Ω),

ou

- trouver $x_m^C \in C$ tel que $f(x_m^C) = \min\limits_{x \in C} f(x)$, où $C \subset \Omega$ est un *ensemble de contrainte* donné, si un tel x_m^C existe.

Nous allons étudier des algorithmes permettant le calcul effectif des solutions de ces deux problèmes à l'aide d'un ordinateur. Nous cherchons des algorithmes efficaces, rapides et peu coûteux en nombre d'opérations, et stables, peu influencés par les erreurs sur les données (notamment car lorsque l'on cherche à les implémenter sur ordinateur, des erreurs d'arrondi sont commises).

On utilise le calcul numérique dans la simulation de modèles issus d'autres disciplines, telles que la physique, la biologie, la mécanique, l'économie,... La simulation numérique de ces modèles conduit souvent à des modèles discrets correspondant à des systèmes linéaires ou à des problèmes d'optimisation de grande taille. Il est alors important d'avoir des algorithmes robustes pour leur résolution, ce qui est un des objectifs de l'analyse numérique matricielle et de l'optimisation numérique.

Merci de signaler toute faute que vous trouvez dans ces notes.

CHAPITRE 1

Rappels de calcul matriciel

Contents

Ce chapitre est un rappel, ou un approfondissement, de certaines notions et de certains résultats d'analyse matricielle qu'il est indispensable de connaître.

Dans toute la suite $\mathbb{K}$ désigne le corps $\mathbb{R}$ des nombres réelles ou le corps $\mathbb{C}$ des nombres complexes.

Soient $n,\ m \in \mathbb{N}^*$. On note $\mathcal{M}_{n,m}(\mathbb{K})$, ou $\mathcal{M}_n(K)$ si $n = m$, l'ensemble des matrices à n lignes et m colonnes, à coefficients dans le corps K.

L'ensemble $\mathcal{M}_{n,m}(\mathbb{K})$ muni des opérations usuelles d'addition de matrice et de produit d'une matrice par un scalaire, $(\mathcal{M}_{n,m}(\mathbb{K}), +, \cdot)$, est un espace vectoriel sur $\mathbb{K}$, de dimension $n \times m$.

Si $A \in \mathcal{M}_{n,m}(\mathbb{K})$, on note $A = [A_{ij}]_{i\in\{1,\ldots,n\},\, j\in\{1,\ldots,n\}}$, A_{ij} étant le coefficient de la ligne i et de la colonne j de A. On notera parfois la matrice A en considérant la famille de vecteurs de $\mathbb{K}^n$ correspondant à ses lignes et à ses colonnes : si $l_1, \ldots, l_n \in \mathbb{K}^n$ sont les lignes de la matrice A et $c_1, \ldots, c_n \in \mathbb{K}^n$ sont les colonnes de la matrice A, on notera

$$A = \left[\begin{array}{c} l_1 \\ \hline \vdots \\ \hline l_n \end{array}\right] = \left[\begin{array}{c|c|c} c_1 & \cdots & c_n \end{array}\right].$$

Produit matriciel.
Soient $n,\ m,\ p \in \mathbb{N}^*$, $A \in \mathcal{M}_{n,m}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$. Le produit de A par B est la matrice $AB \in \mathcal{M}_{n,p}(\mathbb{K})$ définie par

$$(AB)_{ij} = \sum_{k=1}^{m} A_{ik}B_{kj},\ \forall\, i \in \{1,\ldots,n\},\ \forall\, j \in \{1,\ldots,p\}.$$

Quelques rappels de base sur les nombres complexes.
Soit $z \in \mathbb{C}$, $z = a + bi$, avec $a,\ b \in \mathbb{R}$ et $i^2 = -1$. On appelle a la partie réelle de z ($a = \mathrm{Re}(z)$) et b la partie imaginaire de z ($b = \mathrm{Im}(z)$).

Le nombre complexe $\bar{z} = a - bi$ est le conjugué de z.

Le module de z est le nombre réel positif $|z| = \sqrt{a^2 + b^2}$.

On a :

$|\bar{z}| = |z|$;

$z\bar{z} = |z|^2$;

$\mathrm{Re}(z) \leq |\mathrm{Re}(z)| \leq |z|$ (car $a^2 \leq a^2 + b^2$ donc $|a| \leq \sqrt{a^2 + b^2}$) ;

$\mathrm{Im}(z) \leq |\mathrm{Im}(z)| \leq |z|$;

$z + \bar{z} = 2\mathrm{Re}(z)$;

$z - \bar{z} = 2\mathrm{Im}(z)$.

1.1 Matrices inversibles

Soit $n \in \mathbb{N}^*$.

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est **inversible** s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ tel que $AB = BA = I_n$. Dans ce cas la matrice B s'appelle matrice inverse de A et est notée A^{-1}.

Pour $A \in \mathcal{M}_n(\mathbb{K})$, on note :

- $\mathrm{Ker}(A) = \{u \in \mathbb{K}^n \,|\, Au = 0\}$ le **noyau** de A ;
- $\mathrm{Im}(A) = \{Au \,|\, u \in \mathbb{K}^n\}$ l'**image** de A.

On a alors que A est inversible

ssi $\mathrm{Ker}(A) = \{0_{\mathbb{K}^n}\}$
ssi $\mathrm{Im}(A) = \mathbb{K}^n$
ssi $\det(A) \neq 0$
ssi la famille $(c_1, \dots, c_n)$ formée des colonnes de A est une famille libre de $\mathbb{K}^n$
ssi la famille $(l_1, \dots, l_n)$ formée des lignes de A est une famille libre de $\mathbb{K}^n$.

Il existe un autre critère utile pour savoir si une matrice est inversible.

Définition. *Matrice à diagonale strictement dominante.*
On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est à diagonale strictement dominante si pour tout $i \in \{1, \dots, n\}$,

$$|A_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |A_{ij}|.$$

Une matrice à diagonale strictement dominante est alors une matrice telle que sur chaque ligne, l'élément diagonal en valeur absolue est supérieur à la somme des valeurs absolues des éléments non diagonaux.

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice à diagonale strictement dominante. Alors A est inversible.

Démonstration. Montrons que $\ker(A) = \{0_{K^n}\}$.

Soit $x \in \mathbb{K}^n$ tel que $Ax = 0$. Soit $i_0 \in \{1, \dots, n\}$ tel que $x_{i_0} = \max_{i \in \{1, \dots, n\}} |x_i|$. Nous allons montrer que $|x_{i_0}| = 0$. On aura alors que $|x_j| = 0$, pour tout $j \in \{1, \dots, n\}$, et donc que $x = 0$.

On a $(Ax)_{i_0} = 0$, et donc $|(Ax)_{i_0}| = 0$, c'est-à-dire que

$$0 = \Big| \sum_{j=1}^{n} A_{i_0 j} x_j \Big| = \Big| A_{i_0 i_0} x_{i_0} + \sum_{\substack{j=1 \\ j \neq i_0}}^{n} A_{i_0 j} x_j \Big|.$$

En appliquant deux fois l'inégalité triangulaire, on obtient

$$0 \geq \Big| A_{i_0 i_0} x_{i_0} \Big| - \Big| \sum_{\substack{j=1 \\ j \neq i_0}}^{n} A_{i_0 j} x_j \Big| \geq \big| A_{i_0 i_0} \big| \big| x_{i_0} \big| - \sum_{\substack{j=1 \\ j \neq i_0}}^{n} \big| A_{i_0 j} \big| \big| x_j \big|.$$

Comme $|x_{i_0}| \geq |x_j|$, pour tout $j \in \{1, \ldots, n\}$, on obtient

$$0 \geq \left|A_{i_0 i_0}\right| \left|x_{i_0}\right| - \sum_{\substack{j=1 \\ j \neq i_0}}^{n} \left|A_{i_0 j}\right| \left|x_{i_0}\right| = \Big(\left|A_{i_0 i_0}\right| - \sum_{\substack{j=1 \\ j \neq i_0}}^{n} \left|A_{i_0 j}\right|\Big) \left|x_{i_0}\right|.$$

Comme A est à diagonale strictement dominante, on a

$$|A_{i_0 i_0}| > \sum_{\substack{j=1 \\ j \neq i_0}}^{n} |A_{i_0 j}|,$$

c'est-à-dire que

$$\Big(\left|A_{i_0 i_0}\right| - \sum_{\substack{j=1 \\ j \neq i_0}}^{n} \left|A_{i_0 j}\right|\Big) > 0.$$

De l'inégalité

$$\Big(\left|A_{i_0 i_0}\right| - \sum_{\substack{j=1 \\ j \neq i_0}}^{n} \left|A_{i_0 j}\right|\Big) \left|x_{i_0}\right| \leq 0$$

on déduit alors que $\left|x_{i_0}\right| = 0$. □

On finit cette section en rappelant la notion de matrices semblables.

Définition. *Matrices semblables.*
Deux matrices $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ sont dites **semblables** s'il existe une matrice $P \in \mathcal{M}_n(\mathbb{K})$ inversible tel que $A = PBP^{-1}$.

1.2 Matrices triangulaires et matrices diagonales

Soit $D \in \mathcal{M}_n(\mathbb{K})$. On dit que D est **diagonale** si pour tous $i, j \in \{1, \ldots, n\}$ tels que $i \neq j$, $D_{ij} = 0$.

Notation : Si $d_1, \ldots, d_n \in \mathbb{K}$, on note par $\mathrm{diag}(d_1, \ldots, d_n)$ la matrice diagonale $D \in \mathcal{M}_n(\mathbb{K})$ telle que $D_{ii} = d_i$, pour tout $i \in \{1, \ldots, n\}$, et $D_{ij} = 0$, pour tous $i, j \in \{1, \ldots, n\}$, $i \neq j$.

Soit $T \in \mathcal{M}_n(\mathbb{K})$. On dit que :

T est **triangulaire supérieure (stricte)** si $T_{ij} = 0$ pour tous $i, j \in \{1, \ldots, n\}$ tels que $i > j$ $(i \geq j)$;

T est **triangulaire inférieure (stricte)** si $T_{ij} = 0$ pour tous $i, j \in \{1, \ldots, n\}$ tels que $i < j$ $(i \leq j)$.

Si $M \in \mathcal{M}_n(\mathbb{K})$ est diagonale ou triangulaire, on a $\det(M) = \prod_{i=1}^{n} M_{ii} = M_{11} \times \cdots \times M_{nn}$.

On montrera comme exercice en Td que l'inverse d'une matrice triangulaire inférieure (supérieure) inversible est encore une matrice triangulaire inférieure (supérieure), et que le produit de deux matrices triangulaires inférieures (supérieures) est aussi une matrice triangulaire inférieure (supérieure).

1.3 Matrice transposée et matrice adjointe.

Soient $n,\ m,\ p \in \mathbb{N}^*$.

Soit $A \in \mathcal{M}_{n,m}(\mathbb{K})$. La matrice **transposée** de A est la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$, que l'on note A^T (ou A^t, ${}^TA$ ou ${}^tA$), définie par

$$A^T{}_{ij} = A_{ji},\ i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}.$$

Ainsi, les lignes de A^T sont les colonnes de A et les colonnes de A^T sont les lignes de A.

Si $A \in \mathcal{M}_{n,p}(\mathbb{K}),\ B \in \mathcal{M}_{p,m}(\mathbb{K})$, on a $A^T \in \mathcal{M}_{p,n}(\mathbb{K}),\ B^T \in \mathcal{M}_{m,p}(\mathbb{K})$ et :

$$\begin{aligned} &(A^T)^T = A; \\ &(AB)^T = B^T A^T. \end{aligned}$$

Soit $A \in \mathcal{M}_{n,m}(\mathbb{K})$. La matrice **adjointe** de A est la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$, que l'on note A^*, définie par

$$A^*{}_{ij} = \overline{A_{ji}},\ i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}.$$

Les coefficients de A^* sont alors les conjugués des coefficients de A^T. On remarque que si $A \in \mathcal{M}_{n,m}(\mathbb{R})$, alors $A^* = A^T$.

Si $A \in \mathcal{M}_{n,p}(\mathbb{K}),\ B \in \mathcal{M}_{p,m}(\mathbb{K})$ et $\lambda \in \mathbb{C}$, on a :

$$\begin{aligned} &(\lambda A)^* = \overline{\lambda} A^*; \\ &(A^*)^* = A; \\ &(AB)^* = B^* A^*. \end{aligned}$$

Exercice. Montrer les trois propriétés précédentes.

Notation. Si $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$, on identifie le vecteur x à la matrice colonne

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$$

et on note alors x^T la matrice ligne

$$x^T = [x_1 \cdots x_n] \in \mathcal{M}_{1,n}(\mathbb{K}).$$

De même, si $x \in \mathbb{K}$ est un scalaire, on identifie x à la matrice de taille 1×1, $[x] \in \mathcal{M}_{1,1}(\mathbb{K})$.

1.4 Produit scalaire et produit hermitien.

1.4.1 Produit scalaire.

Définition. *Produit scalaire.*
Soit E un espace vectoriel sur $\mathbb{R}$.

On appelle **produit scalaire** sur E une application

$$\begin{array}{rccc} (\cdot|\cdot) : & E \times E & \longrightarrow & \mathbb{R} \\ & (x, y) & \mapsto & (x|y) \end{array}$$

vérifiant :

$(\cdot|\cdot)$ est bilinéaire : pour tous $\alpha,\ \beta,\ \gamma,\ \delta \in \mathbb{R}$, pour tous $x,\ y,\ w,\ z \in E$,

$(\alpha x + \beta y|z) = \alpha(x|z) + \beta(y|z)$,

$(x|\gamma w + \delta z) = \gamma(x|w) + \delta(x|z)$;

$(\cdot|\cdot)$ est symétrique : pour tous $x,\ y \in E,\ (x|y) = (y|x)$;

$(\cdot|\cdot)$ est définie positive : pour tout $x \in E,\ x \neq 0,\ (x|x) > 0$.

Exemple. *Le produit scalaire euclidien.*
Le produit scalaire euclidien sur $\mathbb{R}^n$ est l'application $(\cdot|\cdot) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

$$(x|y) = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n,$$

pour tous $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{R}^n$.

Remarque.
En utilisant les identifications et notations introduites à la section 1.3, entre les vecteurs de $\mathbb{R}^n$ et les matrices ligne ou colonne à coefficients dans $\mathbb{R}$, on a, pour tous $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{R}^n$,

$$(x|y) = y^T x = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x_1 y_1 + \cdots + x_n y_n] \in \mathcal{M}_1(\mathbb{R}) \equiv \mathbb{R}.$$

1.4.2 Produit hermitien.

Dans toute cette section on considère des espaces vectoriels sur le corps $\mathbb{C}$.

Définition. *Produit scalaire complexe ou produit hermitien.*
Soit E un espace vectoriel sur $\mathbb{C}$.

On appelle **produit hermitien** (ou **produit scalaire complexe**) sur E une application

$$\begin{array}{rccc} (\cdot|\cdot) : & E \times E & \longrightarrow & \mathbb{C} \\ & (x, y) & \mapsto & (x|y) \end{array}$$

vérifiant :

$(\cdot|\cdot)$ est sesqui-linéaire : pour tous $\alpha,\ \beta,\ \gamma,\ \delta \in \mathbb{C}$, pour tous $x,\ y,\ w,\ z \in E$,

$(\alpha x + \beta y|z) = \alpha(x|z) + \beta(y|z)$,

$(x|\gamma w + \delta z) = \overline{\gamma}(x|w) + \overline{\delta}(x|z)$;

$(\cdot|\cdot)$ est hermitienne : pour tous $x,\ y \in E,\ (x|y) = \overline{(y|x)}$;

$(\cdot|\cdot)$ est définie positive : pour tout $x \in E,\ x \neq 0$, $(x|x) > 0$.

Remarque.
Si $(\cdot|\cdot)$ est un produit hermitien sur un espace vectoriel complexe E, alors pour tout $x \in E$, on a $(x|x) = \overline{(x|x)}$ (d'après la propriété $(x|y) = \overline{(y|x)}$, pour tous $x,\ y \in E$). On a donc que $(x|x) \in \mathbb{R}$, pour tout $x \in E$.

Exemple. *Le produit hermitien canonique sur $\mathbb{C}^n$.*
Le produit hermitien canonique sur $\mathbb{C}^n$ est l'application $(\cdot|\cdot)_{\mathbb{C}} : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ définie par

$$(x|y)_{\mathbb{C}} = \sum_{i=1}^{n} x_i \overline{y_i} = x_1 \overline{y_1} + \cdots + x_n \overline{y_n},$$

pour tous $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{C}^n$.

Exercice. Vérifier que l'application $(x, y) \in \mathbb{C}^n \times \mathbb{C}^n \mapsto (x|y)_{\mathbb{C}}$ est un produit hermitien sur $\mathbb{C}^n$.

Remarque.
Si on note, pour $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$, $\bar{x}$ le vecteur de $\mathbb{C}^n$ $(\overline{x_1}, \ldots, \overline{x_n})$, que l'on identifie à la matrice de $\mathcal{M}_{n,1}(\mathbb{C})$

$$\bar{x} = \begin{bmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{bmatrix},$$

on a, pour tous $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{C}^n$,

$$(x|y)_{\mathbb{C}} = \bar{y}^T x = \begin{bmatrix} \overline{y_1} & \cdots & \overline{y_n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \overline{y_1} + \cdots + x_n \overline{y_n}] \in \mathcal{M}_1(\mathbb{C}) \equiv \mathbb{C}.$$

On remarque que si E est un espace vectoriel sur $\mathbb{K} = \mathbb{R}$, les trois propriétés que le produit hermitien vérifie correspondent aux trois propriétés d'un produit scalaire. On peut donc, pour un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, parler sans ambiguité de produit hermitien, car dans le cas $\mathbb{K} = \mathbb{R}$ les deux produits sont définis de la même manière.

On rappelle maintenant que l'on peut définir une norme à partir de tout produit produit hermitien (ou scalaire) dans un espace vectoriel complexe (ou réel).

Définition - proposition. *Norme induite par un poduit scalaire ou hermitien.*
Soit E un espace vectoriel sur $\mathbb{K}$ et soit $(\cdot|\cdot) : E \times E \longrightarrow \mathbb{K}$ un produit hermitien sur E (et donc un produit scalaire si $K = \mathbb{R}$, d'après la remarque ci-dessus). Alors l'application $\|\cdot\| : E \longrightarrow \mathbb{R}^+$ définie par

$$\|x\| = \sqrt{(x|x)}, \text{ pour tout } x \in E,$$

est une norme sur E. On a en plus l'**inégalité de Cauchy-Schwarz** :

$$|(x|y)| \leq \|x\|\|y\|, \text{ pour tous } x,\ y \in E,$$

et $|(x|y)| = \|x\|\|y\|$ si et seulement si x et y sont colinéaires.

Démonstration. On a d'abord, pour $x \in E$,

$$\|x\| = 0 \text{ ssi } (x|x) = 0 \text{ ssi } x = 0.$$

Ensuite, si $\lambda \in \mathbb{K}$ et $x \in E$,

$$\|\lambda x\| = \sqrt{(\lambda x|\lambda x)} = \sqrt{\lambda\bar{\lambda}(x|x)} = \sqrt{|\lambda|^2(x|x)} = |\lambda|\|x\|.$$

Pour montrer que l'application $\|\cdot\|$ vérifie l'inégalité triangulaire, on utilise l'inégalité de Cauchy-Schwarz, que l'on montrera ensuite. Supposons alors que l'inégalité de Cauchy-Schwarz est vérifiée pour tous $x,\ y \in E$. On a alors, si $x,\ y \in E$,

$$\begin{aligned}
\|x+y\|^2 &= (x+y|x+y) \\
&= (x|x) + (x|y) + (y|x) + (y|y) \\
&= (x|x) + (x|y) + \overline{(x|y)} + (y|y) \\
&= (x|x) + 2\mathrm{Re}\big((x|y)\big) + (y|y), && \text{car } z + \bar{z} = 2\mathrm{Re}(z), \text{ pour tout } z \in \mathbb{C}, \\
&\leq \|x\|^2 + 2|(x|y)| + \|y\|^2, && \text{car } \mathrm{Re}(z) \leq |z|, \text{ pour tout } z \in \mathbb{C}, \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2, && \text{par l'inégalité de Cauchy-Schwarz,} \\
&= \big(\|x\| + \|y\|\big)^2,
\end{aligned}$$

et donc $\|x+y\| \leq \|x\| + \|y\|$.

Preuve de l'inégalité de Cauchy-Schwarz.

Soient $x,\ y \in E$. Si $x = 0$, il est immédiat que l'inégalité est vérifiée (car $(x|y) = 0$ et $\|x\|\|y\| = 0$).
Supposons ainsi que $x \neq 0$. On a, quelque soit $\alpha \in \mathbb{K}$,

$$y = \alpha x + (y - \alpha x).$$

On cherche $\alpha \in \mathbb{K}$ tel que $z = y - \alpha x$ vérifie $(z|x) = 0$ (on cherche alors à écrire y comme combinaison linéaire d'un vecteur αx, colinéaire à x, et d'un vecteur $z \in E$ tel que $(z|x) = 0$). On doit alors avoir $(y - \alpha x|x) = 0$, c'est-à-dire $(y|x) - \alpha(x|x) = 0$ et on obtient $\alpha = \dfrac{(y|x)}{(x|x)} = \dfrac{(y|x)}{\|x\|^2}$. On a alors $y = \alpha x + z$ avec

$$\alpha = \frac{(y|x)}{\|x\|^2}, \quad z = y - \frac{(y|x)}{\|x\|^2}x \text{ vérifiant } (z|x) = 0.$$

On en déduit que

$$\begin{aligned}
\|y\|^2 &= (\alpha x + z|\alpha x + z) \\
&= \alpha\bar{\alpha}\|x\|^2 + \alpha(x|z) + \bar{\alpha}(z|x) + \|z\|^2 \\
&= |\alpha|^2\|x\|^2 + \|z\|^2, && \text{car } (z|x) = 0 \text{ et } (x|z) = \overline{(z|x)} = 0, \\
&\geq |\alpha|^2\|x\|^2
\end{aligned}$$

(avec égalité si et seulement si $\|z\| = 0$). On a donc que

$$|\alpha|^2 \leq \frac{\|y\|^2}{\|x\|^2},$$

c'est-à-dire que

$$\frac{|(y|x)|^2}{\|x\|^4} \leq \frac{\|y\|^2}{\|x\|^2},$$

c'est-à-dire que

$$|(y|x)|^2 \leq \|y\|^2\|x\|^2,$$

et donc $|(y|x)| \leq \|y\|\|x\|$. On a par ailleurs que l'égalité est obtenue si et seulement si $z = 0$, c'est-à-dire si et seulement si $y = \alpha x$ c'est à dire si et seulement si y et x sont colinéaires. $\square$

Définition. *Vecteurs orthogonaux et famille orthonormale.*
Soit E un espace vectoriel sur $\mathbb{K}$ et $(\cdot|\cdot)$ un produit hermitien (scalaire si $\mathbb{K} = \mathbb{R}$) sur E. On dit que deux vecteurs $x,\ y \in E$ sont orthogonaux par rapport au produit $(\cdot|\cdot)$ si $(x|y) = 0$. On dit qu'une famille $(x_1, \ldots, x_p),\ p \in \mathbb{N},\ p \geq 2$, de vecteurs de E, est orthonormale si pour tous $i, j \in \{1, \ldots, p\},\ i \neq j$, $(x_i|x_j) = 0$ et $(x_i|x_i) = 1$.

Exercice. Soit E un espace vectoriel sur $\mathbb{K}$, $(\cdot|\cdot)$ un produit hermitien sur E et $x,\ y \in E$ deux vecteurs orthogonaux non nuls. Montrer que la famille (x, y) est libre. Conclure que dans un espace vectoriel E de dimension finie égale à $n \in \mathbb{N}$, toute famille orthonormale de n vecteurs est une base de E.

Pour finir cette section, on rappelle le procédé de Gram-Schmidt permettant d'obtenir une famille de vecteurs orthonormale dans un espace E muni d'un produit hermitien.

Proposition. *Procédé d'orthonormalisation de Gram-Schmidt.*

Soit E un espace vectoriel sur $\mathbb{K}$ muni d'un produit hermitien $(\cdot|\cdot)$ et $\|\cdot\|$ la norme induite par ce produit hermitien.

Soit $(x_1, \dots, x_n)$ une famille libre de E.

Soit $(y_1, \dots, y_n)$ la famille de vecteurs de E définie par récurrence par :

$$y_1 = \frac{x_1}{\|x_1\|},$$

$$\text{pour } i \in \{2, \dots, n\}, \qquad y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}, \text{ où}$$

$$\tilde{y}_i = x_i - \big((x_i|y_1)y_1 + \cdots + (x_i|y_{i-1})y_{i-1}\big).$$

Alors la famille $(y_1, \dots, y_n)$ est une famille orthonormale de vecteurs de E.

La preuve est laissée comme exercice.

1.4.3 Quelques propriétés des matrices en lien avec le produit scalaire et hermitien.

Dans toute la suite, $(\cdot|\cdot)$ désigne le produit scalaire euclidien sur $\mathbb{R}^n$ et $(\cdot|\cdot)_{\mathbb{C}}$ le produit hermitien canonique sur $\mathbb{C}^n$.

Nous allons montrer un résultat qui permet de caractériser la transposée d'une matrice réelle et de l'adjointe d'une matrice complexe via respectivement le produit scalaire euclidien et le produit hermitien.

Proposition. *Caractérisation de la tranposée par le produit scalaire, de l'adjointe par le produit hermitien.*

Cas réel.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ et $A^T \in \mathcal{M}_n(\mathbb{R})$ sa transposée, définie par ${A^T}_{i,j} = A_{j,i}$, pour tous $i,\ j \in \{1, \ldots, n\}$.
Soit $(\cdot|\cdot)$ le produit scalaire euclidien sur $\mathbb{R}^n$. Alors on a

$$(Au|v) = (u|A^T v), \text{ pour tous } u, v \in \mathbb{R}^n,$$

et la matrice A^T est l'unique matrice $B \in \mathcal{M}_n(\mathbb{R})$ tel que

$$(Au|v) = (u|Bv), \text{ pour tous } u, v \in \mathbb{R}^n.$$

Cas complexe.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $A^* \in \mathcal{M}_n(\mathbb{C})$ son adjointe, définie par ${A^*}_{i,j} = \overline{A_{j,i}}$, pour tous $i,\ j \in \{1, \ldots, n\}$.
Soit $(\cdot|\cdot)_{\mathbb{C}}$ le produit hermitien canonique sur $\mathbb{C}^n$. Alors on a

$$(Au|v)_{\mathbb{C}} = (u|A^* v)_{\mathbb{C}}, \text{ pour tous } u, v \in \mathbb{C}^n,$$

et la matrice A^* est l'unique matrice $B \in \mathcal{M}_n(\mathbb{C})$ tel que

$$(Au|v)_{\mathbb{C}} = (u|Bv)_{\mathbb{C}}, \text{ pour tous } u, v \in \mathbb{C}^n.$$

Démonstration. On va faire la preuve dans le cas complexe, le cas réel étant une conséquence immédiate du fait que si $A \in \mathcal{M}_n(\mathbb{R})$, alors $A^* = A^T$ et si $u, v \in \mathbb{R}^n$ alors $(u|v)_{\mathbb{C}} = (u|v)$.

Soient $u = (u_1, \ldots, u_n) \in \mathbb{C}^n$ et $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$. On a

$$(Au|v)_{\mathbb{C}} = \sum_{i=1}^{n} (Au)_i \overline{v_i} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} u_j \overline{v_i}.$$

D'autre part,

$$\begin{aligned}
(u|A^*v)_{\mathbb{C}} &= \sum_{j=1}^{n} u_j \overline{(A^*v)_j} = \sum_{j=1}^{n} u_j \overline{\sum_{i=1}^{n} {A^*}_{j,i} v_i} \\
&= \sum_{j=1}^{n} u_j \sum_{i=1}^{n} \overline{{A^*}_{j,i}}\, \overline{v_i} \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} u_j \overline{\overline{A_{i,j}}}\, \overline{v_i} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} u_j \overline{v_i}.
\end{aligned}$$

On a donc $(Au|v)_{\mathbb{C}} = (u|A^*v)_{\mathbb{C}}$.

Soit maintenant $B \in \mathcal{M}_n(\mathbb{C})$ une matrice telle que

$$(Au|v)_{\mathbb{C}} = (u|Bv)_{\mathbb{C}}, \text{ pour tous } u, v \in \mathbb{C}^n.$$

Montrons que $B = A^*$, c'est-à-dire que $B_{i,j} = {A^*}_{i,j} = \overline{A_{j,i}}$, pour tous $i, j \in \{1, \ldots, n\}$.

Soient $i, j \in \{1, \ldots, n\}$ et e_i et e_j respectivement les ième et jème vecteurs de la base canonique de $\mathbb{C}^n$, définis respectivement par $(e_i)_k = 0$, si $k \neq i$, $(e_i)_i = 1$, et par $(e_j)_k = 0$, si $k \neq j$, $(e_j)_j = 1$. On a alors

$$(Ae_i|e_j)_{\mathbb{C}} = (e_i|Be_j)_{\mathbb{C}}.$$

Or

$$\begin{aligned}
(Ae_i|e_j)_{\mathbb{C}} &= \sum_{k=1}^{n} (Ae_i)_k \overline{(e_j)_k} \\
&= (Ae_i)_j, && \text{car } (e_j)_k = 0 \text{ si } k \neq j \text{ et } (e_j)_j = 1, \\
&= \sum_{k=1}^{n} A_{j,k} (e_i)_k \\
&= A_{j,i}, && \text{car } (e_i)_k = 0 \text{ si } k \neq i \text{ et } (e_i)_i = 1,
\end{aligned}$$

et

$$\begin{aligned}
(e_i|Be_j)_{\mathbb{C}} &= \sum_{k=1}^{n} (e_i)_k \overline{(Be_j)_k} \\
&= \overline{(Be_j)_i} \\
&= \overline{\sum_{k=1}^{n} B_{i,k} (e_j)_k} \\
&= \overline{B_{i,j}}.
\end{aligned}$$

On conclut que $A_{j,i} = \overline{B_{i,j}}$ et donc que $B_{i,j} = \overline{A_{j,i}}$. $\square$

La proposition précédente montre que l'on peut définir la transposée d'une matrice réelle A comme étant l'unique matrice B vérifiant $(Au|v) = (u|Bv)$, pour tous $u, v \in \mathbb{R}^n$. De même, l'adjointe d'une matrice complexe A est l'unique matrice B vérifiant $(Au|v)_{\mathbb{C}} = (u|Bv)_{\mathbb{C}}$, pour tous $u, v \in \mathbb{C}^n$.

On donne maintenant quelques définitions pour lesquelles on peut aussi donner une caractérisation par le produit scalaire ou hermitien. On commence par le cas réel.

Définition.
Soit $A \in \mathcal{M}_n(\mathbb{R})$.

- On dit que A est **symétrique** si $A^T = A$, autrement dit si $A_{i,j} = A_{j,i}$, pour tous $i,\ j \in \{1, \ldots, n\}$.
- On dit que A est **orthogonale, ou unitaire,** si $A^T A = AA^T = I_n$, autrement dit si A est inversible et $A^{-1} = A^T$.
- On dit que A est **normale** si $AA^T = A^T A$.

Supposons $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique. On a alors $A = A^T$ et, d'après la proposition précédente,

$$(Au|v) = (u|Av), \text{ pour tous } u, v \in \mathbb{R}^n.$$

D'autre part, si $A \in \mathcal{M}_n(\mathbb{R})$ vérifie

$$(Au|v) = (u|Av), \text{ pour tous } u, v \in \mathbb{R}^n,$$

comme A^T est l'unique matrice B vérifiant

$$(Au|v) = (u|Bv), \text{ pour tous } u, v \in \mathbb{R}^n,$$

on conclut que $A = A^T$.

On a alors le résultat suivant, qui donne une autre caractérisation possible pour les matrices symétriques réelles.

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors A est symétrique si et seulement si

$$(Au|v) = (u|Av), \text{ pour tous } u, v \in \mathbb{R}^n.$$

Supposons maintenant $A \in \mathcal{M}_n(\mathbb{R})$. Soient $l_1, \ldots, l_n \in \mathbb{R}^n$ les lignes de la matrice A et $c_1, \ldots, c_n \in \mathbb{R}^n$ les colonnes de la matrice A :

$$A = \left[\begin{array}{c} l_1 \\ \hline \vdots \\ \hline l_n \end{array} \right] = \left[\begin{array}{c|c|c} c_1 & \cdots & c_n \end{array} \right].$$

On a

$$AA^T = I_n \Longleftrightarrow \begin{bmatrix} l_1 \\ \hline \vdots \\ \hline l_n \end{bmatrix} \left[\begin{array}{c|c|c} l_1 & \cdots & l_n \end{array}\right] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

et

$$A^T A = I_n \Longleftrightarrow \begin{bmatrix} c_1 \\ \hline \vdots \\ \hline c_n \end{bmatrix} \left[\begin{array}{c|c|c} c_1 & \cdots & c_n \end{array}\right] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}.$$

On a donc que

$$\begin{aligned} AA^T = I_n &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ (AA^T)_{i,j} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ \sum_{k=1}^{n} (l_i)_k (l_j)_k = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ (l_i | l_j) = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \end{aligned}$$

et

$$\begin{aligned} A^T A = I_n &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ (A^T A)_{i,j} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ \sum_{k=1}^{n} (c_i)_k (c_j)_k = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\Longleftrightarrow \text{ pour tous } i,\ j \in \{1, \ldots, n\},\ (c_i | c_j) = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} . \end{aligned}$$

On conclut ainsi la caractérisation suivante des matrices orthogonales.

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors A est orthogonale, c'est-à-dire que $AA^T = A^T A = I_n$, si et seulement si la famille $(l_1, \ldots, l_n)$ de vecteurs de $\mathbb{R}^n$ formée par les lignes de A est une famille orthonormale de $\mathbb{R}^n$, et donc une base de $\mathbb{R}^n$, si et seulement si la famille $(c_1, \ldots, c_n)$ de vecteurs de $\mathbb{R}^n$ formée par les colonnes de A est une famille orthonormale de $\mathbb{R}^n$, et donc une base de $\mathbb{R}^n$.

Donnons maintenant les définitions et propriétés analogues pour les matrices complexes.

Définition.
Soit $A \in \mathcal{M}_n(\mathbb{C})$.

- On dit que A est **hermitienne** ou **auto-adjointe** si $A^* = A$, autrement dit si $A_{i,j} = \overline{A_{j,i}}$, pour tous $i,\ j \in \{1, \dots, n\}$.
- On dit que A est **unitaire,** si $A^*A = AA^* = I_n$, autrement dit si A est inversible et $A^{-1} = A^*$.
- On dit que A est **normale** si $AA^* = A^*A$.

On peut argumenter comme dans le cas réel pour conclure le résultat suivant, qui donne une autre caractérisation possible pour les matrices complexes hermitiennes.

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors A est hermitienne si et seulement si

$$(Au|v)_{\mathbb{C}} = (u|Av)_{\mathbb{C}}, \text{ pour tous } u, v \in \mathbb{C}^n.$$

Supposons maintenant $A \in \mathcal{M}_n(\mathbb{C})$ et notons à nouveau $l_1, \dots, l_n \in \mathbb{C}^n$ les lignes de la matrice A et $c_1, \dots, c_n \in \mathbb{C}^n$ les colonnes de la matrice A.

On a

$$AA^* = I_n \iff \left[\begin{array}{c} l_1 \\ \hline \vdots \\ \hline l_n \end{array}\right] \left[\begin{array}{c|c|c} \overline{l_1} & \cdots & \overline{l_n} \end{array}\right] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

et

$$A^*A = I_n \iff \left[\begin{array}{c} \overline{c_1} \\ \hline \vdots \\ \hline \overline{c_n} \end{array}\right] \left[\begin{array}{c|c|c} c_1 & \cdots & c_n \end{array}\right] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix},$$

où l'on rappelle que pour un vecteur $x = (x_1, \dots, x_n) \in \mathbb{C}^n$, $\overline{x}$ dénote le vecteur $(\overline{x_1}, \dots, \overline{x_n})$.

On a donc que

$$\begin{aligned} AA^* = I_n &\iff \text{ pour tous } i,\ j \in \{1, \dots, n\},\ (AA^*)_{i,j} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\iff \text{ pour tous } i,\ j \in \{1, \dots, n\},\ \sum_{k=1}^{n} (l_i)_k \overline{(l_j)_k} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \\ &\iff \text{ pour tous } i,\ j \in \{1, \dots, n\},\ (l_i|l_j)_{\mathbb{C}} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} \end{aligned}$$

et, de manière analogue,

$$A^*A = I_n \iff \text{ pour tous } i,\ j \in \{1, \dots, n\},\ (c_j|c_i)_{\mathbb{C}} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases} .$$

On conclut ainsi la caractérisation suivante des matrices unitaires complexes.

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors A est unitaire, c'est-à-dire que $AA^* = A^*A = I_n$, si et seulement si la famille $(l_1, \dots, l_n)$ de vecteurs de $\mathbb{C}^n$ formée par les lignes de A est une famille orthonormale de $\mathbb{C}^n$, si et seulement si la famille $(c_1, \dots, c_n)$ de vecteurs de $\mathbb{C}^n$ formée par les colonnes de A est une famille orthonormale de $\mathbb{R}^n$ (et donc une base de $\mathbb{C}^n$).

1.5 Matrices symétriques ou hermitiennes définies positives.

On rappelle dans cette section la définition de matrice symétrique réelle définie positive et de matrice hermitienne complexe définie positive.

Définition. *Matrice définie positive - cas réel.*
Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique réelle.
On dit que A est **définie positive** si

$$\text{pour tout } x \in \mathbb{R}^n,\ x \neq 0,\ x^T A x > 0.$$

On dit que A est **positive** si

$$\text{pour tout } x \in \mathbb{R}^n,\ x^T A x \geq 0.$$

Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $(\cdot|\cdot)$ le produit scalaire euclidien sur $\mathbb{R}^n$. Alors

$$\begin{aligned} x^T A x &= \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} (Ax)_1 \\ \vdots \\ (Ax)_n \end{bmatrix} \\ &= \sum_{i=1}^n x_i (Ax)_i \\ &= (Ax|x). \end{aligned}$$

On a ainsi qu'une matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$ est définie positive (resp. positive) si et seulement si

$$(Ax|x) > 0, \text{ pour tout } x \in \mathbb{R}^n, \; x \neq 0, \quad \big(\text{resp. } (Ax|x) \geq 0, \text{ pour tout } x \in \mathbb{R}^n\big).$$

La définie positivité des matrices à coefficients complexes se définit de la même manière, mais justifions d'abord que si $A \in \mathcal{M}_n(\mathbb{C})$ est une matrice hermitienne, c'est-à-dire telle que $A^* = A$, alors pour tout $x \in \mathbb{C}^n$, $\bar{x}^T A x \in \mathbb{R}$.

Soit $(\cdot|\cdot)_{\mathbb{C}}$ le produit hermitien canonique sur $\mathbb{C}^n$ et $x \in \mathbb{C}^n$. On a

$$\begin{aligned}
\bar{x}^T A x &= \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{bmatrix} \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{bmatrix} \begin{bmatrix} (Ax)_1 \\ \vdots \\ (Ax)_n \end{bmatrix} \\
&= \sum_{i=1}^{n} \overline{x_i} (Ax)_i \\
&= (Ax|x)_{\mathbb{C}} = (x|A^* x)_{\mathbb{C}} \\
&= (x|Ax)_{\mathbb{C}}, && \text{car } A \text{ est hermitienne,} \\
&= \overline{(Ax|x)_{\mathbb{C}}} = \overline{\bar{x}^T A x}.
\end{aligned}$$

On a alors que $\bar{x}^T A x$ est égal à son conjugué et donc $\bar{x}^T A x \in \mathbb{R}$.

Définition. *Matrice définie positive - cas complexe.*
Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice complexe hermitienne.

On dit que A est **définie positive** si

$$\text{pour tout } x \in \mathbb{C}^n, \; x \neq 0, \; \bar{x}^T A x > 0.$$

On dit que A est **positive** si

$$\text{pour tout } x \in \mathbb{C}^n, \; \bar{x}^T A x \geq 0.$$

On vient de montrer plus haut que pour tout $x \in \mathbb{C}^n$, $\bar{x}^T A x = (Ax|x)_{\mathbb{C}}$. On a donc qu'une matrice $A \in \mathcal{M}_n(\mathbb{C})$ hermitienne est définie positive (resp. positive) si et seulement si

$$(Ax|x)_{\mathbb{C}} > 0, \text{ pour tout } x \in \mathbb{C}^n, \; x \neq 0, \quad \big(\text{resp. } (Ax|x)_{\mathbb{C}} \geq 0, \text{ pour tout } x \in \mathbb{C}^n\big).$$

1.6 Théorie spectrale des matrices. Vecteurs propres et valeurs propres.

Définition. *Valeurs propres, vecteurs propres, spectre, et polynôme caractéristique.*
Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On dit que $\lambda \in \mathbb{K}$ est une **valeur propre** de A s'il existe $u \in \mathbb{K}^n,\ u \neq 0$, tel que $Au = \lambda u$. Le vecteur u s'appelle **vecteur propre** associé à la valeur propre λ.

On appelle **spectre** de A l'ensemble des valeurs propres de A, et on le note spec(A) ou $\sigma(A)$:

$$\mathrm{spec}(A) = \{\lambda \in \mathbb{K} \,|\, \exists u \in \mathbb{K}^n,\ u \neq 0,\ Au = \lambda u\}.$$

On appelle **polynôme caractéristique** de A le polynôme $P_A \in \mathbb{K}^n[\lambda]$ défini par

$$P_A(\lambda) = \det(A - \lambda I_n).$$

Si $\lambda \in \mathbb{K}$ est une valeur propre de A et $u \in \mathbb{K}^n$ un vecteur propre associé, le sous-espace vectoriel de $\mathbb{K}^n$ vect(u) engendré par u est stable par A, vue en tant qu'application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^n$, et celle-ci agit comme une dilatation sur ce sous-espace.

On remarque que l'on a

$$\begin{aligned}
\lambda \in \mathbb{K} \text{ est une valeur propre de } A \quad & \text{ssi} \quad \exists u \in \mathbb{K}^n,\ u \neq 0,\ Au = \lambda u \\
& \text{ssi} \quad \exists u \in \mathbb{K}^n,\ u \neq 0,\ (A - \lambda I_n)u = 0 \\
& \text{ssi} \quad \ker(A - \lambda I_n) \neq \{0_{\mathbb{K}^n}\} \\
& \text{ssi} \quad A - \lambda I_n \text{ n'est pas inversible} \\
& \text{ssi} \quad \det(A - \lambda I_n) = 0 \\
& \text{ssi} \ \lambda \text{ est une racine du polynôme } P_A,\ i.\ e.\ P_A(\lambda) = 0.
\end{aligned}$$

Toute matrice A à coefficients dans $\mathbb{R}$ est aussi à coefficients dans $\mathbb{C}$. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ peut admettre ou non des valeurs propres dans $\mathbb{R}$, mais en tant que matrice de $\mathcal{M}_n(\mathbb{C})$, elle admet n valeurs propres dans $\mathbb{C}$, distinctes ou non, qui sont les n racines complexes du polynôme caractéristique P_A de A, comptées avec leur multiplicité en tant que racines de P_A.

Par exemple la matrice de $\mathcal{M}_2(\mathbb{R})$

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

a comme valeurs propres i et $-i$; on a donc spec$(A) \subseteq \mathbb{C} \backslash \mathbb{R}$, et A n'a pas de valeurs propres dans $\mathbb{R}$.

Dans la suite de cette section, et sauf mention en contraire, on considérera toute matrice à coefficients réels ou complexes comme un élément de $\mathcal{M}_n(\mathbb{C})$.

Par ailleurs on notera $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ les n valeurs propres de $A \in \mathcal{M}_n(\mathbb{C})$ (non nécessairement distinctes), comptées avec leur multiplicité en tant que racines de $P_A(\lambda)$. Par exemple si $A \in \mathcal{M}_4(\mathbb{R})$ est la matrice diagonale dont la diagonale est $(1, 1, 1, 2)$, on a $P_A(\lambda) = (\lambda - 1)^3(\lambda - 2)$ et on dira que les 4 valeurs propres de A sont $1,\ 1,\ 1$ et 2.

On va par la suite montrer deux propriétés que l'on utilisera régulièrement dans la suite du cours.

Proposition. *Deux matrices semblables ont les mêmes valeurs propres.*
Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ tel qu'il existe $P \in \mathcal{M}_n(\mathbb{K})$ inversible tel que $B = P^{-1}AP$. Alors A et B ont les mêmes valeurs propres.

Démonstration. Soit $\lambda \in \mathbb{C}$ et $u \in \mathbb{C}^n,\ u \neq 0$. Alors on a, puisque P est inversible,

$$\begin{aligned} Au = \lambda u \ &\text{ssi}\ APP^{-1}u = \lambda PP^{-1}u \\ &\text{ssi}\ P^{-1}APP^{-1}u = \lambda P^{-1}PP^{-1}u \\ &\text{ssi}\ P^{-1}AP(P^{-1}u) = \lambda P^{-1}u. \end{aligned}$$

On a donc que $\lambda \in \mathbb{C}$ est une valeur propre de A, et $u \in \mathbb{C}^n \backslash \{0\}$ un vecteur propre de A associé à la valeur propre λ, si et seulement si λ est une valeur propre de $B = P^{-1}AP$, et $v = P^{-1}u \in \mathbb{C}^n \backslash \{0\}$ est un vecteur propre de B associé à la valeur propre λ (on remarque que comme P^{-1} est une matrice inversible, on a $P^{-1}u \neq 0$ ssi $u \neq 0$). $\square$

Proposition. *Les valeurs propres d'une matrice symétrique ou d'une matrice hermitienne sont réelles.*
Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique (c'est-à-dire telle que $A = A^T$) ou $A \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne (c'est-à-dire telle que $A = A^*$). Alors $\mathrm{spec}(A) \subseteq \mathbb{R}$, autrement dit les valeurs propres de A sont réelles.

Démonstration. Supposons $A \in \mathcal{M}_n(\mathbb{K})$ une telle matrice. Soit $\lambda \in \mathbb{C}$ une valeur propre de A. Montrons que $\lambda = \bar{\lambda}$, ce qui prouve que $\lambda \in \mathbb{R}$.

Il existe $u \in \mathbb{C}^n,\ u \neq 0$, tel que $Au = \lambda u$. On a alors, d'une part,

$$(Au|u)_{\mathbb{C}} = (\lambda u|u)_{\mathbb{C}} = \lambda(u|u)_{\mathbb{C}},$$

et, d'autre part,

$$(Au|u)_{\mathbb{C}} = (u|A^*u)_{\mathbb{C}} = (u|\lambda u)_{\mathbb{C}} = \bar{\lambda}(u|u)_{\mathbb{C}}.$$

On conclut que

$$\lambda(u|u)_{\mathbb{C}} = \bar{\lambda}(u|u)_{\mathbb{C}},$$

et, comme $u \neq 0$, on a $(u|u)_{\mathbb{C}} > 0$ et donc on obtient $\lambda = \bar{\lambda}$. $\square$

Remarque.
On vient de voir que si $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique alors les n valeurs propres de A sont réelles. On remarque alors que si λ est une valeur propre de A, il existe forcément un vecteur propre $u \in \mathbb{R}^n$ associé à λ. En effet, soit $u \in \mathbb{C}^n,\ u \neq 0$, tel que $Au = \lambda u$. Alors $u = u_a + iu_b$, avec $u_a,\ u_b \in \mathbb{R}^n$, et comme $u \neq 0$ on a $u_a \neq 0$ ou $u_b \neq 0$. On a aussi $Au_a + iAu_b = \lambda u_a + i\lambda u_b$ et, comme $\lambda \in \mathbb{R}$, $Au_a = \lambda u_a$ et $Au_b = \lambda u_b$, c'est-à-dire que $u_a \in \mathbb{R}^n$ ou $u_b \in \mathbb{R}^n$ est un vecteur propre associé à λ.

1.6.1 Trigonalisation et diagonalisation.

Définition.
Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On dit que A est **trigonalisable** dans $\mathbb{K}$ s'il existe une matrice $P \in \mathcal{M}_n(\mathbb{K})$ inversible et une matrice $T \in \mathcal{M}_n(\mathbb{K})$ triangulaire tel que $A = PTP^{-1}$, autrement dit si A est semblable dans $\mathbb{K}$ à une matrice triangulaire.

On dit que A est **diagonalisable** dans $\mathbb{K}$ s'il existe une matrice $P \in \mathcal{M}_n(\mathbb{K})$ inversible et une matrice $D \in \mathcal{M}_n(\mathbb{K})$ diagonale tel que $A = PDP^{-1}$, autrement dit si A est semblable dans $\mathbb{K}$ à une matrice diagonale.

Remarque.
Si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable et semblable à la matrice triangulaire $T \in \mathcal{M}_n(\mathbb{K})$, alors les valeurs propres de A et de T sont les mêmes, et donc les valeurs propres de A sont les éléments diagonaux de la matrice T (car si T est triangulaire, on a $P_T(\lambda) = \prod_{i=1}^{n}(\lambda - T_{ii})$).
De même, si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable et semblable à la matrice diagonale $D \in \mathcal{M}_n(\mathbb{K})$, alors les valeurs propres de A et de D sont les mêmes, et donc les valeurs propres de A sont les éléments diagonaux de la matrice D.

On a que toute matrice A est trigonalisable dans $\mathbb{C}$:

Proposition.
Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable dans $\mathbb{C}$: pour tout $A \in \mathcal{M}_n(\mathbb{K})$, il existe $P \in \mathcal{M}_n(\mathbb{C})$ inversible, $T \in \mathcal{M}_n(\mathbb{C})$ triangulaire, tel que $A = PTP^{-1}$.

Démonstration. On va montrer par induction sur $n \in \mathbb{N}^*$ que pour tout $A \in \mathcal{M}_n(\mathbb{K})$ il existe une matrice triangulaire supérieure $T \in \mathcal{M}_n(\mathbb{C})$ tel que A est semblable à T dans $\mathbb{C}$.

Si $n = 1$ le résultat est vrai : si $A = [a] \in \mathcal{M}_1(\mathbb{K})$, avec $a \in \mathbb{K}$, on a $A = PTP^{-1}$, avec $P = [1]$ et $T = [a]$ matrice triangulaire.

Supposons que le résultat est vrai pour toute matrice $A \in \mathcal{M}_{n-1}(\mathbb{K})$, $n \in \mathbb{N},\ n \geq 2$. Montrons qu'il est vrai pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{C}$ une valeur propre de A. Il existe $v \in \mathbb{C}^n \backslash \{0\}$ tel que $Av = \lambda v$. Soient alors $(u_2, \ldots, u_n) \in \mathbb{C}^n$ tel que $(v, u_2, \ldots, u_n)$ est une base de $\mathbb{C}^n$. On remarque que comme

$$\left[\begin{array}{c} \\ A \\ \\ \end{array}\right]\left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right] = \left[\begin{array}{c|c|c|c} Av & Au_2 & \cdots & Au_n \end{array}\right] = \left[\begin{array}{c|c|c|c} \lambda v & Au_2 & \cdots & Au_n \end{array}\right],$$

il existe alors $a_2, \ldots, a_n \in \mathbb{C}$ et $\tilde{A} \in \mathcal{M}_{n-1}(\mathbb{C})$ tel que

$$\left[\begin{array}{c} \\ A \\ \\ \end{array}\right]\left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right] = \left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right]\left[\begin{array}{c|ccc} \lambda & a_2 & \cdots & a_n \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right].$$

Comme $\tilde{A} \in \mathcal{M}_{n-1}(\mathbb{C})$, par hypothèse d'induction il existe une matrice $\tilde{P} \in \mathcal{M}_{n-1}(\mathbb{C})$ inversible et une matrice $\tilde{T} \in \mathcal{M}_{n-1}(\mathbb{C})$ triangulaire supérieure tel que $\tilde{A} = \tilde{P}\tilde{T}\tilde{P}^{-1}$. On en déduit que

$$\begin{aligned}\left[\begin{array}{c} \\ A \\ \\ \end{array}\right]\left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right] &= \left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right]\left[\begin{array}{c|ccc} \lambda & a_2 & \cdots & a_n \\ \hline 0 & & & \\ \vdots & & \tilde{P}\tilde{T}\tilde{P}^{-1} & \\ 0 & & & \end{array}\right] \\ &= \left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right]\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{array}\right]\left[\begin{array}{c|ccc} \lambda & a_2 & \cdots & a_n \\ \hline 0 & & & \\ \vdots & & \tilde{T} & \\ 0 & & & \end{array}\right]\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}^{-1} & \\ 0 & & & \end{array}\right].\end{aligned}$$

On pose

$$U := \left[\begin{array}{c|c|c|c} v & u_2 & \cdots & u_n \end{array}\right], \quad \tilde{U} := \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{array}\right] \quad \text{et} \quad T := \left[\begin{array}{c|ccc} \lambda & a_2 & \cdots & a_n \\ \hline 0 & & & \\ \vdots & & \tilde{T} & \\ 0 & & & \end{array}\right].$$

Comme $(v, u_2, \ldots, u_n)$ est une base de $\mathbb{C}^n$, $U \in \mathcal{M}_n(\mathbb{C})$ est inversible. Comme $\det(\tilde{U}) = \det(\tilde{P})$ et $\tilde{P}$ est inversible, on a que $\tilde{U} \in \mathcal{M}_n(\mathbb{C})$ est inversible, avec

$$\tilde{U}^{-1} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}^{-1} & \\ 0 & & & \end{array}\right].$$

Finalement, comme $\tilde{T}$ est triangulaire supérieure, $T \in \mathcal{M}_n(\mathbb{C})$ l'est aussi. On a alors $AU = U\tilde{U}T\tilde{U}^{-1}$, c'est-à-dire que

$$A = (U\tilde{U})T(\tilde{U}^{-1}U^{-1}) = (U\tilde{U})T((U\tilde{U})^{-1}) = PTP^{-1},$$

avec $P = U\tilde{U} \in \mathcal{M}_n(\mathbb{C})$ inversible et $T \in \mathcal{M}_n(\mathbb{C})$ triangulaire. □

Attention, il n'est pas vrai que toute matrice est diagonalisable. Par exemple la matrice

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

n'est semblable à aucune matrice diagonale D, car si elle l'était, cette matrice diagonale D ne pourrait être que la matrice nulle (car D contiendrait dans la diagonale les valeurs propres de A et la valeur propre double 0 est l'unique valeur propre de A) ; on aurait alors que A serait semblable à la matrice nulle et A serait alors elle même la matrice nulle, ce qui n'est pas le cas.

Nous verrons par la suite que les matrices symétriques réelles et hermitiennes complexes, et plus généralement les matrices normales, sont diagonalisables. Mais d'abord, faisons le lien entre la matrice de passage P et les vecteurs propres d'une matrice diagonalisable A.

Soient $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ et $D \in \mathcal{M}_n(\mathbb{C})$ la matrice diagonale telle que $D_{ii} = \lambda_i$. Soit $P \in \mathcal{M}_n(\mathbb{C})$ une matrice inversible et notons $u_1, \dots, u_n \in \mathbb{C}^n$ les n colonnes de la matrice P. Comme P est inversible, la famille $(u_1, \dots, u_n)$ est une base de $\mathbb{C}^n$. Soit $A \in \mathcal{M}_n(\mathbb{C})$ la matrice $A = PDP^{-1}$. On a que

$$\begin{aligned} A = PDP^{-1} &\iff AP = PD \\ &\iff \left[\begin{array}{c} \\ \quad A \quad \\ \\ \end{array}\right] \left[\begin{array}{c|c|c} u_1 & \cdots & u_n \end{array}\right] = \left[\begin{array}{c|c|c} u_1 & \cdots & u_n \end{array}\right] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \\ &\iff \left[\begin{array}{c|c|c} Au_1 & \cdots & Au_n \end{array}\right] = \left[\begin{array}{c|c|c} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{array}\right], \end{aligned}$$

car chaque colonne j de la matrice AP, avec $j \in \{1, \dots, n\}$, est obtenue en multipliant la matrice A par la colonne j de la matrice P, et chaque colonne j de la matrice PD est obtenue en multipliant λ_j par la colonne j de P.

On conclut ainsi que

$$A = PDP^{-1} \text{ ssi } Au_i = \lambda_i u_i, \text{ pour tout } i \in \{1, \dots, n\},$$

où, pour chaque $i \in \{1, \dots, n\}$, $\lambda_i \in \mathbb{C}$ est le ième coefficient diagonal de la matrice D et $u_i \in \mathbb{C}^n$ la ième colonne de P. Autrement dit, $A = PDP^{-1}$ si et seulement si les coefficients de D sont des valeurs propres de la matrice A et les colonnes de P des vecteurs propres de A, associés, respectivement pour chaque $i \in \{1, \dots, n\}$, à la valeur propre $\lambda_i = D_{ii}$. On a ainsi la propriété suivante.

Proposition. *Diagonalisation et base de vecteurs propres.*
Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors A est diagonalisable si et seulement si il existe $(u_1, \dots, u_n)$ base de $\mathbb{C}^n$ formée de vecteurs propres de A. Dans ce cas, si $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de A associées respectivement aux vecteurs propres $u_1, \dots, u_n$, on a $A = PDP^{-1}$, avec $P \in \mathcal{M}_n(\mathbb{C})$ la matrice inversible dont les colonnes sont les vecteurs $u_1, \dots, u_n$, et $D \in \mathcal{M}_n(\mathbb{C})$ la matrice diagonale telle que $D_{ii} = \lambda_i$.

Le résultat ci-dessus montre qu'**être diagonalisable, pour une matrice $A \in \mathcal{M}_n(\mathbb{C})$, équivaut à l'existence d'une base de $\mathbb{C}^n$ constituée de vecteurs propres de** A.

Remarque.
Supposons que $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable et que $A = QDQ^{-1}$ avec D matrice diagonale et Q **unitaire** (c'est-à-dire inversible et vérifiant $Q^{-1} = Q^*$). Alors d'après ce que l'on vient de voir les colonnes de Q sont des vecteurs propres de A qui forment une base de $\mathbb{C}^n$. Comme en plus Q est unitaire cette base de vecteurs propres est **orthonormale** (cf. la section 1.4.3).

Réciproquement, s'il existe une base orthonormale $(u_1, \dots, u_n)$ de $\mathbb{C}^n$ constituée de vecteurs propres de A, alors en définissant Q la matrice de $\mathcal{M}_n(\mathbb{C})$ dont les colonnes sont les vecteurs $u_1, \dots, u_n$, on a d'une part que Q est une matrice unitaire, d'après la section 1.4.3 car la famille $(u_1, \dots, u_n)$ est orthonormale, et d'autre part que $A = QDQ^{-1} = QDQ^*$, avec D la matrice diagonale dont les éléments diagonaux sont les valeurs propres de A associées respectivement aux vecteurs $u_1, \dots, u_n$.

On a ainsi que :

$A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable avec $A = QDQ^*$, $D \in \mathcal{M}_n(\mathbb{C})$ matrice diagonale et $Q \in \mathcal{M}_n(\mathbb{C})$ matrice unitaire (*i. e.* vérifiant $QQ^* = Q^*Q = I_n$), si et seulement si il existe une base orthonormale $(u_1, \dots, u_n)$ de $\mathbb{C}^n$ formée de vecteurs propres de A.

On dit dans ce cas que A **est diagonalisable dans une base orthonormale (de vecteurs propres)**.

Le théorème suivant donne des cas particuliers de matrices diagonalisables.

Théorème. *Théorème spectral - version matricielle.*

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice **symétrique** (c'est-à-dire telle que $A = A^T$). Alors les **valeurs propres de A sont réelles** et A est **diagonalisable dans une base orthonormale** de **vecteurs propres réels**, c'est-à-dire qu'il existe une base orthonormale de $\mathbb{R}^n$ constituée de vecteurs propres de A. De manière équivalente, il existe $D \in \mathcal{M}_n(\mathbb{R})$ matrice diagonale, $Q \in \mathcal{M}_n(\mathbb{R})$ matrice orthogonale (*i. e.* vérifiant $QQ^T = Q^TQ = I_n$), tel que $A = QDQ^T$.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice **hermitienne** (c'est-à-dire telle que $A = A^*$). Alors les **valeurs propres de A sont réelles** et A est **diagonalisable dans une base orthonormale** de **vecteurs propres complexes**, c'est-à-dire qu'il existe une base orthonormale de $\mathbb{C}^n$ constituée de vecteurs propres de A. De manière équivalente, il existe $D \in \mathcal{M}_n(\mathbb{R})$ matrice diagonale, $Q \in \mathcal{M}_n(\mathbb{C})$ matrice unitaire (*i. e.* vérifiant $QQ^* = Q^*Q = I_n$), tel que $A = QDQ^*$.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice **normale** (c'est-à-dire telle que $AA^* = A^*A$). Alors les **valeurs propres de A appartiennent à $\mathbb{C}$** et A est **diagonalisable dans une base orthonormale** de **vecteurs propres complexes**, c'est-à-dire qu'il existe une base orthonormale de $\mathbb{C}^n$ constituée de vecteurs propres de A. De manière équivalente, il existe $D \in \mathcal{M}_n(\mathbb{C})$ matrice diagonale, $Q \in \mathcal{M}_n(\mathbb{C})$ matrice unitaire tel que $A = QDQ^*$.

Démonstration. (*Cas des matrices symétriques réelles*).

Nous allons faire la preuve du théorème dans le cas des matrices symétriques réelles. Le cas des matrices hermitiennes complexes est une simple adaptation de la preuve au cas complexe. Le cas général des matrices normales est aussi une adaptation, mais ce résultat est admis ici. On pourra trouver une preuve par d'autres arguments dans la référence *Algèbre linéaire numérique*, de G. Allaire et S. M. Kaber, Editions Ellipses, 2002.

Nous avons déjà montré que les n valeurs propres d'une matrice symétrique réelle sont réelles, et remarqué qu'une matrice réelle ayant une valeur propre réelle admet obligatoirement un vecteur propre dans $\mathbb{R}^n$ associé à cette valeur propre.

On va montrer à nouveau par induction sur $n \in \mathbb{N}^*$ que pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$ symétrique, il existe $Q \in \mathcal{M}_n(\mathbb{R})$ orthogonale et $D \in \mathcal{M}_n(\mathbb{R})$ diagonale tel que $A = QDQ^T$.

Si $n = 1$ le résultat est vrai : si $A = [a] \in \mathcal{M}_1(\mathbb{R})$, avec $a \in \mathbb{R}$, on a $A = QDQ^T$, avec $Q = [1]$ matrice orthogonale et $D = [a]$ matrice diagonale.

Supposons que le résultat est vrai pour toute matrice symétrique $A \in \mathcal{M}_{n-1}(\mathbb{R})$, avec $n \in \mathbb{N},\ n \geq 2$. Montrons qu'il est vrai pour toute matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$.

Soit $\lambda \in \mathbb{R}$ une valeur propre de A et $u \in \mathbb{R}^n$ un vecteur propre de A associé à la valeur propre λ tel que $\|u\| = 1$. Soit

$$E = \langle u \rangle^\perp := \{v \in \mathbb{R}^n \mid (v|u) = 0\}.$$

E est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n-1$, soient $u_2, \dots, u_n \in \mathbb{R}^n$ tel que $(u_2, \dots, u_n)$ est une base orthonormale de E. Alors la famille $(u, u_2, \dots, u_n)$ est une base de $\mathbb{R}^n$, car il s'agit d'une famille de n vecteurs linéairement indépendants (car orthogonaux deux à deux entre eux). En plus la base $(u, u_2, \dots, u_n)$ est orthonormale ; il en découle que la matrice

$U \in \mathcal{M}_n(\mathbb{R})$ définie par

$$U = \left[\begin{array}{c|c|c|c} u & u_2 & \cdots & u_n \end{array}\right]$$

est orthogonale, c'est-à-dire que $UU^T = U^TU = I_n$.

On a d'autre part que l'espace E est stable par A, c'est-à-dire que si $v \in E$, alors $Av \in E$. En effet, si $v \in E$, on a $(v|u) = 0$. Or la matrice A étant symétrique, on a

$$(Av|u) = (v|Au) = (v|\lambda u) = \lambda(v|u) = 0,$$

et donc $Av \in E$. On a alors que, $u_2, \ldots, u_n$ étant des vecteurs de E, $Au_2, \ldots, Au_n$ appartiennent aussi à E. On a donc, d'une part,

$$Au = \lambda u$$

et, d'autre part, comme $Au_2, \ldots, Au_n \in E = \langle u_2, \ldots, u_n \rangle$, il existe $\mu^{i,j} \in \mathbb{R},\ i,\ j \in \{2, \ldots, n\}$, tel que

$$\begin{aligned} Au_2 &= \mu^{2,2}u_2 + \cdots + \mu^{2,n}u_n \\ &\vdots \\ Au_n &= \mu^{n,2}u_2 + \cdots + \mu^{n,n}u_n, \end{aligned}$$

autrement dit, on a la relation

$$\left[\begin{array}{c} \quad A \quad \end{array}\right] \left[\begin{array}{c|c|c|c} u & u_2 & \cdots & u_n \end{array}\right] = \left[\begin{array}{c|c|c|c} u & u_2 & \cdots & u_n \end{array}\right] \left[\begin{array}{c|ccc} \lambda & & 0 & \\ \hline & \mu^{2,2} & \cdots & \mu^{n,2} \\ 0 & \vdots & & \vdots \\ & \mu^{2,n} & \cdots & \mu^{n,n} \end{array}\right],$$

que l'on peut écrire

$$AU = U \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & C \end{array}\right],$$

où $C \in \mathcal{M}_{n-1}(\mathbb{R})$ est la matrice

$$C = \begin{bmatrix} \mu^{2,2} & \cdots & \mu^{n,2} \\ \vdots & & \vdots \\ \mu^{2,n} & \cdots & \mu^{n,n} \end{bmatrix}.$$

Montrons que C est une matrice symétrique. On note $M \in \mathcal{M}_n(\mathbb{R})$ la matrice

$$M = \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & C \end{array}\right].$$

On a $AU = UM$ et donc, U étant orthogonale, $A = UMU^T$. Comme A est symétrique, on a $A = A^T$ et donc

$$UMU^T = UM^TU^T.$$

La matrice U étant inversible, on déduit que $M = M^T$. Comme

$$M^T = \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & C^T \end{array}\right],$$

on conclut que $C = C^T$.

La matrice C est alors une matrice symétrique réelle de dimension $n-1$. Par hypothèse d'induction, il existe une matrice $\tilde{Q} \in \mathcal{M}_{n-1}(\mathbb{R})$ orthogonale et une matrice $\tilde{D} \in \mathcal{M}_{n-1}(\mathbb{R})$ diagonale tel que $C = \tilde{Q}\tilde{D}\tilde{Q}^T$. On en déduit que

$$A = U \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{Q}\tilde{D}\tilde{Q}^T \end{array}\right] U^T.$$

Mais on remarque que

$$\left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{Q}\tilde{D}\tilde{Q}^T \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right] \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{D} \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q}^T \end{array}\right]$$

et on a donc

$$\begin{aligned} A &= \left(U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right]\right) \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{D} \end{array}\right] \left(\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q}^T \end{array}\right] U^T\right) \\ &= \left(U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right]\right) \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{D} \end{array}\right] \left(\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right]^T U^T\right) \\ &= QDQ^T, \end{aligned}$$

où

$$Q = U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right]$$

et

$$D = \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \tilde{D} \end{array}\right].$$

La matrice D est une matrice de $\mathcal{M}_n(\mathbb{R})$ diagonale car la matrice $\tilde{D}$ est une matrice de $\mathcal{M}_{n-1}(\mathbb{R})$ diagonale. La matrice Q est une matrice de $\mathcal{M}_n(\mathbb{R})$ orthogonale car

$$\begin{aligned} QQ^T &= U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q} \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q}^T \end{array}\right] U^T \\ &= U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{Q}\tilde{Q}^T \end{array}\right] U^T \\ &= U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & I_n \end{array}\right] U^T, && \text{car } \tilde{Q} \text{ est orthogonale,} \\ &= U I_n U^T \\ &= I_n, && \text{car } U \text{ est orthogonale.} \end{aligned}$$

$\square$

1.6.2 Matrices positives et définies positives et valeurs propres.

On va dans cette section montrer qu'une matrice hermitienne ou symétrique (ayant donc des valeurs propres réelles) définie positive a des valeurs propres strictement positives, et qu'une matrice hermitienne ou symétrique positive a des valeurs propres positives. On montrera que la réciproque est aussi vraie : si une matrice hermitienne ou symétrique a des valeurs propres réelles strictement positives (positives), alors elle est définie positive (positive).

Proposition.
Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne, *i. e.* vérifiant $A = A^*$. Alors A est définie positive (respectivement positive) si et seulement si toutes ses valeurs propres sont strictement positives (respectivement positives).

En particulier, une matrice $A \in \mathcal{M}_n(\mathbb{R})$ symétrique est définie positive (respectivement positive) si et seulement si ses valeurs propres sont strictement positives (respectivement positives).

Démonstration. On montre l'équivalence entre matrice définie positive et valeurs propres strictement positives. La preuve dans le cas seulement positif est analogue.

Supposons d'abord $A \in \mathcal{M}_n(\mathbb{C})$ hermitienne définie positive. A étant hermitienne, on a déjà montré que les valeurs propres de A sont réelles, montrons qu'elles sont strictement positives. Soit $\lambda \in \mathbb{R}$ valeur propre de A. Il existe alors $u \in \mathbb{C}^n,\ u \neq 0$, tel que $Au = \lambda u$. On a alors

$$(Au|u)_{\mathbb{C}} = (\lambda u|u)_{\mathbb{C}} = \lambda (u|u)_{\mathbb{C}}.$$

Comme A est définie positive et $u \neq 0$, on a $(Au|u)_{\mathbb{C}} > 0$. Comme $u \neq 0$, on a aussi $(u|u)_{\mathbb{C}} > 0$. On conclut que

$$\lambda = \frac{(Au|u)_{\mathbb{C}}}{(u|u)_{\mathbb{C}}} > 0.$$

Supposons maintenant que A est une matrice hermitienne telle que toutes ses valeurs propres sont strictement positives. Montrons que A est définie positive, en montrant que $(Ax|x)_{\mathbb{C}} > 0$, pour tout $x \in \mathbb{C}^n,\ x \neq 0$.

On note $\lambda_1, \dots, \lambda_n$ les n valeurs propres de A (non nécessairement distinctes), vérifiant $\lambda_i > 0$, pour tout $i \in \{1, \dots, n\}$. Comme la matrice A est hermitienne, il existe une base orthonormale de $\mathbb{C}^n$ constituée de vecteurs propres de A. Soit $(u_1, \dots, u_n)$ une telle base, avec u_i vecteur propre associé à la valeur propre λ_i, pour tout $i \in \{1, \dots, n\}$.

Soit $x \in \mathbb{C}^n,\ x \neq 0$. Comme $(u_1, \ldots, u_n)$ est une base de $\mathbb{C}^n$, il existe $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ tel que $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$. Puisque $x \neq 0$, il existe $i \in \{1, \ldots, n\}$ tel que $\alpha_i \neq 0$. On a alors

$$\begin{aligned}
(Ax|x)_{\mathbb{C}} &= \big(A(\alpha_1 u_1 + \cdots + \alpha_n u_n)|\alpha_1 u_1 + \cdots + \alpha_n u_n\big)_{\mathbb{C}} \\
&= \big(\alpha_1 A u_1 + \cdots + \alpha_n A u_n|\alpha_1 u_1 + \cdots + \alpha_n u_n\big)_{\mathbb{C}} \\
&= \big(\alpha_1 \lambda_1 u_1 + \cdots + \alpha_n \lambda_n u_n|\alpha_1 u_1 + \cdots + \alpha_n u_n\big)_{\mathbb{C}} \\
&= \Big(\sum_{i=1}^{n} \alpha_i \lambda_i u_i \Big| \sum_{j=1}^{n} \alpha_j u_j\Big)_{\mathbb{C}} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \lambda_i \overline{\alpha_j} (u_i|u_j)_{\mathbb{C}} \\
&= \sum_{i=1}^{n} \alpha_i \lambda_i \overline{\alpha_i},
\end{aligned}$$

car les vecteurs $u_1, \ldots, u_n$ sont de norme égale à un et deux à deux orthogonaux. On conclut alors que

$$(Ax|x)_{\mathbb{C}} = |\alpha_1|^2 \lambda_1 + \cdots + |\alpha_n|^2 \lambda_n \geq 0.$$

Comme il existe $i \in \{1, \ldots, n\}$ tel que $\alpha_i > 0$ et donc tel que $|\alpha_i| > 0$ et comme pour tout $i \in \{1, \ldots, n\}$, $\lambda_i > 0$, on en déduit que $(Ax|x)_{\mathbb{C}} > 0$. $\square$