Notes de cours de l'ISIMA, première année, http://www.isima.fr/~leborgne

Diagonalisation : valeurs propres, valeurs propres généralisées

Gilles Leborgne

14 septembre 2021

Table des matières

1 Rappels matriciels

1.1 Rappels vectoriels

1.1.1 Base

Pour la diagonalisation (recherche de valeurs propres) on utilisera $\mathbb{R}$ (l'ensemble des réels) ou $\mathbb{C}$ (l'ensemble des complexes)). On note notera K pour désigner $\mathbb{R}$ ou $\mathbb{C}$.

Soit V un espace vectoriel de dimension n sur le corps K.

Une base $(\vec{e}_i)_{i=1,\dots,n}$ dans V est une famille de vecteurs de V qui est libre et génératrice. On notera simplement $(\vec{e}_i)_{i=1,\dots,n} \overset{\text{noté}}{=} (\vec{e}_i)$ si la dimension n est implicite.

Si $(\vec{e}_i)_{i=1,\dots,n}$ est une base de V, si $\vec{x} \in V$, alors $\vec{x}$ admet n uniques composantes $x_i \in K$:

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i. \tag{1.1}$$

Définition 1.1 La représentation en “matrice colonne” du vecteur $\vec{x}$ dans la base $(\vec{e}_i)$ est :

$$[\vec{x}]_{|\vec{e}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{1.2}$$

et cette matrice colonne $[\vec{x}]_{|(\vec{e})}$ est également appelée un “vecteur colonne” (vocabulaire dangereux : une matrice n'est pas un vecteur).

En particulier :

$$[\vec{e}_i]_{|\vec{e}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \ \leftarrow \ \text{ligne } i \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{1.3}$$

Remarque 1.2 Le vocabulaire “vecteur colonne” est trompeur : un “vecteur colonne” n'est pas un vecteur, c'est une matrice. Donc on utilise “vecteur colonne” pour dire “matrice colonne” dans une base. ▪

1.1.2 Base canonique dans K^n

Soit $K = \mathbb{R}$ ou $\mathbb{C}$. Un élément de K est appelé un scalaire.

$K^n = K \times ... \times K$ (n fois) est le produit cartésien de K par lui-même n fois. (Donc $K^n = \mathbb{R}^n$ ou bien $K^n = \mathbb{C}^n$.) On le munit de l'addition usuelle et de la multiplication par un scalaire usuelle qui en fait un espace vectoriel sur le corps K (immédiat).

Définition 1.3 La base canonique de K^n, notée ici $(\vec{E}_i)_{i=1,...,n}$, est définie par :

$$\vec{E}_1 = (1, 0, ..., 0) \in K \times ... \times K, \qquad ..., \qquad \vec{E}_n = (0, ..., 0, 1) \in K \times ... \times K.$$

où on a utilisé 1 fois l'élément neutre 1 de la multiplication, et $n-1$ fois l'élément neutre 0 de l'addition.

Lorsqu'on utilise la base canonique, on abrège la notation (1.2) : pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i \in K^n$:

$$[\vec{x}]_{|\vec{E}} \overset{\text{noté}}{=} [\vec{x}] \qquad (= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}). \tag{1.4}$$

Ainsi $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i = (x_1, ..., x_n) \in K \times ... \times K$ est représenté matriciellement par (1.4) dans la base canonique. En particulier :

$$[\vec{E}_i] = \begin{pmatrix} \vdots \\ 0 \\ 1 \leftarrow \text{ ligne } i \\ 0 \\ \vdots \end{pmatrix}. \tag{1.5}$$

Lorsqu'on travaillera simultanément avec K^m et K^n, on notera $(\vec{E}_i^{(m)})_{i=1,...,m} = (\vec{E}_i^{(m)})$ et $(\vec{E}_i^{(n)})_{i=1,...,n} = (\vec{E}_i^{(n)})$ les bases canoniques respectives de K^m et de K^n.

Remarque 1.4 Un espace vectoriel n'a pas nécessairement de base canonique : il n'est pas nécessairement décomposable de manière “naturelle” en un produit cartésien d'un corps. Exemple : l'espace vectoriel P_1 des fonctions affines par morceaux, cf. § 5.7, utilisé en éléments finis, en imagerie, en dessin informatique... ▪

Remarque 1.5 « La » base canonique est une notion mathématique. En pratique elle dépend :

a) du choix du 1 (!) : 1 pied ? 1 mètre ? 1 nanomètre ? 1 Ampère ? 1 degré Kelvin ?...

b) et de l'orientation choisie : dans $\mathbb{R}^3$ la “verticale” dépend du lieu où on est...

D'où l'importance des formules de changement de bases (pour que les différents utilisateurs puissent échanger leurs résultats et les comparer). ▪

1.1.3 Produit scalaire usuel dans $\mathbb{R}^n$

On rappelle que les symboles de Kronecker δ_{ij} sont définis par :

$$\delta_{ij} = \begin{cases} 1 \text{ si } i = j, \\ 0 \text{ si } i \neq j. \end{cases} \tag{1.6}$$

Définition 1.6 Un produit scalaire est une forme bilinéaire symétrique définie positive (sert à comparer deux vecteurs). Le produit scalaire usuel dans $\mathbb{R}^n$ est la forme bilinéaire $(\cdot, \cdot)_{\mathbb{R}^n}$ définie à l'aide de la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$ par :

$$\forall i, j = 1, ..., n, \qquad (\vec{E}_i, \vec{E}_j)_{\mathbb{R}^n} = \delta_{ij}. \tag{1.7}$$

Par bilinéarité, pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i$ et $\vec{y} = \sum_{i=1}^n y_i \vec{E}_i$, on déduit immédiatement : $(\vec{x}, \vec{y})_{\mathbb{R}^n} = \sum_{i,j=1}^n x_i y_j (\vec{E}_i, \vec{E}_j)_{\mathbb{R}^n} = \sum_{i,j=1}^n x_i y_j \delta_{ij}$, donc :

$$(\vec{x}, \vec{y})_{\mathbb{R}^n} = \sum_{i=1}^n x_i y_i. \tag{1.8}$$

Définition 1.7 Une base $(\vec{e}_i)$ de $\mathbb{R}^n$ est une base orthonormée (b.o.n.) dans $(\mathbb{R}^n, (\cdot, \cdot)_{\mathbb{R}^n})$ (orthonormée relativement au produit scalaire usuel), ssi :

$$\forall i, j = 1, ..., n, \qquad (\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = \delta_{ij}. \tag{1.9}$$

Exemple : la base canonique de $\mathbb{R}^n$ est une b.o.n. dans $(\mathbb{R}^n, (\cdot, \cdot)_{\mathbb{R}^n})$, par définition du produit scalaire usuel, cf. (1.7).

1.1.4 Norme usuelle dans $\mathbb{R}^n$ et Pythagore

Définition 1.8 La norme usuelle dans $\mathbb{R}^n$ est la norme associée au produit scalaire usuel : pour tout $\vec{x} \in \mathbb{R}^n$:

$$||\vec{x}||_{\mathbb{R}^n} = \sqrt{(\vec{x}, \vec{x})_{\mathbb{R}^n}}. \tag{1.10}$$

Théorème 1.9 (Pythagore). *Le théorème de Pythagore exprime la norme usuelle dans la base canonique $(\vec{E}_i)$: pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i \in \mathbb{R}^n$ (décomposition sur la base canonique) on a :*

$$||\vec{x}||_{\mathbb{R}^n}^2 = \sum_{i=1}^n x_i^2 \qquad (Pythagore). \tag{1.11}$$

Et cela reste vrai si on remplace la base canonique $(\vec{E}_i)$ par une b.o.n. $(\vec{e}_i)$ de $(\mathbb{R}^n, (\cdot, \cdot)_{\mathbb{R}^n})$.

Preuve. $||\vec{x}||_{\mathbb{R}^n}^2 = (\vec{x}, \vec{x})_{\mathbb{R}^n} = \sum_{i,j=1}^n x_i x_j (\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = \sum_{i,j=1}^n x_i x_j \delta_{ij} = \sum_{i=1}^n x_i^2$ pour $(\vec{e}_i)$ b.o.n.. ▪

1.1.5 Produit scalaire usuel dans $\mathbb{C}^n$

On aura essentiellement besoin des complexes et du produit scalaire usuel dans $\mathbb{C}^n$ pour démontrer le théorème de diagonalisation des matrices symétriques réelles (existence des racines du polynôme caractéristique).

Soit $(\vec{E}_i)$ la base canonique de $\mathbb{C}^n = \mathbb{C} \times ... \times \mathbb{C}$ (espace vectoriel sur $\mathbb{C}$).

Pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i \in \mathbb{C}^n$, on définit le vecteur complexe conjugué par $\overline{\vec{x}} = \sum_{i=1}^n \overline{x}_i \vec{E}_i$.

Définition 1.10 Le produit scalaire usuel dans $\mathbb{C}^n$ est l'application $(\cdot, \cdot)_{\mathbb{C}^n} : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ définie par, pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i$ et $\vec{y} = \sum_{i=1}^n y_i \vec{E}_i$. :

$$(\vec{x}, \vec{y})_{\mathbb{C}^n} = \sum_{i=1}^n x_i \overline{y_i}, \tag{1.12}$$

où $\overline{y_i}$ est le complexe conjugué du complexe y_i.

N.B. : certains auteurs notent $(\vec{x}, \vec{y})_{\mathbb{C}^n} = \sum_{i=1}^n \overline{x_i}\, y_i$ (conjugué de (1.12)).

Donc :

$$(\vec{y}, \vec{x})_{\mathbb{C}^n} = \overline{(\vec{x}, \vec{y})_{\mathbb{C}^n}} \qquad (= \sum_{i=1}^n y_i \overline{x_i}). \tag{1.13}$$

1.1.6 Norme usuelle dans $\mathbb{C}^n$

Définition 1.11 La norme usuelle $||.||_{\mathbb{C}^n} : \mathbb{C}^n \to \mathbb{R}_+$ est la norme associée au produit scalaire usuel $(\cdot, \cdot)_{\mathbb{C}^n}$: pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i \in \mathbb{C}^n$, c'est le réel donnée par :

$$||\vec{x}||_{\mathbb{C}^n} = \sqrt{(\vec{x}, \vec{x})_{\mathbb{C}^n}} = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}, \tag{1.14}$$

où $|z| = \sqrt{z\overline{z}}$ est le module du complexe z.

1.2 Matrices $m*n$

1.2.1 L'espace vectoriel des matrices $m*n$

Définition 1.12 Soit $m,n \in \mathbb{N}^*$. Une matrice $m*n$ dans K est un tableau :

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix} = [A_{ij}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \tag{1.15}$$

où $A_{ij} \in K$ pour tout i,j. Le premier indice dans A_{ij} désigne la ligne, ici i désigne la "ligne i", et le second indice désigne la colonne, ici j désigne "la colonne j", le scalaire A_{ij} se trouvant à l'intersection de la ligne i et de la colonne j.

On note simplement $A \overset{\text{noté}}{=} [A_{ij}]$ si m et n sont implicites.

On note $\mathcal{M}_{m,n}(K) = K^{m\times n}$ l'ensemble de matrices $m*n$ dans K.

Définition 1.13 Une matrice ligne est une matrice $A = \begin{pmatrix} A_{11} & \dots & A_{1n} \end{pmatrix}$, ici matrice ligne $1*n$.
Une telle matrice est également appelé un "vecteur ligne". Attention : un "vecteur ligne" n'est pas un vecteur : c'est une matrice.

Définition 1.14 Une matrice colonne est une matrice $A = \begin{pmatrix} A_{11} \\ \vdots \\ A_{n1} \end{pmatrix}$, ici matrice colonne $n*1$.

Une telle matrice est également appelé un "vecteur colonne". Attention : un "vecteur colonne" n'est pas un vecteur : c'est une matrice.

Définition 1.15 Une matrice est dite carrée ssi $m=n$; sinon elle est dite rectangle.

On note aussi $\mathcal{M}_{n,n}(K) = K^{n^2}$ l'ensemble des matrices carrées dans K.

Définition 1.16 Une matrice $D = [D_{ij}]$ est diagonale ssi elle est carrée et $D_{ij} = 0$ pour tout $i \neq j$. Et on note alors :

$$D = \text{diag}(\lambda_1, ..., \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & & ... & 0 \\ 0 & \lambda_2 & 0 & ... & \\ \vdots & 0 & \ddots & & \\ & & & & 0 \\ 0 & ... & & 0 & \lambda_n \end{pmatrix} = \text{diag}(\lambda_i), \tag{1.16}$$

où donc $\lambda_i = D_{ii}$ pour tout i, et $D_{ij} = 0$ si $i \neq j$.

Définition 1.17 La matrice $n*n$ identité est la matrice carrée définie par :

$$I = [\delta_{ij}]_{\substack{i=1,\dots,n \\ j=1,\dots,n}} = \text{diag}(1, ..., 1), \tag{1.17}$$

matrice diagonale dont les éléments diagonaux valent tous 1.

Définition 1.18 Une matrice $L = [L_{ij}] \in \mathcal{M}_{nn}$ est triangulaire inférieure (Lower triangular matrix) ssi $i < j$ implique $L_{ij} = 0$ (des zéros au dessus de la diagonale).
Une matrice $U = [U_{ij}] \in \mathcal{M}_{nn}$ est triangulaire supérieure (Upper triangular matrix) ssi $i > j$ implique $U_{ij} = 0$ (des zéros en dessous de la diagonale).

Définition 1.19 Soit $A = [A_{ij}]$ et $B = [B_{ij}]$ deux matrices $m*n$ dans K, et soit $\lambda \in K$. On définit la somme $A+B = [(A+B)_{ij}]$ et le produit $\lambda A = [(\lambda A)_{ij}]$ par, pour tout i,j :

$$(A+B)_{ij} = A_{ij} + B_{ij} \qquad \text{et} \qquad (\lambda A)_{ij} = \lambda A_{ij}. \tag{1.18}$$

Proposition 1.20 *L'ensemble $\mathcal{M}_{m,n}(K)$ des matrices $m*n$ dans K, muni de l'addition des matrices et de la multiplication par un scalaire, est un espace vectoriel.*

Preuve. Il suffit d'appliquer la définition d'un espace vectoriel : vérifications élémentaires. ▪

1.2.2 Matrice transposée

Définition 1.21 Soit $A=[A_{ij}]_{\substack{i=1,...,m\\ j=1,...,n}} \in \mathcal{M}_{mn}(K)$. Sa matrice transposée est la matrice $A^T = [(A^T)_{ij}]_{\substack{i=1,...,n\\ j=1,...,m}} \in \mathcal{M}_{nm}(K)$ définie par :

$$\forall i=1,...,n, \quad \forall j=1,...,m, \quad (A^T)_{ij}=A_{ji}, \tag{1.19}$$

autrement dit, la transposition échange les lignes et les colonnes.

Exemple 1.22 $A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}$ a pour matrice transposée $A^T=\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}$. ∎

Proposition 1.23 *Pour toute matrice A on a $(A^T)^T=A$.*

Preuve. $((A^T)^T)_{ij}=(A^T)_{ji}=A_{ij}$, pour tout i,j. ∎

1.2.3 Produit d'une matrice $1*n$ et d'une matrice $n*1$

Définition 1.24 Soit $[\ell]=(\,\ell_1 \quad ... \quad \ell_n\,)$ une matrice ligne $1*n$, et soit $[c]=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$ une matrice colonne $n*1$. On définit le produit matriciel $[\ell].[c]$ comme donnant le réel définie par :

$$[\ell].[c]=\sum_{i=1}^n \ell_i c_i \quad \in \mathbb{R}. \tag{1.20}$$

Exemple : si $(\vec{v}_i)$ une base de $\mathbb{R}^n$, si $\vec{x}=\sum_{i=1}^n x_i\vec{v}_i$ et si $\vec{y}=\sum_{i=1}^n y_i\vec{v}_i$, alors avec la notation (1.2) :

$$[\vec{x}]_{|(\vec{v})}^T.[\vec{y}]_{|(\vec{v})}=\sum_{i=1}^n x_iy_i. \tag{1.21}$$

Proposition 1.25 *Soit $(\vec{v}_i)$ une b.o.n. de $(\mathbb{R}^n,(\cdot,\cdot)_{\mathbb{R}^n})$. Si $\vec{x}=\sum_{i=1}^n x_i\vec{v}_i$ et $\vec{y}=\sum_{i=1}^n y_i\vec{v}_i$, alors :*

$$(\vec{x},\vec{y})_{\mathbb{R}^n}=[\vec{x}]_{|(\vec{v})}^T.[\vec{y}]_{|(\vec{v})}, \qquad \textit{si } (\vec{v}_i) \textit{ b.o.n..} \tag{1.22}$$

C'est faux si $(\vec{v}_i)$ n'est pas une b.o.n..

Preuve. Soit $(\vec{v}_i)$ une b.o.n. : $(\vec{x},\vec{y})_{\mathbb{R}^n}=\sum_{i,j=1}^n x_iy_j(\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n}=\sum_{i,j=1}^n x_iy_j\delta_{ij}=\sum_{i=1}^n x_iy_i=[\vec{x}]_{|(\vec{v})}^T.[\vec{y}]_{|(\vec{v})}$, cf. (1.2) et (1.20).

Si $(\vec{v}_i)$ n'est pas une b.o.n., il existe deux vecteurs $\vec{v}_i$ et $\vec{v}_j$ t.q. $(\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n}\neq\vec{0}$. Donc, avec (1.3) et (1.20), $[\vec{v}_i]_{|\vec{v}}^T.[\vec{v}_j]_{|\vec{v}}=0\neq(\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n}$. ∎

Cas particulier de $(\vec{E}_i)$ la base canonique : (1.22) est écrite simplement :

$$(\vec{x},\vec{y})_{\mathbb{R}^n}=[\vec{x}]^T.[\vec{y}], \qquad \text{si base canonique et produit scalaire usuel.} \tag{1.23}$$

1.2.4 Produit d'une matrice $m*n$ et d'une matrice $n*1$

Soit $A=[A_{ij}]_{\substack{i=1,...,m\\ j=1,...,n}} \in \mathcal{M}_{mn}$ une matrice $m*n$. Notons :

$$A=\begin{pmatrix}[\ell_1]\\ \vdots\\ [\ell_m]\end{pmatrix}, \qquad \text{où donc} \qquad [\ell_i]=(\,A_{i1} \quad ... \quad A_{in}\,), \tag{1.24}$$

où donc $[\ell_i]=(\,A_{i1} \quad ... \quad A_{in}\,)$ est la i-ème matrice ligne de A (le i-ème "vecteur ligne"). Soit $[c]=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$ une matrice colonne $n*1$.

Définition 1.26 On définit le produit matriciel $A.[c]$ comme donnant la matrice colonne $m*1$ définie par :

$$A.[c] \stackrel{\text{def}}{=} \begin{pmatrix} [\ell_1].[c] \\ \vdots \\ [\ell_m].[c] \end{pmatrix} \quad \in \mathcal{M}_{m1}. \tag{1.25}$$

Donc :

$$A.[c] = \begin{pmatrix} A_{11} & ... & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & ... & A_{mn} \end{pmatrix} . \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n A_{1j}c_j \\ \vdots \\ \sum_{j=1}^n A_{mj}c_j \end{pmatrix}. \tag{1.26}$$

En particulier, pour $\vec{x} = \sum_{i=1}^n x_i \vec{E}_i^{(n)}$ et $[\vec{x}] = [\vec{x}]_{|(\vec{E})}$ (utilisation de la base canonique) :

$$A.[\vec{x}] = \begin{pmatrix} \sum_{j=1}^n A_{1j}x_j \\ \vdots \\ \sum_{j=1}^n A_{mj}x_j \end{pmatrix}. \tag{1.27}$$

En particulier :

$$A.[\vec{E}_j^{(n)}] = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{pmatrix} = \text{colonne } j \text{ de } A. \tag{1.28}$$

1.2.5 Définition du vecteur $A.\vec{x}$

Multiplier une matrice et un vecteur n'a pas de sens : l'existence d'un vecteur ne dépend du choix d'une base faite par un expérimentateur (un vecteur est "intrinsèque"), alors qu'une matrice est juste un tableau de scalaires.

Quand on utilise les bases canoniques, on abuse tout de même des notations (c'est pratique au sens où ça allège les notations) : on représente un vecteur $\vec{x}$ dans la base canonique par sa matrice colonne $[\vec{x}]$ pour définir $A.\vec{x}$:

Définition 1.27 Soit $A = [A_{ij}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$. Quand on utilise les bases canoniques :

$$\text{si} \quad \vec{x} = \sum_{i=1}^n x_i \vec{E}_i^{(n)} \quad \text{alors (définition) :} \qquad A.\vec{x} \stackrel{\text{def}}{=} \sum_{i=1}^m (\sum_{j=1}^n A_{ij}x_j) \vec{E}_i^{(m)}. \tag{1.29}$$

Donc :

$$[A.\vec{x}] = \begin{pmatrix} \sum_{j=1}^n A_{1j}x_j \\ \vdots \\ \sum_{j=1}^n A_{mj}x_j \end{pmatrix} = A.[\vec{x}], \qquad \text{quand on utilise les bases canoniques,} \tag{1.30}$$

i.e. les composantes du vecteur $A.\vec{x}$ dans la base canonique sont stockées dans la matrice colonne $A.[\vec{x}]$ (cf. (1.27)). Notation abusive :

$$A.[\vec{x}] \stackrel{\text{noté}}{=} A.\vec{x} \qquad \text{quand on utilise les bases canoniques,} \tag{1.31}$$

et donc un vecteur est "identifié" à une matrice (!) : $[A.\vec{x}] = A.\vec{x}$. Attention uniquement si on utilise les bases canoniques (notation dangereuse si on fait des changements de base, et notation inadmissible si deux utilisateurs calculent indépendamment, cf. remarque 1.5).

Remarque 1.28 Définition non abusive : soit $\mathcal{L} : \vec{x} \in \mathbb{R}^n \to \mathcal{L}.\vec{x} \in \mathbb{R}^m$ une application linéaire. Une application linéaire est parfaitement déterminée si, étant donnée une base $(\vec{v}_i)$ dans $\mathbb{R}^n$ et une base $(\vec{w}_i)$ dans $\mathbb{R}^m$, on connaît les images des $\mathcal{L}.\vec{v}_i$ dans la base $(\vec{w}_i)$, i.e. si on connaît les scalaires A_{ij} t.q. $\mathcal{L}.\vec{v}_j = \sum_{i=1}^m A_{ij}\vec{w}_i$, et alors $[\mathcal{L}]_{|(\vec{v}),(\vec{w})} = [A_{ij}] = A$ est la matrice de $\mathcal{L}$ relativement à ces bases. Dans le cas particulier où les bases sont les bases canoniques, on note $A.\vec{x} \stackrel{\text{def}}{=} \mathcal{L}.\vec{x}$. Voir le § 2.2. ▪

Exemple 1.29 $A.\vec{E}_j^{(n)}$ est donc le vecteur dont les composantes dans la base canonique sont données par la colonne j de A. C'est le j-ème "vecteur colonne" de A où on utilise ici les bases canoniques. ▪

La définition (1.31) donne, pour tout $\vec{x},\vec{y}\in K^n$ et tout $\lambda\in\mathbb{R}$:

$$A.(\vec{x}+\lambda\vec{y}) = A.\vec{x}+\lambda A.\vec{y}, \tag{1.32}$$

vérification immédiate avec (1.27). D'où :

Définition 1.30 On dit que A est linéaire.

Proposition 1.31 *Soit $A\in\mathcal{M}_{nn}(K)$. S'il existe une base $(\vec{v}_i)$ dans K^n telle que $A.\vec{v}_j=\vec{v}_j$ pour tout $j=1,...,n$ (donc au sens $A.[\vec{v}_j]=[\vec{v}_j]$ pour tout $j=1,...,n$, cf. (1.31)), alors $A=I$ (la matrice identité dans K^n).*

Preuve. Soit $\vec{x}\in\mathbb{R}^n$. Soit x_i ses composantes sur la base $(\vec{v}_i)$: $\vec{x}=\sum_{i=1}^n x_i\vec{v}_i$ Donc $A.\vec{x}=\sum_{j=1}^n x_jA.\vec{v}_j=\sum_{j=1}^n x_j\vec{v}_j=\vec{x}$ pour tout $\vec{x}\in\mathbb{R}^n$. En particulier $A.\vec{E}_j=\vec{E}_j$ pour tout j. Donc le j-ème vecteur colonne de A est $A.[\vec{E}_j]=[\vec{E}_j]$, cf. (1.5). Donc $A=I$. ▪

1.2.6 Produit d'une matrice $m*n$ et d'une matrice $n*p$

Soit $A=[A_{ij}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}}=\begin{pmatrix}[\ell_1]\\ \vdots\\ [\ell_m]\end{pmatrix}\in\mathcal{M}_{mn}$ (matrice $m*n$), cf. (1.24). Soit :

$$B=[B_{ij}]_{\substack{i=1,\dots,n\\ j=1,\dots,p}}=\begin{pmatrix}[\vec{b}_1] & \dots & [\vec{b}_p]\end{pmatrix}\in\mathcal{M}_{np},\qquad \text{où donc}\qquad [\vec{b}_j]=\begin{pmatrix}B_{1j}\\ \vdots\\ B_{nj}\end{pmatrix}, \tag{1.33}$$

donc $[\vec{b}_j]\in\mathcal{M}_{np}$ est le "j-ème vecteur colonne" de B pour j=1,...,p.

Définition 1.32 Le produit matriciel $A.B$ est la matrice $m*p$ définie par :

$$A.B\overset{\text{def}}{=}\begin{pmatrix}[\ell_1].[\vec{b}_1] & \dots & [\ell_1].[\vec{b}_p]\\ \vdots & & \vdots\\ [\ell_m].[\vec{b}_1] & \dots & [\ell_m].[\vec{b}_p]\end{pmatrix}=\left[[\ell_i].[\vec{b}_j]\right]_{\substack{i=1,\dots,m\\ j=1,\dots,p}}, \tag{1.34}$$

matrice des produits "lignes×colonnes".

Donc :

$$A.B=\begin{pmatrix}\sum_{j=1}^n A_{1j}B_{j1} & \dots & \sum_{j=1}^n A_{1j}B_{jp}\\ \vdots & & \vdots\\ \sum_{j=1}^n A_{mj}B_{j1} & \dots & \sum_{j=1}^n A_{mj}B_{jp}\end{pmatrix}=[\sum_k A_{ik}B_{kj}]_{\substack{i=1,\dots,m\\ j=1,\dots,p}}, \tag{1.35}$$

et :

$$A.B\overset{\text{def}}{=}\left(\begin{pmatrix}A.[\vec{b}_1]\end{pmatrix}\quad\dots\quad\begin{pmatrix}A.[\vec{b}_p]\end{pmatrix}\right), \tag{1.36}$$

i.e. la matrice $A.B$ stocke le "vecteur colonne" $A.[\vec{b}_j]$ dans sa j-ème colonne, cf. (1.27).

Exemple 1.33 Si une machine fabrique p produits $[\vec{b}_j]$ caractérisés par n valeurs $(B_{ij})_{i=1,\dots,n}$, si A est un indicateur de "qualité", la matrice $A.B$ stocke dans ses colonnes les "qualités $[A.\vec{b}_j]$" des produits $[\vec{b}_j]$ fabriqués. ▪

Proposition 1.34 *Pour $A\in\mathcal{M}_{m,n}(K)$ et $B\in\mathcal{M}_{n,p}(K)$ on a :*

$$(A.B)^T=B^T.A^T. \tag{1.37}$$

Preuve. $((A.B)^T)_{ij}=(A.B)_{ji}=\sum_k A_{jk}B_{ki}$ et $(B^T.A^T)_{ij}=\sum_k(B^T)_{ik}(A^T)_{kj}=\sum_k B_{ki}A_{jk}$, d'où $((A.B)^T)_{ij}=(B^T.A^T)_{ij}$ pour tout i,j. ▪

1.2.7 Relation $(A.\vec{x},\vec{y})_{\mathbb{R}^n} = [\vec{y}]^T.A.[\vec{x}]$ (utilisation des bases canoniques)

Soit $(\vec{E}_i^{(k)})$ la base canonique de $\mathbb{R}^k$. Pour $\vec{x} = \sum_{i=1}^k x_i \vec{E}_i^{(k)}$ on note $[\vec{x}] = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$. Et soit $(\cdot,\cdot)_{\mathbb{R}^k}$ le produit scalaire usuel dans $\mathbb{R}^k$.

Proposition 1.35 *Si $A \in \mathcal{M}_{mn}(\mathbb{R})$, si $\vec{x} \in \mathbb{R}^n$, si $\vec{y} \in \mathbb{R}^m$, alors avec la notation (1.31) :*

$$(A.\vec{x},\vec{y})_{\mathbb{R}^m} = [\vec{y}]^T.A.[\vec{x}] \qquad (= [\vec{x}]^T.A^T.[\vec{y}] = (\vec{x},A.\vec{y})_{\mathbb{R}^n}), \tag{1.38}$$

où donc on a utilisé les représentations des vecteurs dans les bases canoniques. En particulier :

$$A_{ij} = (A.\vec{E}_j^{(n)},\vec{E}_i^{(m)})_{\mathbb{R}^m} \qquad (= [\vec{E}_i^{(m)}]^T.A.[\vec{E}_j^{(n)}]). \tag{1.39}$$

Preuve. Avec (1.31), où donc on utilise les bases canoniques, avec $\vec{x} = \sum_{i=1}^m x_i \vec{E}_i^{(m)}$ et $\vec{y} = \sum_{i=1}^n y_i \vec{E}_i^{(n)}$, on a $A.\vec{y} = \sum_{i=1}^m (\sum_{j=1}^n A_{ij}y_j)\vec{E}_i^{(m)}$, donc $(\vec{x},A.\vec{y})_{\mathbb{R}^m} = \sum_{i=1}^m \sum_{j=1}^n x_i A_{ij} y_j$ qui est bien égal à $[\vec{x}]^T.A.[\vec{y}]$. ▪

1.2.8 Associativité du produit matriciel et algèbre des matrices

Proposition 1.36 *Si $A \in \mathcal{M}_{m,n}(K)$, $B \in \mathcal{M}_{n,p}(K)$ et $C \in \mathcal{M}_{p,q}(K)$ sont trois matrices, alors :*

$$(A.B).C = A.(B.C). \tag{1.40}$$

D'où $\mathcal{M}_{n,n}(K)$ muni de l'addition des matrices et de la multiplication matricielle est un anneau. Cet anneau est non commutatif (pour $n \geq 2$).

Et $\mathcal{M}_{n,n}(K)$ muni de l'addition des matrices, de la multiplication matricielle, et de la multiplication par un scalaire est une algèbre non commutative.

Preuve. Vérification de l'associativité (1.40) : exercice facile. L'élément neutre de la multiplication est la matrice identité (immédiat). La distributivité est immédiate : $(\mathcal{M}_{n,n}(K),+,.)$ est un anneau.

Mais cet anneau n'est pas commutatif : prendre $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ qui donnent $A.B = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ et $B.A = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \neq A.B$

Et $\mathcal{M}_{n,n}(K)$ est une algèbre : exercice facile. ▪

1.2.9 Échange de lignes et de colonnes

Soit $(\vec{E}_i)$ la base canonique de K^n.

Définition 1.37 La matrice $n * n$ de transposition pour les indices i et j est la matrice P définie par (ses colonnes) :

$$P.[\vec{E}_i] = [\vec{E}_j], \qquad P.[\vec{E}_j] = [\vec{E}_i], \qquad \text{et} \qquad P.[\vec{E}_k] = [\vec{E}_k] \qquad \forall k \neq i,j. \tag{1.41}$$

Exemple 1.38 $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est la matrice de transposition pour les deux premiers indices. ▪

Proposition 1.39 *Une matrice P de transposition est symétrique d'inverse elle-même : $P^T = P$ et $P.P = I$.*

Preuve. Soit $i < j$ et $P = \big([\vec{E}_1] \ \ ... \ \ [\vec{E}_{i-1}] \ \ [\vec{E}_j] \ \ [\vec{E}_{i+1}] \ \ ... \ \ [\vec{E}_{j-1}] \ \ [\vec{E}_i] \ \ [\vec{E}_{j+1}] \ \ ...[\vec{E}_n]\big)$, matrice de transposition des indices i et j. Avec (1.41) on a : $P_{kk} = 1$ pour tout $k \neq i,j$, $P_{ij} = P_{ji} = 1$, et les autres termes nuls. Donc $P_{ij} = P_{ji}$ pour tout i,j.

Et $P.P.[\vec{E}_j] = P.[\vec{E}_i] = [\vec{E}_j]$ pour tout j, d'où $P.P = I$, cf. proposition 1.31. ▪

Proposition 1.40 *Soit P une matrice $n*n$ de transposition pour les indices i et j. Soit A une matrice $n*n$. Alors :*
$A.P$ est la matrice qui dans A a échangé les colonnes i et j.
$P.A$ est la matrice qui dans A a échangé les lignes i et j.

Preuve. Soit $A = (\,[\vec{a}_1] \quad ... \quad [\vec{a}_n]\,)$. Avec (1.36) et (1.41) on a :

$$\begin{aligned} A.P &= \big(\,A.[\vec{E}_1] \quad ... \quad A.[\vec{E}_{i-1}] \quad A.[\vec{E}_j] \quad A.[\vec{E}_{i+1}] \quad ... \quad A.[\vec{E}_{j-1}] \quad A.[\vec{E}_i] \quad A.[\vec{E}_{j+1}] \quad ... \quad A.[\vec{E}_n]\,\big) \\ &= (\,[\vec{a}_1] \quad ... \quad [\vec{a}_{i-1}] \quad [\vec{a}_j] \quad [\vec{a}_{i+1}] \quad ... \quad [\vec{a}_{j-1}] \quad [\vec{a}_i] \quad [\vec{a}_{j+1}] \quad ... \quad [\vec{a}_n]\,), \end{aligned}$$

et on a bien échangé les colonnes i et j. Donc $A^T.P$ échange les colonnes i et j de A^T. Donc $(A^T.P)^T = P.A$ échange les lignes i et j de A. ▪

Définition 1.41 Une matrice $n*n$ de permutation est une matrice $P = \big(\,[\vec{E}_{\sigma(1)}] \quad ... \quad [\vec{E}_{\sigma(n)}]\,\big)$ où σ est une permutation sur $[1,n]_{\mathbb{N}}$ (bijection de $\{1,...,n\}$ dans lui-même).

Proposition 1.42 *Toute matrice de permutation est une composée de matrice de transposition.*

Preuve. Il s'agit de démontrer que toute permutation sur $[1,n]_{\mathbb{N}}$ est une composée de transpositions. Cas $n=2$: une permutation est une transposition. Puis récurrence : Considérons le cas $n{+}1$: on commence par la transposition qui envoie $\sigma(n{+}1)$ sur $n{+}1$: on s'est ramené au cas n. ▪

1.3 Matrice symétrique, adjointe, hermitienne

1.3.1 Matrice symétrique

Définition 1.43 Soit A une matrice carrée $n*n$. La matrice $A = [A_{ij}]$ est dite symétrique ssi :

$$A^T = A, \tag{1.42}$$

(si A n'est pas carrée (1.42) n'a pas de sens), i.e. ssi :

$$\forall i,j = 1,...,n, \quad (A^T)_{ij} = A_{ij}. \tag{1.43}$$

Exemple 1.44 $A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$ est symétrique. ▪

Exemple 1.45 Toute matrice diagonale est symétrique : immédiat. ▪

Proposition 1.46 *Si $A \in \mathcal{M}_{nn}$, alors, avec l'utilisation implicite des bases canoniques :*

$$A \text{ symétrique} \iff (\vec{x}, A.\vec{y})_{\mathbb{R}^n} = (A.\vec{x}, \vec{y})_{\mathbb{R}^n}, \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n. \tag{1.44}$$

Preuve. Pour un réel r, on a $[r]^T = [r] \overset{\text{noté}}{=} r$ (matrice 1*1), donc $([\vec{x}]^T.A.[\vec{y}])^T = [\vec{x}]^T.A.[\vec{y}]$, donc $[\vec{y}]^T.A^T.[\vec{x}] = [\vec{x}]^T.A.[\vec{y}]$, donc $[\vec{y}]^T.A.[\vec{x}] = [\vec{x}]^T.A.[\vec{y}]$, car $A^T = A$, ce pour tout $\vec{x}, \vec{y}$. Donc, avec (1.38) on obtient donne (1.44). ▪

Exercice 1.47 Soit $A \in \mathcal{M}_{mn}$. Montrer que les matrices $A^T.A \in \mathcal{M}_{nn}$ et $A.A^T \in \mathcal{M}_{mm}$ sont symétriques.

Réponse. $(A^T.A)^T = A^T.(A^T)^T = A^T.A$, et $(A.A^T)^T = A.A^T$. ▪

1.3.2 Matrices conjuguées et transconjuguées

Définition 1.48 Soit $A \in \mathcal{M}_{m,n}(\mathbb{C})$ une matrice complexe. Sa matrice conjuguée est la matrice $\overline{A} \in \mathcal{M}_{m,n}(\mathbb{C})$ définie par :

$$\overline{A} = [(\overline{A})_{ij}] \qquad \text{où} \qquad (\overline{A})_{ij} = \overline{A_{ij}} \overset{\text{noté}}{=} \overline{A}_{ij}. \tag{1.45}$$

Définition 1.49 Soit $A \in \mathcal{M}_{m,n}(\mathbb{C})$ une matrice complexe. Sa matrice transconjuguée est la matrice $A^* \in \mathcal{M}_{n,m}(\mathbb{C})$ donnée par :

$$A^* \stackrel{\text{def}}{=} (\overline{A})^T \qquad \text{i.e.} \qquad \forall i,j=1,...,n, \quad (A^*)_{ij} = \overline{A}_{ji}, \tag{1.46}$$

donc A^* est la transposée de la conjuguée. (A^* est également appelée la matrice adjointe de A.)

Exemple 1.50 Si $A = \begin{pmatrix} 1+i & 2+2i \\ 3 & -4i \end{pmatrix}$, on a $\overline{A} = \begin{pmatrix} 1-i & 2-2i \\ 3 & 4i \end{pmatrix}$ et $A^* = \begin{pmatrix} 1-i & 3 \\ 2-2i & 4i \end{pmatrix}$. ∴

Exemple 1.51 Si A est réelle alors $A^* = A^T$. ∴

Proposition 1.52 *La transposition commute avec la conjugaison :*

$$(\overline{A})^T = A^* = \overline{A^T}. \tag{1.47}$$

Et si $A \in \mathcal{M}_{m,n}(\mathbb{C})$ et $B \in \mathcal{M}_{n,p}(\mathbb{C})$ alors :

$$\overline{A.B} = \overline{A}.\overline{B}, \qquad \text{et} \qquad (A.B)^* = B^*.A^*. \tag{1.48}$$

Preuve. (1.47) car : $(A^*)_{ij} = ((\overline{A})^T)_{ij} = (\overline{A})_{ji} \stackrel{\text{def}}{=} \overline{A_{ji}} = \overline{(A^T)_{ij}}$ pour tout i,j.
Puis $\overline{\sum_k A_{ik}B_{kj}} = \sum_k \overline{A}_{ik}\overline{B}_{kj}$ pour tout i,j.
Et $(A.B)^* = \overline{(A.B)^T} = \overline{B^T.A^T} = \overline{B^T}.\overline{A^T} = B^*.A^*$. ∴

Remarque 1.53 Attention : si $D = \text{diag}(\lambda_1,...,\lambda_n)$ alors $D^* \neq D$ quand au moins un des λ_i n'est pas réel, car $D^* = \text{diag}(\overline{\lambda}_1,...,\overline{\lambda}_n)$. ∴

1.3.3 Matrice complexe et relation $(A.\vec{x},\vec{y})_{\mathbb{C}^n} = [\overline{\vec{y}}]^T.A.[\vec{x}] = (\vec{x},A^*.\vec{y})_{\mathbb{C}^n}$

Soit $(\vec{E}_i^{(k)})$ la base canonique de $\mathbb{C}^k$. Pour $\vec{x} = \sum_{i=1}^k x_i\vec{E}_i^{(k)}$ on note $[\vec{x}] = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$. Et soit $(\cdot,\cdot)_{\mathbb{C}^k}$ le produit scalaire usuel dans $\mathbb{C}^k$.

Corollaire 1.54 *Si $A \in \mathcal{M}_{mn}(\mathbb{C})$, si $\vec{x} \in \mathbb{C}^n$, si $\vec{y} \in \mathbb{C}^m$, alors :*

$$(A.\vec{x},\vec{y})_{\mathbb{C}^m} = [\overline{\vec{y}}]^T.[A].[\vec{x}] = (\vec{x},A^*.\vec{y})_{\mathbb{C}^n} \qquad (= \overline{(A^*.\vec{y},\vec{x})_{\mathbb{C}^n}} = \overline{[\overline{\vec{x}}]^T.A^*.[\vec{y}]}). \tag{1.49}$$

En particulier :

$$(A.\vec{E}_j^{(n)},\vec{E}_i^{(m)})_{\mathbb{C}^m} = A_{ij} = [\vec{E}_i]^T.[A].[\vec{E}_j], \tag{1.50}$$

car $\vec{E}_i = \overline{\vec{E}}_i$.

Preuve. Avec (1.31), où donc on utilise les bases canoniques, avec $\vec{x} = \sum_{i=1}^m x_i\vec{E}_i^{(m)}$ et $\vec{y} = \sum_{i=1}^n y_i\vec{E}_i^{(n)}$, $(A.\vec{x},\vec{y})_{\mathbb{C}^m} = \sum_{i=1}^m (A.\vec{x})_i\overline{y}_i = \sum_{i=1}^m\sum_{j=1}^n \overline{y}_iA_{ij}x_j = [\overline{\vec{y}}]^T.[A].[\vec{x}]$ et $(\vec{x},A^*.\vec{y})_{\mathbb{C}^n} = \sum_{j=1}^n x_j\overline{(A^*.\vec{y})_j} = \sum_{j=1}^n x_j(A^T.\overline{\vec{y}})_j = \sum_{j=1}^n\sum_{i=1}^m x_j(A^T)_{ji}\overline{y}_i = [\overline{\vec{y}}]^T.[A].[\vec{x}]$. ∴

1.3.4 Matrice hermitienne

Définition 1.55 $A \in \mathbb{C}^{n^2}$ (matrice complexe $n*n$) est dite hermitienne (ou auto-adjointe) ssi elle est égale à sa transconjuguée :

$$A \text{ hermtienne} \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad A^* = A, \qquad \text{i.e.} \qquad \forall i,j=1,...,n, \quad (A^*)_{ij} = A_{ij}. \tag{1.51}$$

En particulier une matrice hermitienne a ses éléments diagonaux qui sont réels : $\overline{A}_{ii} = A_{ii}$.

Exemple 1.56 $A = \begin{pmatrix} 1 & 2+2i \\ 2-2i & 3 \end{pmatrix}$ est hermitienne. ∴

Exemple 1.57 Une matrice symétrique réelle est hermitienne car alors $\overline{A}_{ji} = A_{ji} = A_{ij}$. ∴

Corollaire 1.58 *Avec l'utilisation implicite des bases canoniques :*

$$A \text{ hermitienne} \iff \forall \vec{x}, \vec{y} \in \mathbb{C}^n, \quad (A.\vec{x}, \vec{y})_{\mathbb{C}^n} = (\vec{x}, A.\vec{y})_{\mathbb{C}^n} \quad (= [\overline{\vec{y}}]^T.[A].[\vec{x}]). \tag{1.52}$$

Preuve. $(A.\vec{x}, \vec{y})_{\mathbb{C}^m} = [\overline{\vec{y}}]^T.[A].[\vec{x}]$, et $(\vec{x}, A.\vec{y})_{\mathbb{C}^n} = \overline{(A.\vec{y}, \vec{x})_{\mathbb{C}^n}} = (\overline{A}.\overline{\vec{y}}, \overline{\vec{x}})_{\mathbb{C}^n} = [\vec{x}]^T.\overline{A}.[\overline{\vec{y}}] = [\overline{\vec{y}}]^T.\overline{A}^T.[\vec{x}] = [\overline{\vec{y}}]^T.A.[\vec{x}]$, car $A^* = A$ (matrice hermitienne), d'où (1.52). ∎

Exercice 1.59 Soit $A \in \mathcal{M}_{mn}$. Montrer que les matrices $A^*.A \in \mathcal{M}_{nn}$ et $A.A^* \in \mathcal{M}_{mm}$ sont hermitiennes.

Réponse. $(A^*.A)^* = A^*.(A^*)^* = A^*.A$, idem pour $A.A^*$. ∎

Exercice 1.60 Montrer que si A est antisymétrique réelle (matrice telle que $A^T = -A$), alors iA est hermitienne. (Ainsi une matrice antisymétrique réelle sera diagonalisable dans $\mathbb{C}$, ses valeurs propres étant dans $i\mathbb{R}$, car les valeurs propres d'une matrice hermitienne sont réelles).

Réponse. Soit $B = iA$, donc $\overline{B} = -i\overline{A} = -iA$ car A est réelle. Donc $B^* = (\overline{B})^T = -iA^T = +iA$ car A est antisymétrique, soit $B^* = B$.

Exemple : $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ donne $iA = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, donc $(iA)^T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$, donc $(iA)^* = iA$. ∎

1.4 Matrice inversible

Soit $K = \mathbb{R}$ ou $\mathbb{C}$. Soit V un espace vectoriel de dimension n sur K.

1.4.1 Définition

Définition 1.61 Soit $A = [A_{ij}] \in K^{n^2}$ une matrice $n * n$ dans K. La matrice A est inversible ssi il existe $B \in K^{n^2}$ (matrice $n * n$) telle que :

$$A.B = I \qquad \text{et} \qquad B.A = I. \tag{1.53}$$

Et on note alors $B = A^{-1}$.

Exemple 1.62 Si $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2^2}$ matrice 2×2 alors $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. ∎

Soit $(\vec{E}_i)$ la base canonique de K^n.

Proposition 1.63 *Si A est inversible et si $A.\vec{x} = \vec{0}$ alors $\vec{x} = \vec{0}$.*

Preuve. $A.\vec{x} = \vec{0}$ avec A inversible implique $A^{-1}.A.\vec{x} = \vec{0}$, soit $I.\vec{x} = \vec{0}$, soit $\vec{x} = \vec{0}$. ∎

Proposition 1.64 *Soit A et B deux matrices $n * n$:*

$$\text{si} \quad A.B = I \quad \text{alors} \quad (B.\vec{E}_i)_{i=1,\dots,n} \quad \text{est une base de } K^n. \tag{1.54}$$

D'où :

$$\text{si} \quad A.B = I \quad \text{alors} \quad B.A = I. \tag{1.55}$$

Et donc A est inversible ssi il existe B telle que $A.B = I$, ou ssi il existe B telle que $B.A = I$.

En particulier, si A est inversible, alors ses "vecteurs colonnes" $A.\vec{E}_j$ forment une base.

N.B. : le résultat (1.55) est faux en dimension infinie.

Preuve. Supposons $A.B = I$. Posons $\vec{b}_j = B.\vec{E}_j$, et donc $A.\vec{b}_j = \vec{E}_j$ pour tout j. Supposons $(\vec{b}_i)$ famille liée. Quitte à renommer les $\vec{b}_i$, supposons $\vec{b}_n$ combinaison linéaire des autres : $\vec{b}_n = \sum_{i=1}^{n-1} \alpha_i \vec{b}_i$ avec au moins un des α_i non nul. Donc $\vec{E}_n = A.\vec{b}_n = \sum_{i=1}^{n-1} \alpha_i A.\vec{b}_i = \sum_{i=1}^{n-1} \alpha_i \vec{E}_i$, donc $(\vec{E}_i)$ famille liée : absurde. Donc $(\vec{b}_i)$ est une famille libre, et est composée de n vecteurs, donc c'est une base dans K^n : on a montré (1.54).

Et donc on a $(B.A).\vec{b}_j = (B.A).B.\vec{E}_j = B.(A.B).\vec{E}_j = B.\vec{E}_j = \vec{b}_j$ pour tout j. Donc $B.A = I$, cf proposition 1.31 : on a montré (1.55).

Donc pour vérifier que A est inversible, il suffit de vérifier qu'il existe une matrice B telle que $A.B = I$ puisqu'alors $B.A = I$. Ou encore il suffit de vérifier que $B.A = I$ puisqu'alors $A.B = I$.

Contre-exemple de (1.55) en dimension infinie : soit V l'espace vectoriel des suites réelles $\vec{x} = (x_i)_{i\in\mathbb{N}^*} = (x_1, x_2, ...)$. Soit A l'application linéaire de "décalage à gauche" : $A : \vec{x} = (x_1, x_2, ...) \to A.\vec{x} = (x_2, x_3, ...)$ (on a perdu x_1). Soit B l'application linéaire de "décalage à droite" : $B : \vec{y} = (y_1, y_2, ...) \to B.\vec{y} = (0, y_1, y_2, ...)$.

On a bien $A.B = I$ (on a bien $A.B.\vec{y} = \vec{y}$ pour tout $\vec{y}$), mais on n'a pas $B.A = I$ puisque $(B.A).\vec{x} = (0, x_2, x_3, ...)$ (on a perdu x_1). Ici la famille $(\vec{b}_j = B.\vec{E}_j)_{j\in\mathbb{N}^*}$ n'est pas une base de V : elle est bien libre, mais elle n'est pas génératrice (on a $B.\vec{E}_j = \vec{E}_{j+1}$ et $\vec{E}_1$ n'est pas combinaison linéaire des $B.\vec{E}_j$). Ou encore la famille $(\vec{a}_i = A.\vec{E}_i)_{i\in\mathbb{N}^*}$ n'est pas une base (ce n'est pas une famille libre car $\vec{a}_1 = \vec{0}$). ∎

Proposition 1.65 *Si A et B sont deux matrices $n*n$ inversibles, alors $A.B$ est inversible et :*

$$(A.B)^{-1} = B^{-1}.A^{-1}. \tag{1.56}$$

Preuve. $(B^{-1}.A^{-1}).(A.B) = B^{-1}.(A^{-1}.A).B = B^{-1}.B = I$, d'où (1.56) grâce à (1.55). ∎

Proposition 1.66 *Si $A = [a_{ij}] \in K^{n^2}$ (matrice $n\times n$) est inversible alors A^T est inversible et :*

$$(A^T)^{-1} = (A^{-1})^T. \tag{1.57}$$

Et on peut ainsi noter (non ambigu) :

$$A^{-T} \stackrel{\text{def}}{=} (A^T)^{-1} = (A^{-1})^T. \tag{1.58}$$

En particulier, si $A \in \mathbb{R}^{n^2}$ est symétrique et inversible alors A^{-1} est symétrique.

Si $A \in \mathbb{C}^{n^2}$ est inversible alors A^* est inversible et :

$$(A^*)^{-1} = (A^{-1})^*. \tag{1.59}$$

Et on peut ainsi noter (non ambigu) :

$$A^{-*} \stackrel{\text{def}}{=} (A^*)^{-1} = (A^{-1})^*. \tag{1.60}$$

Et si A est hermitienne et inversible alors A^{-1} est hermitienne.

Preuve. A inversible, donc il existe B telle que $A.B = I$, donc $B^T.A^T = I^T = I$, donc, cf. proposition 1.64, A^T est inversible d'inverse $B^T = (A^{-1})^T$.

Et $B^T.A^T = I$ donne $(A^T)^{-1}.(B^T)^{-1} = I$, cf. (1.56), donc $(A^T)^{-1} = B^T = (A^{-1})^T$.

Donc si A est réelle symétrique, $(A^{-1})^T = (A^T)^{-1} = A^{-1}$, donc A^{-1} est symétrique.

Dans $\mathbb{C}^n$, A inversible, donc il existe B telle que $A.B = I$, donc $B^*.A^* = I^* = I$ (cf. (1.48)), donc, cf. proposition 1.64, A^* est inversible d'inverse $B^* = (A^{-1})^*$.

Et $B^*.A^* = I$ donne $(A^*)^{-1}.(B^*)^{-1} = I$, cf. (1.56), donc $(A^*)^{-1} = B^* = (A^{-1})^*$.

Donc si A est hermitienne, $(A^{-1})^* = (A^*)^{-1} = A^{-1}$, donc A^{-1} est hermitienne. ∎

1.4.2 Matrice de passage

Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ sont deux bases. Soit P_{ij} la i-ème composante de $\vec{b}_j$ dans la base $(\vec{a}_i)$, et $P = [P_{ij}]$:

$$\vec{b}_j = \sum_{i=1}^n P_{ij}\vec{a}_i, \quad \text{i.e.} \quad [\vec{b}_j]_{|\vec{a}} = \begin{pmatrix} P_{ij} \\ \vdots \\ P_{nj} \end{pmatrix}, \quad \text{et} \quad P = [P_{ij}] := \big([\vec{b}_1]_{|\vec{a}} \quad ... \quad [\vec{b}_n]_{|\vec{a}} \big). \tag{1.61}$$

Définition 1.67 P est appelée la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$.

Exemple 1.68 Dans $\mathbb{R}^2$. Si $\vec{b}_1 = \vec{a}_1$ et $\vec{b}_2 = \vec{a}_1 + \vec{a}_2$, i.e., $[\vec{b}_1]_{|\vec{a}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $[\vec{b}_2]_{|\vec{a}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, alors $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$ (lire les colonnes). ▪

Autrement dit, la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$ est la matrice de l'application linéaire $\mathcal{P} : V \to V$ (endomorphisme) caractérisée par $\mathcal{P}.\vec{a}_j := \vec{b}_j$ pour tout j, i.e.

$$\forall j \in [1,n]_{\mathbb{N}}, \quad \mathcal{P}.\vec{a}_j = \sum_{i=1}^{n} P_{ij}\vec{a}_i, \quad \text{et} \quad [\mathcal{P}]_{|\vec{a}} := [P_{ij}] = P. \tag{1.62}$$

Proposition 1.69 *Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases. Soit $P = [P_{ij}]$ la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$. Soit $Q = [Q_{ij}]$ la matrice de passage de la base $(\vec{b}_i)$ à la base $(\vec{a}_i)$. Alors les matrices P et Q sont (inversibles et) inverses l'une de l'autre : on a*

$$\forall j \in [1,n]_{\mathbb{N}}, \quad \vec{b}_j = \sum_{i=1}^{n} P_{ij}\vec{a}_i \quad \text{et} \quad \vec{a}_j = \sum_{i=1}^{n} Q_{ij}\vec{b}_i \quad \Longrightarrow \quad P^{-1} = Q. \tag{1.63}$$

Autrement dit : si $[\vec{b}_j]_{|(\vec{a})} = \begin{pmatrix} P_{ij} \\ \vdots \\ P_{nj} \end{pmatrix}$ et $[\vec{a}_j]_{|(\vec{b})} = \begin{pmatrix} Q_{ij} \\ \vdots \\ Q_{nj} \end{pmatrix}$ pour tout j, alors $P^{-1} = Q$.

Preuve. On a :

$$\vec{a}_j = \sum_{i=1}^{n} Q_{ij}\vec{b}_i = \sum_{i=1}^{n} Q_{ij}(\sum_{k=1}^{n} P_{ki}\vec{a}_k) = \sum_{k=1}^{n}(\sum_{i=1}^{n} P_{ki}Q_{ij})\vec{a}_k = \sum_{k=1}^{n}(PQ)_{kj}\vec{a}_k.$$

Et $(\vec{a}_i)$ est une base, donc $(PQ)_{kj} = \delta_{kj}$ (pour tout j,k), soit $PQ = I$. ▪

Exemple 1.70 $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est la matrice de passage d'une base $(\vec{a}_1, \vec{a}_2)$ vers la base (lecture par colonne) $\vec{b}_1 = \vec{a}_1$ et $\vec{b}_2 = \vec{a}_1 + \vec{a}_2$. Et $Q = P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ est la matrice de passage de la base $(\vec{b}_i)$ vers la base $(\vec{a}_i)$: on a $\vec{a}_1 = \vec{b}_1$ et $\vec{a}_2 = -\vec{b}_1 + \vec{b}_2$ (lecture des colonnes de Q). ▪

Proposition 1.71 *Toute matrice inversible est une matrice de passage.*

Preuve. Soit P une matrice inversible. Soit $(\vec{a}_i)$ une base. Soit $\vec{b}_j = \sum_{i=1}^n P_{ij}\vec{a}_i$, autrement dit, soit $P = \begin{pmatrix} [\vec{b}_1]_{|(\vec{a})} & \dots & [\vec{b}_n]_{|(\vec{a})} \end{pmatrix}$. Donc $I = P^{-1}.P = \begin{pmatrix} P^{-1}.[\vec{b}_1]_{|(\vec{a})} & \dots & P^{-1}.[\vec{b}_n]_{|(\vec{a})} \end{pmatrix}$. On en déduit que les n vecteurs $\vec{b}_i$ sont indépendants (ils forment une base) : sinon, quitte à renommer les $\vec{b}_i$, supposons $\vec{b}_n$ combinaison linéaire des autres : $\vec{b}_n = \sum_{i=1}^{n-1} \alpha_i\vec{b}_i$. Donc $[\vec{b}_n]_{|(\vec{a})} = \sum_{i=1}^{n-1} \alpha_i[\vec{b}_i]_{|(\vec{a})}$. Donc $P^{-1}.[\vec{b}_n]_{|(\vec{a})} = \sum_{i=1}^{n-1} \alpha_i P^{-1}.[\vec{b}_i]_{|(\vec{a})}$, soit $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = [\vec{a}_n]_{|(\vec{a})} = \sum_{i=1}^{n-1} \alpha_i[\vec{a}_i]_{|(\vec{a})}$: c'est absurde : dans le membre de gauche la dernière composante vaut 1 alors que dans le membre de droite la dernière composante vaut 0. Donc P est une matrice de passage : de la base $(\vec{a}_i)$ vers la base $(\vec{b}_i)$. ▪

Exercice 1.72 Soit $\sigma : j \in [1,n]_{\mathbb{N}} \to \sigma(j) \in [1,n]_{\mathbb{N}}$ une permutation (une application bijective de $\{1,...,n\}$ dans lui-même). Soit $(\vec{E}_i)$ la base canonique de $\mathbb{R}^n$ et $[\vec{E}_j]_{|(\vec{E})} \overset{\text{noté}}{=} [\vec{E}_j]$ la matrice colonne dont tous les éléments sont nuls sauf le j-ème qui vaut 1. Montrer que la matrice de passage de la base $(\vec{E}_i)$ à la base $(\vec{E}_{\sigma(i)})$ (la base "réordonnée") est la matrice :

$$P = \begin{pmatrix} [\vec{E}_{\sigma(1)}] & \dots & [\vec{E}_{\sigma(n)}] \end{pmatrix}, \tag{1.64}$$

autrement dit, $[\vec{E}_{\sigma(j)}] = P.[\vec{E}_j]$ pour tout j.

En déduire que si $(\vec{a}_i)$ une base d'un espace vectoriel de V de dimension n, alors la matrice de passage de la base $(\vec{a}_i)$ à la base réordonnée $(\vec{a}_{\sigma(i)})$ est la matrice P donnée en (1.64) : pour tout j :

$$[\vec{a}_{\sigma(j)}] = P.[\vec{a}_j], \tag{1.65}$$

où $[\vec{a}_j] = [\vec{a}_j]_{|(\vec{a})} = [\vec{E}_j]$ et $[\vec{a}_{\sigma(j)}] = [\vec{a}_{\sigma(j)}]_{|(\vec{a})} = [\vec{E}_{\sigma(j)}]$ (égalités de matrices colonnes).

Réponse. P est une matrice de changement de base car les "vecteurs colonnes" $\vec{b}_j = \vec{E}_{\sigma(j)}$ forment une base, avec :

$$\vec{b}_j = \vec{E}_{\sigma(j)} = \sum_{i=1}^{n} P_{ij}\vec{E}_i \quad = \text{"colonne } j\text{" de } P.$$

Et dans V de dimension n, on a égalité des matrices colonnes $[\vec{a}_j]_{|(\vec{a})} = [\vec{E}_j]_{|(\vec{E})} = \begin{pmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{pmatrix}$ puisque $\vec{a}_j = \sum_i \delta_{ij}\vec{a}_i$ et $\vec{E}_j = \sum_i \delta_{ij}\vec{E}_i$. D'où (1.65). ∎

Proposition 1.73 *Si P est inversible et si $(\vec{a}_i)$ est une base, alors les $\vec{b}_j = \sum_{i=1}^n P_{ij}\vec{a}_i$ forment une base.*

Preuve. Si $(\vec{b}_i)_{i=1,\dots,n}$ est une famille liée, alors l'un des vecteurs est combinaison linéaire des autres. Quitte à les renommer supposons $\vec{b}_n$ combinaison linéaire des autres : $\vec{b}_n = \sum_{i=1}^{n-1} \alpha_i\vec{b}_i$. Donc $P.[\vec{a}_n]_{|\vec{a}} = [\vec{b}_n]_{|\vec{a}} = \sum_{i=1}^{n-1} \alpha_i[\vec{b}_i]_{|\vec{a}} = \sum_{i=1}^{n-1} \alpha_i P.[\vec{a}_i]_{|\vec{a}}$, donc $[\vec{a}_n]_{|\vec{a}} = \sum_{i=1}^{n-1} \alpha_i[\vec{a}_i]_{|\vec{a}}$: absurde : dans le membre de gauche la dernière composante vaut 1 alors que dans le membre de droite la dernière composante vaut 0. ∎

1.4.3 Changement de base pour les vecteurs : $[\vec{x}]_{|(\vec{b})} = P^{-1}.[\vec{x}]_{|(\vec{a})}$

Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases d'un espace vectoriel V de dimension n.

Soit $\vec{v} \in V$ (un vecteur dans V). On note x_i et y_i ses composantes sur les bases $(\vec{a}_i)$ et $(\vec{b}_i)$, i.e.

$$\vec{v} = \sum_{i=1}^{n} x_i\vec{a}_i = \sum_{i=1}^{n} y_i\vec{b}_i, \quad \text{i.e.} \quad [\vec{v}]_{|\vec{a}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad [\vec{v}]_{|\vec{b}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{1.66}$$

Proposition 1.74 *On a (formule de changement de base = relation entre les coordonnées d'un vecteur :*

$$[\vec{v}]_{|\vec{b}} = P^{-1}.[\vec{v}]_{|\vec{a}}. \tag{1.67}$$

Preuve. $\vec{v} = \sum_{i=1}^n y_i\vec{b}_i = \sum_{i=1}^n y_i \sum_{k=1}^n P_{ki}\vec{a}_k = \sum_{k=1}^n (\sum_{i=1}^n P_{ki}y_i)\vec{a}_k = \vec{x} = \sum_{k=1}^n x_k\vec{a}_k$, et $(\vec{a}_i)$ est une base, donc $\sum_{i=1}^n P_{ki}y_i = x_k$ pour tout k, soit $[\vec{v}]_{|\vec{a}} = P.[\vec{x}]_{|\vec{b}}$. ∎

1.4.4 Matrice orthogonale (ou orthonormale) et matrice unitaire

Définition 1.75 Une matrice $A \in \mathbb{R}^{n^2}$ est orthogonale (ou orthonormale) ssi :

$$A^T.A = I. \tag{1.68}$$

Une matrice $A \in \mathbb{C}^{n^2}$ est unitaire ssi :

$$A^*.A = I. \tag{1.69}$$

Proposition 1.76 *Les matrices orthogonales et unitaires sont inversibles.*

Dans $\mathbb{R}^{n^2}$:

$$A^T.A = I \quad \Longrightarrow \quad A^{-1} = A^T \quad \text{et} \quad A.A^T = I. \tag{1.70}$$

Dans $\mathbb{C}^{n^2}$:

$$A^*.A = I \quad \Longrightarrow \quad A^{-1} = A^* \quad \text{et} \quad A.A^* = I. \tag{1.71}$$

Preuve. $A^T.A = I$ implique A inversible et $A^{-1} = A^T$; et $A.A^{-1} = I$ donne $A.A^T = I$.

$A^*.A = I$ implique A inversible et $A^{-1} = A^*$; et $A.A^{-1} = I$ donne $A.A^* = I$. ∎

1.4.5 B.o.n. : relations $P^T.P = I$ dans $\mathbb{R}$ et $P^*.P = I$ dans $\mathbb{C}$

Soit $(\vec{v}_i)_{i=1,\dots,n}$ une base dans K^n et $P = [P_{ij}] = (\,[\vec{v}_1] \quad \dots \quad [\vec{v}_n]\,)$ la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{v}_i)$: on a posé $\vec{v}_j = \sum_{i=1}^n P_{ij}\vec{E}_i$, et $[\vec{v}_j] = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}$, pour tout j.

Proposition 1.77 *Cas $\mathbb{R}^n$: on a :*

$$(\vec{v}_i)_{i=1,\dots,n} \text{ b.o.n. dans } (\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n}) \iff P^T.P = I. \tag{1.72}$$

Cas $\mathbb{C}^n$: on a :

$$(\vec{v}_i)_{i=1,\dots,n} \text{ b.o.n. dans } (\mathbb{C}^n, (\cdot,\cdot)_{\mathbb{C}^n}) \iff P^*.P = I. \tag{1.73}$$

Preuve. $(\vec{v}_i, \vec{v}_j)_{\mathbb{C}^n} = (\sum_k P_{ki}\vec{E}_k, \sum_\ell P_{\ell i}\vec{E}_\ell)_{\mathbb{C}^n} = \sum_{k\ell} P_{ki}\overline{P}_{\ell j}(\vec{E}_k, \vec{E}_\ell)_{\mathbb{C}^n} = \sum_{k\ell} P_{ki}\overline{P}_{\ell j}\delta_{k\ell} = \sum_k (P^*)_{jk}P_{ki} = (P^*.P)_{ji}$. Donc $(\vec{v}_i, \vec{v}_j)_{\mathbb{C}^n} = \delta_{ji}$ pour tout i, j ssi $(P^*.P)_{ji} = \delta_{ji}$ pour tout i, j, i.e. ssi $P^*.P = I$. Et cas $\mathbb{R}^n$: on a $P^* = P^T$. ▪

Proposition 1.78 *Si $P^T.P = I$, alors $P.P^T = I$.*
Si $P^*.P = I$, alors $P.P^* = I$.

1.4.6 Décompositions $L.U$ (Gauss), $L.L^T$ (Cholesky), $Q.R$ (Gram–Schmidt)

Théorème 1.79 *(Décomposition de Gauss.) Si $A \in \mathcal{M}_{nn}(\mathbb{R})$ est inversible alors il existe une matrice triangulaire inférieure L, une matrice triangulaire supérieure U et une matrice de permutation P telles que :*

$$P.A = L.U. \tag{1.74}$$

(La matrice de permutation sert à permuter les lignes, cf. propositions 1.40 et 1.42, de manière à avoir des pivots non nuls.)

Preuve. Exercice. ▪

Théorème 1.80 *(Décomposition de Cholesky.) Si $A \in \mathcal{M}_{nn}(\mathbb{R})$ est symétrique inversible, alors il existe une matrice triangulaire inférieure L telle que :*

$$A = L.L^T. \tag{1.75}$$

Si de plus A est définie positive alors cette décomposition est unique.

Preuve. Exercice. ▪

Théorème 1.81 *(Décomposition $Q.R$ et Gram–Schmidt.) Si $A \in \mathcal{M}_{nn}(\mathbb{R})$ est inversible, alors il existe une matrice orthogonale Q (donc t.q. $Q^T.Q = I$) et une matrice triangulaire supérieure R (Right matrix) telle que :*

$$A = Q.R. \tag{1.76}$$

Et les vecteurs colonnes $\vec{q}_j$ de $Q = (\,[\vec{q}_1] \quad \dots \quad [\vec{q}_n]\,)$ constituent une b.o.n. de $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$ dit base de Gram–Schmidt (procédé de diagonalisation de Gram–Schmidt).

Preuve. Si $A = (\,[\vec{a}_1] \quad \dots \quad [\vec{a}_n]\,)$, $Q = (\,[\vec{q}_1] \quad \dots \quad [\vec{q}_n]\,)$ et $R = [R_{ij}] = (\,[\vec{r}_1] \quad \dots \quad [\vec{r}_n]\,)$, on veut $\vec{a}_j = Q.\vec{r}_j$ pour tout j. On commence par $j = 1$: on veut $\vec{a}_1 = Q.(R_{11}\vec{E}_1) = R_{11}Q.\vec{E}_1 = R_{11}\vec{q}_1$ avec $||\vec{q}_1||_{\mathbb{R}^n} = 1$, d'où $\vec{q}_1 = \frac{\vec{a}_1}{||\vec{a}_1||_{\mathbb{R}^n}}$ et $R_{11} = ||\vec{a}_1||_{\mathbb{R}^n}$; on continue avec $j = 2$: on veut $\vec{a}_2 = Q.(R_{21}\vec{E}_1 + R_{22}\vec{E}_2) = R_{21}\vec{q}_1 + R_{22}\vec{q}_2$, avec $(\vec{q}_i, \vec{q}_j)_{\mathbb{R}^n} = \delta_{ij}$ pour $i, j = 1, 2...$ ▪

1.5 Le déterminant

1.5.1 Définition

Définition 1.82 Une forme n-linéaire $F : (K^n)^n \to K$ est une fonction linéaire par rapport à chaque variable, i.e. si on note :

$$F_{\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_{i+1},...,\vec{x}_n}(\vec{x}) = F(\vec{x}_1, ..., \vec{x}_{i-1}, \vec{x}, \vec{x}_{i+1}, ..., \vec{x}_n), \tag{1.77}$$

alors pour tout $\vec{x}_1, ..., \vec{x}_n \in K^n$, tout $i = 1, ..., n$, tout $\vec{v}_1, \vec{v}_2 \in K^n$, tout $\alpha \in K$:

$$F_{\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_{i+1},...,\vec{x}_n}(\vec{v}_1 + \alpha\vec{v}_2) = F_{\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_{i+1},...,\vec{x}_n}(\vec{v}_1) + \alpha F_{\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_{i+1},...,\vec{x}_n}(\vec{v}_2). \tag{1.78}$$

Définition 1.83 Une forme n-linéaire $F : K^n \times ... \times K^n \to K$ est alternée ssi, pour tout $\vec{v}_1, ..., \vec{v}_n \in K^n$ et pour tout $i, j \in [1, n]_{\mathbb{N}}$:

$$F(..., \vec{v}_i, ..., \vec{v}_j, ...) = -F(..., \vec{v}_j, ..., \vec{v}_i, ...), \tag{1.79}$$

les autres $\vec{v}_k$ aux autres emplacements étant inchangés.

Proposition 1.84 *Soit $\vec{v}_1, ..., \vec{v}_n \in K^n$ et soit F une forme n-linéaire alternée. Si l'un des $\vec{v}_i$ est nul, alors $F(\vec{v}_1, ..., \vec{v}_n) = 0$.*

Preuve. On a $F(\vec{0}, \vec{v}_2, ..., \vec{v}_n) = F(2 \times \vec{0}, \vec{v}_2, ..., \vec{v}_n) = 2F(\vec{0}, \vec{v}_2, ..., \vec{v}_n)$ par linéarité relativement au premier emplacement, donc $F(\vec{0}, \vec{v}_2, ..., \vec{v}_n) = 0$. Idem pour les autres emplacements. ▪

Proposition 1.85 *Si F est n-linéaire alternée sur K^n, si $(\vec{E}_i)$ est la base canonique de K^n, alors F est complètement déterminée par la donnée du réel $F(\vec{E}_1, ..., \vec{E}_n)$.*

Preuve. Une forme n-linéaire F est déterminée ssi on connaît les images pour tous les vecteurs d'une base, i.e. ssi en particulier pour la base canonique $(\vec{E}_i)$ on connaît les $F(\vec{E}_{i_1}, ..., \vec{E}_{i_n})$ pour tout $i_1, ..., i_n \in [1, n]_{\mathbb{N}}$. De plus si F est alternée alors soit $F(\vec{E}_{i_1}, ..., \vec{E}_{i_n}) = 0$ (si deux des vecteurs au moins sont égaux), soit $F(\vec{E}_{i_1}, ..., \vec{E}_{i_n}) = \varepsilon(\sigma)F(\vec{E}_1, ..., \vec{E}_n)$ où $\varepsilon(\sigma)$ est la signature de la permutation (bijection) $\sigma : k \in [1, n]_{\mathbb{N}} \to \sigma(k) = i_k \in [1, n]_{\mathbb{N}}$: la donnée du seul réel $F(\vec{E}_1, ..., \vec{E}_n)$ détermine F. ▪

Définition 1.86 Le déterminant est la forme n-linéaire alternée $\det : (K^n)^n \to K$ telle que $\det(\vec{E}_1, ..., \vec{E}_n) = 1$.

Remarque 1.87 Dans $\mathbb{R}^n$, le volume algébrique du parallélépipède limité par les n vecteurs $\vec{v}_1, ..., \vec{v}_n \in \mathbb{R}^n$ est le réel $\det(\vec{v}_1, ..., \vec{v}_n)$.

Le volume (des mécaniciens) est la valeur absolue de ce déterminant. ▪

Proposition 1.88 *L'ensemble des formes n-linéaires alternées sur K^n est* $\text{Vect}\{\det\}$ *l'espace vectoriel de dimension* 1 *engendré par la forme n linéaire alternée "le déterminant".*

Preuve. Soit F une forme n-linéaire alternée. Soit $\alpha = F(\vec{E}_1, ..., \vec{E}_n)$. Alors $F = \alpha \det$: en effet, soit $F(\vec{E}_{i_1}, ..., \vec{E}_{i_n}) = 0 = \det(\vec{E}_{i_1}, ..., \vec{E}_{i_n})$ (si deux des vecteurs au moins sont égaux), soit $F(\vec{E}_{i_1}, ..., \vec{E}_{i_n}) = \varepsilon(\sigma)F(\vec{E}_1, ..., \vec{E}_n) = \varepsilon(\sigma)\alpha \det(\vec{E}_1, ..., \vec{E}_n) = \alpha \det(\vec{E}_{i_1}, ..., \vec{E}_{i_n})$ où $\varepsilon(\sigma)$ est la signature de la permutation (bijection) $\sigma : k \in [1, n]_{\mathbb{N}} \to \sigma(k) = i_k \in [1, n]_{\mathbb{N}}$. ▪

1.5.2 Déterminant d'une matrice

Définition 1.89 Soit $A \in K^{n^2}$ une matrice $n * n$. Le déterminant $\det(A)$ de la matrice A est le déterminant de "ses vecteurs colonnes", i.e. des vecteurs de K^n dont les composantes dans la base canonique sont stockées dans les colonnes de A.

Donc, notant $(\vec{E}_i)$ est la base canonique, si $A = [A_{ij}] = (\,[\vec{a}_1] \;\; ... \;\; [\vec{a}_n]\,)$, avec donc $\vec{a}_j = \sum_{i=1}^n A_{ij}\vec{E}_i$ et $[\vec{a}_j] = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{pmatrix}$ (la j-ème matrice colonne), alors par définition :

$$\det(A) \stackrel{\text{def}}{=} \det(\vec{a}_1, ..., \vec{a}_n). \tag{1.80}$$

En particulier $\det(I) = 1$ ($= \det(\vec{E}_1, ..., \vec{E}_n)$).

Proposition 1.90 *Si $T = [T_{ij}]$ est une matrice triangulaire supérieure (ou inférieure) alors :*

$$\det(T) = T_{11}T_{22}...T_{nn}, \tag{1.81}$$

produit des éléments diagonaux.

Preuve. Soit $\vec{v}_j$ le j-ème vecteur colonne de T, donc $T = ([\vec{v}_1] \ \ ... \ \ [\vec{v}_n])$. En particulier $\vec{v}_1 = T_{11}\vec{E}_1$. Si $T_{11} = 0$ alors $\vec{v}_1 = 0$, donc $\det(T) = 0$, et (1.81) est vérifiée. Si $T_{11} \neq 0$, alors par multilinéarité, on a $\det(T) = \det(\vec{v}_1, \vec{v}_2, ..., \vec{v}_n) = T_{11} \det(\frac{1}{T_{11}}\vec{v}_1, \vec{v}_2 - \frac{T_{12}}{T_{11}}\vec{v}_1, ..., \vec{v}_n - \frac{T_{1n}}{T_{11}}\vec{v}_1) = T_{11}\det(U)$ où $U = [U_{ij}]$ est la matrice qui vérifie $U_{11} = 1$, $U_{1j} = 0$ pour tout $j \geq 2$, et $U_{ij} = T_{ij}$ pour tout autre i, j (c'est la matrice T dans laquelle on a remplacé la première colonne par $\vec{E}_1$, et la première ligne, sauf le premier élément, par des 0). On poursuit le processus : $\det(T) = T_{11}T_{22}...T_{nn}\det(\vec{E}_1, ..., \vec{E}_n) = T_{11}T_{22}...T_{nn}$.

Si T est triangulaire inférieure, on initialise le processus précédent en partant de $\vec{v}_n = T_{nn}\vec{E}_n$. ∎

Proposition 1.91 *On a :*

$$(\vec{v}_i)_{i=1,...,n} \text{ famille libre} \quad \Longleftrightarrow \quad \det(\vec{v}_1, ..., \vec{v}_n) \neq 0. \tag{1.82}$$

D'où :

$$A \text{ est inversible} \quad \Longleftrightarrow \quad \det(A) \neq 0. \tag{1.83}$$

Et :

$$\det(A^T) = \det(A). \tag{1.84}$$

Preuve. Supposons $(\vec{v}_i)_{i=1,...,n}$ libre. Notons $\vec{v}_j = \sum_{i=1}^n P_{ij}\vec{E}_i$. Quitte à réordonner la famille $(\vec{v}_i)$, supposons $P_{11} \neq 0$ (si tous les P_{1i} sont nuls alors tous les $\vec{v}_j$ sont combinaisons linéaires de $\vec{E}_2, ..., \vec{E}_n$ et donc $(\vec{v}_i)$ est liée). Par multilinéarité, $\det(\vec{v}_1, \vec{v}_2, ..., \vec{v}_n) = P_{11}\det(\frac{1}{P_{11}}\vec{v}_1, \vec{v}_2 - \frac{P_{12}}{P_{11}}\vec{v}_1, ..., \vec{v}_n - \frac{P_{1n}}{P_{11}}\vec{v}_1)$ (on a mis des zéros sur la "première ligne"). On continue le processus : $\det(\vec{v}_1, \vec{v}_2, ..., \vec{v}_n) = P_{11}...P_{nn}\det(T)$ où les P_i sont tous non nuls et T est la matrice triangulaire inférieure avec des 1 sur la diagonale, donc $\det(T) = 1$.

Si $(\vec{v}_i)$ est liée, quitte à réordonner on peut supposer $\vec{v}_n = \sum_{i=1}^{n-1} \alpha_i\vec{v}_i$. Donc, det étant multilinéaire alternée, on a $\det(\vec{v}_1, ..., \vec{v}_{n-1}, \vec{v}_n) = \det(\vec{v}_1, ..., \vec{v}_{n-1}, \vec{v}_n - \sum_{i=1}^{n-1}\alpha_i\vec{v}_i) = 0$. Donc si $\det(\vec{v}_1, ..., \vec{v}_n) \neq 0$ alors $(\vec{v}_i)_{i=1,...,n}$ est une famille libre.

D'où (1.83).

(1.84) est vraie pour les matrices triangulaires (produit des éléments diagonaux). Et pour une matrice $A = ([\vec{v}_1] \ \ ... \ \ [\vec{v}_n])$ on vient de voir que $\det(A) = P_{11}...P_{nn}\det(T)$. De même $\det(A^T) = P_{11}...P_{nn}\det(T^T)$, donc $= P_{11}...P_{nn}\det(T) = \det(A)$. D'où (1.84). ∎

Proposition 1.92 *Soit $A \in K^{n^2}$, et $\vec{v}_1, ..., \vec{v}_n$ dans $\mathbb{R}^n$. Alors (utilisation implicite de la base canonique) :*

$$\det(A.\vec{v}_1, ..., A.\vec{v}_n) = \det(A).\det(\vec{v}_1, ..., \vec{v}_n). \tag{1.85}$$

Donc, si A et B sont des matrices $n * n$, alors :

$$\det(A.B) = \det(A)\det(B). \tag{1.86}$$

Preuve. L'application $F : (\vec{v}_1, ..., \vec{v}_n) \to \det(A.\vec{v}_1, ..., A.\vec{v}_n)$ est une forme multilinéaire alternée (immédiat par "linéarité" de A). Donc il exsite $\alpha \in \mathbb{R}$ t.q. $F = \alpha \det$, cf. proposition 1.88.

Et pour la base canonique $(\vec{E}_i)$, on a $F(\vec{E}_1, ..., \vec{E}_n) = \det(A.\vec{E}_1, ..., A.\vec{E}_n) = \det(A)$ (car $A.\vec{E}_j$ est le j-ème vecteur colonne de la matrice A), d'où $\alpha = \det(A)$, d'où $F = \det(A)\det$, d'où (1.85). D'où (1.86), cf. (1.36). ∎

Proposition 1.93 *Soit S_n est l'ensemble des permutations de $\{1, ..., n\}$ et $\varepsilon(\sigma)$ la signature de la permutation σ. Alors :*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n A_{\sigma(i),i}. \tag{1.87}$$

Preuve. Appliquer le caractère n-alterné, voir cours sur les déterminants. ∎

1.5.3 Exemples

La multilinéarité et (1.84) permettent de calculer un déterminant en développant par exemple par rapport à une de ses colonnes ou une de ses lignes.

Exemple 1.94 Dans K^2, montrer que pour $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ on a :

$$\det(A) = a_{11}a_{22} - a_{21}a_{12}, \tag{1.88}$$

opération appelée de “développement par rapport à la première colonne”.

Réponse. 1- On a $A = (\,[\vec{c}_1] \quad [\vec{c}_2]\,)$, où les $\vec{c}_j$ sont les “vecteurs colonnes de A”, avec en particulier $\vec{c}_1 = a_{11}\vec{E}_1 + a_{21}\vec{E}_2$. Donc par multilinéarité :

$$\begin{aligned}\det A &= a_{11}\det(\,[\vec{E}_1] \quad [\vec{c}_2]\,) + a_{21}\det(\,[\vec{E}_2] \quad [\vec{c}_2]\,) \\ &= a_{11}\det\begin{pmatrix} 1 & a_{12} \\ 0 & a_{22} \end{pmatrix} + a_{21}\det\begin{pmatrix} 0 & a_{12} \\ 1 & a_{22} \end{pmatrix}.\end{aligned}$$

2- Puis en enlevant a_{12} fois la première colonne à la seconde (forme alternée) dans le premier déterminant et a_{22} fois la première colonne à la seconde dans le second déterminant, on a :

$$\begin{aligned}\det A &= a_{11}\det\begin{pmatrix} 1 & 0 \\ 0 & a_{22} \end{pmatrix} + a_{21}\det\begin{pmatrix} 0 & a_{12} \\ 1 & 0 \end{pmatrix} \\ &= a_{11}\det((\,[\vec{E}_1] \quad a_{22}[\vec{E}_2]\,)) + a_{21}\det((\,[\vec{E}_2] \quad a_{12}[\vec{E}_1]\,)) \\ &= a_{11}a_{22}\det((\,[\vec{E}_1] \quad [\vec{E}_2]\,)) + a_{21}a_{12}\det((\,[\vec{E}_2] \quad [\vec{E}_1]\,)),\end{aligned}$$

avec la multilinéarité. ∎

Exercice 1.95 Dans K^3, développer par rapport à la première colonne le déterminant de la matrice $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

Réponse. Montrons :

$$\det(A) = a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}, \tag{1.89}$$

opération appelée de “développement par rapport à la première colonne”.

1- On a $A = (\,[\vec{c}_1] \quad [\vec{c}_2] \quad [\vec{c}_3]\,)$, où les $\vec{c}_j$ sont les “vecteurs colonnes de A”, avec en particulier $\vec{c}_1 = a_{11}\vec{E}_1 + a_{21}\vec{E}_2 + a_{31}\vec{E}_3$. Donc par multilinéarité :

$$\begin{aligned}\det A &= a_{11}\det(\,[\vec{E}_1] \quad [\vec{c}_2] \quad [\vec{c}_3]\,) + a_{21}\det(\,[\vec{E}_2] \quad [\vec{c}_2] \quad [\vec{c}_3]\,) + a_{31}\det(\,[\vec{E}_3] \quad [\vec{c}_2] \quad [\vec{c}_3]\,) \\ &= a_{11}\det\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{21}\det\begin{pmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{31}\det\begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{pmatrix}.\end{aligned}$$

2- Puis en enlevant “une constante fois la première colonne” aux autres colonne, on a par exemple (forme alternée) $\det\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} = \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}$.

3- Et ce dernier déterminant est le “volume de la dernière matrice” : la longueur “horizontale”, ici 1, fois l'aire “verticale” $\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$.

De même pour les autres termes, d'où, sachant qu'un échange de deux lignes change le signe (forme alternée), on obtient (1.89). ∎

2 Applications linéaires et matrices

Soit V et W deux espaces vectoriels de dimension finie sur un corps K (ici $\mathbb{R}$ ou $\mathbb{C}$). On notera $\dim V = n$ et $\dim W = m$.

2.1 Définitions

Une application $\mathcal{L} : V \to W$ est linéaire ssi $\mathcal{L}(\vec{x}+\alpha\vec{y}) = \mathcal{L}(\vec{x}) + \alpha\mathcal{L}(\vec{y})$ pour tout $\vec{x},\vec{y} \in V$ et tout $\alpha \in K$. Pour une application linéaire $\mathcal{L}$, on utilisera la notation :

$$\mathcal{L}(\vec{x}) \overset{\text{noté}}{=} \mathcal{L}.\vec{x}, \tag{2.1}$$

qui est la notation du produit, car la linéarité permet de “distribuer” : $\mathcal{L}.(\vec{x}+\alpha\vec{y}) = \mathcal{L}.\vec{x} + \alpha\mathcal{L}.\vec{y}$.

Définition 2.1 Un endomorphisme est une application linéaire d'un espace vectoriel dans lui-même.

2.2 Représentations matricielles d'une application linéaire

Une application linéaire existe indépendamment du choix d'une base. Le choix d'une base dépend de celui qui fait les calculs (pour quantifier l'action de l'application linéaire sur un vecteur).

Soit $\mathcal{L} : V \to W$ une application linéaire, avec $\dim V = n$ et $\dim W = m$.
Soit $(\vec{a}_i)_{i=1,...,n}$ une base de V et $(\vec{w}_i)_{i=1,...,m}$ une base de W.
Notons A_{ij} les composantes de $\mathcal{L}.\vec{a}_j$ dans la base $(\vec{w}_i)$:

$$\mathcal{L}.\vec{a}_j = \sum_{i=1}^{m} A_{ij}\vec{w}_i, \quad \forall j = 1,...,n, \qquad \text{donc} \qquad [\mathcal{L}.\vec{a}_j]_{|(\vec{w})} = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{pmatrix}, \quad \forall j = 1,...,n. \tag{2.2}$$

Définition 2.2 La matrice :

$$[\mathcal{L}]_{|(\vec{a}),(\vec{w})} = [A_{ij}]_{\substack{i=1,...,m \\ j=1,...,n}}. \tag{2.3}$$

est appelée matrice représentant $\mathcal{L}$ relativement aux bases $(\vec{a}_i)$ et $(\vec{w}_j)$. Ainsi les composantes de $\mathcal{L}.\vec{a}_j$ dans la base $(\vec{w}_i)$ sont donc stockées dans la colonne j de $A = [\mathcal{L}]_{|(\vec{a}),(\vec{w})}$.

Par linéarité, si $\vec{x} = \sum_{j=1}^n x_j\vec{a}_j$, on a avec $A = [\mathcal{L}]_{|(\vec{a}),(\vec{w})}$:

$$\mathcal{L}.\vec{x} = \sum_{j=1}^{n} x_j\mathcal{L}.\vec{a}_j = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}x_j\vec{w}_i, \qquad \text{soit} \qquad [\mathcal{L}.\vec{x}]_{|(\vec{w})} = A.[\vec{x}]_{|(\vec{a})} = \begin{pmatrix} \sum_{j=1}^n A_{1j}x_j \\ \vdots \\ \sum_{j=1}^n A_{nj}x_j \end{pmatrix}. \tag{2.4}$$

Remarque 2.3 Donc une application linéaire a autant de matrices la représentant qu'il y a de bases : une application linéaire a une infinité de représentations matricielles. ▪

Remarque 2.4 Les notations tensorielles lèvent les ambiguïtés de notation, voir polycopié d'algèbre linéaire. On note $a^i : V \to \mathbb{R}$ la forme linéaire définie par $a^i(\vec{a}_j) = \delta_{ij}$. Ainsi (a^i) est la base duale de $(\vec{a}_i)$. On note $A_{ij} = A^i_j$ (convention d'Einstein pour les applications linéaires et leurs composantes) :

$$\mathcal{L} = \sum_{i=1}^{m}\sum_{j=1}^{n} A^i_j\vec{w}_i \otimes a^j, \tag{2.5}$$

Ainsi, notant $\vec{x} = \sum_{i=1}^n x^i\vec{a}_i$ (convention d'Einstein pour les vecteurs et leurs composantes), ayant $a^j.\vec{x} = \sum_{i=1}^n x^i(a^j.\vec{a}_i) = x^j$ (par définition de la base duale) :

$$\mathcal{L}.\vec{x} = (\sum_{i=1}^{m}\sum_{j=1}^{n} A^i_j\vec{w}_i \otimes a^j).\vec{x} = \sum_{i=1}^{m}\sum_{j=1}^{n} A^i_j\vec{w}_i(a^j.\vec{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n} A^i_j\vec{w}_i x^j, \tag{2.6}$$

et on retrouve (2.4). ▪

Définition 2.5 Cas particulier des endomorphismes $\mathcal{L}$ (applications linéaires de V dans lui-même) : on prend pour base dans $W = V$ la base $(\vec{w}_i) = (\vec{a}_i)$ (un observateur décrit un phénomène modélisé par un endomorphisme $\mathcal{L}$ à l'aide d'une base unique dans V), et on note (simplification de (2.3)) :

$$[\mathcal{L}]_{|(\vec{a}),(\vec{a})} = [\mathcal{L}]_{|(\vec{a})}, \tag{2.7}$$

où donc $[\mathcal{L}]_{|(\vec{a})}$ est la matrice représentant l'endomorphisme $\mathcal{L}$ dans la base $(\vec{a}_i)$.

2.3 Compositions d'applications linéaires et matrices

Soit U, V, W trois espaces vectoriels sur K de dimension p, n, m.

Soit $\mathcal{L} : U \to V$ et $\mathcal{M} : V \to W$ deux applications linéaires.

On dispose donc de la composée $\mathcal{M} \circ \mathcal{L} : U \to W$.

Soit $(\vec{a}_i)$, $(\vec{b}_i)$ et $(\vec{c}_i)$ des bases respectives de U, V et W.

Soit $A = [\mathcal{L}]_{|(\vec{a}),(\vec{b})} \in \mathcal{M}_{np}$, $B = [\mathcal{M}]_{|(\vec{b}),(\vec{c})} \in \mathcal{M}_{mn}$ et $C = [\mathcal{M} \circ \mathcal{L}]_{|(\vec{a}),(\vec{c})} \in \mathcal{M}_{mp}$ les matrices des applications linéaires dans les bases indiquées.

Proposition 2.6 *L'application linéaire $\mathcal{M} \circ \mathcal{L}$ vérifie :*

$$C = B.A \qquad \text{i.e.} \qquad [\mathcal{M} \circ \mathcal{L}]_{|(\vec{a}),(\vec{c})} = [\mathcal{M}]_{|(\vec{b}),(\vec{c})}.[\mathcal{L}]_{|(\vec{a}),(\vec{b})}. \tag{2.8}$$

Preuve. $\mathcal{M} \circ \mathcal{L}$ est bien linéaire : $(\mathcal{M} \circ \mathcal{L})(\vec{x} + \alpha\vec{y}) = \mathcal{M}(\mathcal{L}(\vec{x} + \alpha\vec{y})) = \mathcal{M}(\mathcal{L}(\vec{x}) + \alpha\mathcal{L}(\vec{y})) = \mathcal{M}(\mathcal{L}(\vec{x})) + \alpha\mathcal{M}(\mathcal{L}(\vec{y})) = (\mathcal{M} \circ \mathcal{L})(\vec{x}) + \alpha(\mathcal{M} \circ \mathcal{L})(\vec{y})$ par linéarité de $\mathcal{L}$ et de $\mathcal{M}$. Et pour tout j, d'une part $(\mathcal{M} \circ \mathcal{L})(\vec{a}_j)) = \sum_{k=1}^{z} C_{kj}\vec{c}_k$, et d'autre part $\mathcal{M}(\mathcal{L}(\vec{a}_j)) = \mathcal{M}(\sum_{i=1}^{n} A_{ij}\vec{b}_i) = \sum_{i=1}^{n} A_{ij}\mathcal{M}(\vec{b}_i) = \sum_{i=1}^{n} A_{ij} \sum_{k=1}^{z} B_{ki}\vec{c}_k = \sum_{k=1}^{z} (B.A)_{kj}\vec{c}_k$. Donc, $(\vec{c}_k)$ étant une base, $C_{kj} = (B.A)_{kj}$ pour tout k, j, soit $C = BA$. ∎

2.4 Inverse d'une application linéaire

Définition 2.7 Une application linéaire $\mathcal{L} : V \to W$ est inversible ssi il existe une application linéaire $\mathcal{M} : W \to V$ telle que :

$$\mathcal{L} \circ \mathcal{M} = I_W \qquad \text{et} \qquad \mathcal{M} \circ \mathcal{L} = I_V, \tag{2.9}$$

où I_V et I_W sont les identités dans V et W. Et alors on note $\mathcal{M} = \mathcal{L}^{-1}$.

Proposition 2.8 *Si V et W sont de dimension finie, alors une application linéaire $\mathcal{L} : V \to W$ est inversible ssi il existe une application linéaire $\mathcal{M} : W \to V$:*

$$\textit{telle que} \qquad \mathcal{L} \circ \mathcal{M} = I_W, \qquad \textit{ou telle que} \qquad \mathcal{M} \circ \mathcal{L} = I_V. \tag{2.10}$$

Preuve. Adapter la démonstration de la proposition 1.64. ∎

2.5 Matrices de l'endomorphisme de changement de base et son inverse

Supposons que dans V le choix de la base $(\vec{a}_i)$ soit fait par un premier "observateur", et que le choix d'une autre base $(\vec{b}_i)$ soit fait par un second "observateur" : le résultat $\mathcal{L}.\vec{x}$ doit pouvoir être comparé par les deux observateurs, ici à l'aide des "formules de changement de base".

Soit $\mathcal{P} : V \to V$ l'endomorphisme qui permet de passe de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$: pour tout $j = 1, ..., n$:

$$\mathcal{P}.\vec{a}_j = \vec{b}_j \quad \in V. \tag{2.11}$$

Notons P_{ij} les composantes de $\vec{b}_j$ dans la base $(\vec{a}_i)$: pour tout $j = 1, ..., n$:

$$\vec{b}_j = \sum_{i=1}^{n} P_{ij}\vec{a}_i, \qquad [\vec{b}_j]_{|(\vec{a})} = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \qquad [\mathcal{P}]_{|(\vec{a})} = [P_{ij}] = P. \tag{2.12}$$

Donc $P = \begin{pmatrix} [\vec{b}_j]_{|(\vec{a})} & \dots & [\vec{b}_j]_{|(\vec{a})} \end{pmatrix}$ est la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$: sa j-ème colonne contient les composantes de $\vec{b}_j = \mathcal{P}.\vec{a}_j$ dans la base $(\vec{a}_i)$.

(Notatiosn tensorielles : $\mathcal{P} = \sum_{ij} P^i_j \vec{a}_i \otimes a^j$, puisque $\mathcal{P}.\vec{a}_k = \sum_{ij} P^i_j \vec{a}_i (a^j.\vec{a}_k) = \sum_{ij} P^i_j \vec{a}_i \delta^j_k = \sum_i P^i_k \vec{a}_i = \vec{b}_k$ comme il se doit.)

En particulier P est inversible (matrice de passage) d'inverse $\mathcal{P}^{-1} : V \to V$ donné par :

$$\mathcal{P}^{-1}.\vec{b}_j = \vec{a}_j = \sum_{i=1}^{n} Q_{ij}\vec{b}_i, \qquad [\vec{a}_j]_{|(\vec{b})} = \begin{pmatrix} Q_{1j} \\ \vdots \\ Q_{nj} \end{pmatrix}, \qquad [\mathcal{P}^{-1}]_{|(\vec{b})} = [Q_{ij}] = Q. \tag{2.13}$$

En effet on vérifie que $\mathcal{P}^{-1}.(\mathcal{P}.\vec{a}_j) = \vec{a}_j$ pour tout j, et $\mathcal{P}.(\mathcal{P}^{-1}.\vec{b}_j) = \mathcal{P}.\vec{a}_j = \vec{b}_j$ pour tout j.

(Notations tensorielles, $\mathcal{P}^{-1} = \sum_{ij} Q^i_j \vec{b}_i \otimes b^j$.)

Proposition 2.9 *On a :*

$$Q = P^{-1}, \qquad \text{i.e.} \qquad [\mathcal{P}^{-1}]_{|(\vec{b})} = [\mathcal{P}]^{-1}_{|(\vec{a})}. \tag{2.14}$$

Preuve. Pour tout j on a :

$$\vec{a}_j = \mathcal{P}^{-1}.\vec{b}_j = \sum_{i=1}^n Q_{ij}\vec{b}_i = \sum_{i=1}^n Q_{ij}(\sum_{k=1}^n P_{ki}\vec{a}_k) = \sum_{k=1}^n(\sum_{i=1}^n P_{ki}Q_{ij})\vec{a}_k = \sum_{k=1}^n (P.Q)_{kj}\,\vec{a}_k,$$

donc $(P.Q)_{kj} = \delta_{kj}$ pour tout i,j (car $(\vec{a}_i)$ est une base), donc $P.Q = I$ comme souhaité. ∎

Proposition 2.10 *On a aussi :*

$$P = [\mathcal{P}]_{|(\vec{b})} \qquad (= [\mathcal{P}]_{|(\vec{a})}). \tag{2.15}$$

(Notations tensorielles : $\mathcal{P} = \sum_{ij} P^i_j \vec{a}_i \otimes a^j = \sum_{ij} P^i_j \vec{b}_i \otimes b^j$.)

Donc $Q = P$ est la matrice de représentation de $\mathcal{P}$ à la fois dans la base $(\vec{a}_i)$ et la base $(\vec{b}_i)$.

On a aussi :

$$Q = [\mathcal{P}^{-1}]_{|(\vec{a})} \qquad (= [\mathcal{P}^{-1}]_{|(\vec{b})}). \tag{2.16}$$

(Notations tensorielles : $\mathcal{P}^{-1} = \sum_{ij} Q^i_j \vec{b}_i \otimes b^j = \sum_{ij} Q^i_j \vec{a}_i \otimes a^j$.)

Donc $Q = P^{-1}$ est la matrice de représentation de $\mathcal{P}^{-1}$ à la fois dans la base $(\vec{a}_i)$ et la base $(\vec{b}_i)$.

Preuve. Soit $P = [\mathcal{P}]_{|(\vec{a})}$ et soit $R = [\mathcal{P}]_{|(\vec{b})}$. Donc $\mathcal{P}.\vec{a}_j = \sum_i P_{ij}\vec{a}_i$ et $\mathcal{P}.\vec{b}_j = \sum_i R_{ij}\vec{b}_i$, pour tout j. Vérifions $R = P$. Pour tout j, avec (2.13), on a $\vec{b}_j = \mathcal{P}^{-1}(\mathcal{P}.\vec{b}_j) = \mathcal{P}^{-1}(\sum_i R_{ij}\vec{b}_i) = \sum_i R_{ij}\mathcal{P}^{-1}.\vec{b}_i = \sum_i R_{ij}\sum_k Q_{ki}\vec{b}_k = \sum_k(\sum_i Q_{ki}R_{ij})\vec{b}_k = \sum_k (Q.R)_{kj}\vec{b}_k$, et $(\vec{b}_i)$ est une base, d'où $(Q.R)_{kj} = \delta_{kj}$ pour tout i,j, soit $Q.R = I$, donc $R = P$, cf. (2.14).

Soit $Q = [\mathcal{P}^{-1}]_{|(\vec{b})}$ et soit$S = [\mathcal{P}^{-1}]_{|(\vec{a})}$. Donc $\mathcal{P}^{-1}.\vec{b}_j = \sum_i Q_{ij}\vec{b}_i$ et $\mathcal{P}^{-1}.\vec{a}_j = \sum_i S_{ij}\vec{a}_i$, pour tout j. Vérifions $S = Q$. Pour tout j on a $\vec{a}_j = \mathcal{P}(\mathcal{P}^{-1}.\vec{a}_j) = \mathcal{P}(\sum_i S_{ij}\vec{a}_i) = \sum_i S_{ij}\mathcal{P}.\vec{a}_i = \sum_i S_{ij}\sum_k P_{ki}\vec{a}_k = \sum_k(\sum_i P_{ki}S_{ij})\vec{a}_k = \sum_k (P.S)_{kj}\vec{a}_k$, et $(\vec{a}_i)$ est une base, donc $(P.S)_{kj} = \delta_{kj}$ pour tout i,j, soit $PS = I$, donc $S = Q$, cf. (2.14). ∎

2.6 Changement de base pour les endomorphismes : $B = P^{-1}.A.P$

Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases dans V et soit $\mathcal{P}$ l'endomorphisme de changement de bases. Soit :

$$P = [\mathcal{P}]_{|(\vec{a})}, \qquad \text{et donc} \qquad P^{-1} = [\mathcal{P}^{-1}]_{|(\vec{b})}, \tag{2.17}$$

avec (2.14).

Proposition 2.11 *(Formule de changement de base.) Soit $\mathcal{L} : V \to V$ un endomorphisme. Alors :*

$$[\mathcal{L}]_{|(\vec{b})} = P^{-1}.[\mathcal{L}]_{|(\vec{a})}.P. \tag{2.18}$$

Soit, avec $[\mathcal{L}]_{|(\vec{a})} \overset{\text{noté}}{=} A$ et $[\mathcal{L}]_{|(\vec{b})} \overset{\text{noté}}{=} B$:

$$B = P^{-1}.A.P. \tag{2.19}$$

Preuve. On a $\mathcal{L}.\vec{a}_j = \sum_{i=1}^n A_{ij}\vec{a}_i$ et $\mathcal{L}.\vec{b}_j = \sum_{i=1}^n B_{ij}\vec{b}_i$ pour tout $j = 1,...,n$, par définition de $A = [\mathcal{L}]_{|(\vec{a})}$ et $B = [\mathcal{L}]_{|(\vec{b})}$, cf. (2.7). Donc, notant $Q = P^{-1}$, avec (2.13), on a pour tout j :

$$\begin{aligned}\sum_i B_{ij}\vec{b}_i &= \mathcal{L}.\vec{b}_j = \mathcal{L}.(\sum_k P_{kj}\vec{a}_k) = \sum_k P_{kj}\mathcal{L}.\vec{a}_k = \sum_k P_{kj}\sum_\ell A_{\ell k}\vec{a}_\ell \\ &= \sum_k P_{kj}\sum_\ell A_{\ell k}\sum_i Q_{i\ell}\vec{b}_i = \sum_i(\sum_{k\ell} Q_{i\ell}A_{\ell k}P_{kj})\vec{b}_i = \sum_i (Q.A.P)_{ij}\vec{b}_i,\end{aligned}$$

et $(\vec{b}_i)$ est une base, donc $B_{ij} = (Q.A.P)_{ij}$ pour tout i,j : on a (2.19), i.e. (2.18). ∎

Exemple 2.12 Soit $\mathcal{L} : \vec{x} \in \mathbb{R} \to \mathcal{L}(\vec{x}) \in \mathbb{R}$ une application linéaire, donc ici (on est en une variable d'espace) $\mathcal{L}.\vec{x} = \ell\vec{x}$ où $\ell \in \mathbb{R}$ (une multiplication par le réel ℓ). Soit $\vec{a}$ un vecteur représentant un mètre (m) et $\vec{b}$ un vecteur représentant un pied (foot ft). Par définition, 1 ft = 0,3048 m. Donc $[\vec{b}]_{|(\vec{a})} = 0,3048$: on a $\vec{b} = 0,3048\vec{a}$. Donc $P = [0,3048] \overset{\text{noté}}{=} 0,3048$ (matrice $1*1$ du changement de base $\mathcal{P}$ défini par $\vec{b} = \mathcal{P}.\vec{a}$). Donc $P^{-1} = \frac{1}{0{,}3048}$.

Soit $\vec{x}$ donné. Soit α (resp. β) sa longueur en mètre (resp. en pied) : $\vec{x} = \alpha\vec{a} = \beta\vec{b}$, où donc $[\vec{x}]_{|(\vec{a})} = \alpha$ et $[\vec{x}]_{|(\vec{b})} = \beta$, avec $[\vec{x}]_{|(\vec{b})} = P^{-1}.[\vec{x}]_{|(\vec{a})}$, cf. (1.67), soit $\beta = P^{-1}\alpha$, donc $\beta = \frac{1}{0{,}3048}\alpha$. (Exemple si $\beta = 1$ ft, alors $\alpha = 0,3048$ m.)

Et on a, formule de changement de bases des endomorphismes, $[\mathcal{L}]_{|(\vec{b})} = P^{-1}.[\mathcal{L}]_{|(\vec{a})}.P = \frac{1}{0{,}3048}.[\mathcal{L}]_{|(\vec{a})}.0,3048 = [\mathcal{L}]_{|(\vec{a})}$, soit $B = A$ (c'est “évident” : qu'on soit en pieds ou en mètre, une distance 2 fois plus longue double le résultat...).

Vérification : $[\mathcal{L}.\vec{x}]_{|(\vec{b})} = [\mathcal{L}]_{|(\vec{b})}.[\vec{x}]_{|(\vec{b})} = (P^{-1}.[\mathcal{L}]_{|(\vec{a})}.P).(P^{-1}.[\vec{x}]_{|(\vec{a})}) = P^{-1}.[\mathcal{L}]_{|(\vec{a})}.[\vec{x}]_{|(\vec{a})} = P^{-1}.[\mathcal{L}.\vec{x}]_{|(\vec{a})} = \frac{1}{0{,}3048}[\mathcal{L}.\vec{x}]_{|(\vec{a})} = P^{-1}.[\mathcal{L}.\vec{x}]_{|(\vec{a})}$ qui est bien la formule (1.67) de changement de base des vecteurs. ∎

Remarque 2.13 La formule $B = P^{-1}.A.P$ est indépendante de l'existence ou non de produit scalaire : il n'est jamais question de produit scalaire lors de sa justification, et en particulier les bases $(\vec{a}_i)$ et $(\vec{b}_i)$ sont quelconques (elles ne sont pas orthonormées, la notion d'orthogonalité dépend d'un produit scalaire). ∎

2.7 Matrices semblables

Définition 2.14 Deux matrices $A, B \in K^{n^2}$ sont semblables ssi il existe une matrice inversible $P \in K^{n^2}$ telle que :

$$B = P^{-1}.A.P. \tag{2.20}$$

Proposition 2.15 *Des matrices sont semblables ssi elles représentent un même endomorphisme $\mathcal{L} : V \to V$ dans deux bases différentes.*

Preuve. Les matrices d'un endomorphisme sont toutes semblables entre elles, cf. (2.18).

Réciproquement supposons $B = P^{-1}.A.P$. Soit $\mathcal{L}$ définie sur une base $(\vec{a}_i)$ par $[\mathcal{L}]_{|(\vec{a})} = A$, soit $\mathcal{L}.\vec{a}_j = \sum_{i=1}^n A_{ij}\vec{a}_i$ pour tout j. Soit $(\vec{b}_i)$ la base définie par $\vec{b}_j = \sum_{i=1}^n P_{ij}\vec{a}_i$ pour tout j. Alors (2.18) donne $B = [\mathcal{L}]_{|(\vec{b})}$. ∎

2.8 Endomorphisme symétrique

2.8.1 Définition d'un endomorphisme symétrique

Ici on suppose de plus que l'espace vectoriel réel V est muni d'un produit scalaire $(\cdot,\cdot)_V$.

Définition 2.16 Soit $\mathcal{L} : V \to V$ un endomorphisme. L'endomorphisme transposé $\mathcal{L}^T : V \to V$ de $\mathcal{L}$ relativement au produit scalaire $(\cdot,\cdot)_V$ est l'endomorphisme défini par :

$$\forall \vec{v}, \vec{w} \in V, \qquad (\mathcal{L}.\vec{v}, \vec{w})_V = (\vec{v}, \mathcal{L}^T.\vec{w})_V. \tag{2.21}$$

Définition 2.17 L'endomorphisme $\mathcal{L}$ est dit symétrique relativement au produit scalaire $(\cdot,\cdot)_V$ ssi $\mathcal{L}^T = \mathcal{L}$:

$$\forall \vec{v}, \vec{w} \in V, \qquad (\mathcal{L}.\vec{v}, \vec{w})_V = (\vec{v}, \mathcal{L}.\vec{w})_V. \tag{2.22}$$

2.8.2 Matrice d'un endomorphisme symétrique

Soit $(\vec{a}_i)$ une base de V. Soit M la matrice de masse relative à cette base :

$$M = [M_{ij}] \qquad \text{où} \qquad M_{ij} = (\vec{a}_i, \vec{a}_j)_V. \tag{2.23}$$

On suppose que l'endomorphisme $\mathcal{L} : V \to V$ est initialement décrit dans la base $(\vec{a}_i)$, i.e. on connaît $[\mathcal{L}]_{|(\vec{a})} = A = [A_{ij}]$:

$$\mathcal{L}.\vec{a}_j = \sum_{i=1}^n A_{ij}\vec{a}_i, \qquad [\mathcal{L}.\vec{a}_j]_{|(\vec{a})} = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{pmatrix}, \qquad A = [\mathcal{L}]_{|(\vec{a})}. \tag{2.24}$$

Question : si $A = [\mathcal{L}]_{|(\vec{a})}$ est symétrique, a-t-on $\mathcal{L}$ symétrique relativement à $(\cdot,\cdot)_V$?

Et si $\mathcal{L}$ est symétrique relativement à $(\cdot,\cdot)_V$, a-t-on $A = [\mathcal{L}]_{|(\vec{a})}$ symétrique ?
Réponse : ça dépend :

Proposition 2.18 *1- Si $(\vec{a}_i)$ est une b.o.n. dans $(V,(\cdot,\cdot)_V)$, alors : $\mathcal{L}$ est symétrique dans $(V,(\cdot,\cdot)_V)$ ssi sa matrice $A = [\mathcal{L}]_{|(\vec{a})}$ est symétrique.*

2- Si $(\vec{a}_i)$ n'est pas une b.o.n. dans $(V,(\cdot,\cdot)_V)$, alors : $\mathcal{L}$ est symétrique dans $(V,(\cdot,\cdot)_V)$ ssi $A^T.M = M.A$ (et alors la matrice $M.A$ est symétrique).

3- Et en général on peut avoir A symétrique et $\mathcal{L}$ non symétrique, ainsi que $\mathcal{L}$ symétrique et A non symétrique.

4- Un endomorphisme peut être symétrique pour un produit scalaire et être non symétrique pour un autre produit scalaire (la notion de symétrie relativement à un produit scalaire n'est pas intrinsèque).

Preuve. 1- Cas $(\vec{a}_i)$ b.o.n. : on a $(\mathcal{L}.\vec{a}_j, \vec{a}_i)_V = \sum_k A_{kj}(\vec{a}_k, \vec{a}_i)_V = \sum_k A_{kj}\delta_{ki} = A_{ij}$. Donc :
Si $\mathcal{L}$ est symétrique alors : $(\mathcal{L}.\vec{a}_i, \vec{a}_j)_V = (\mathcal{L}.\vec{a}_j, \vec{a}_i)_V$ donne $A_{ji} = A_{ij}$ et A est symétrique.
Si A est symétrique alors : $A_{ji} = A_{ij}$ donne $(\mathcal{L}.\vec{a}_i, \vec{a}_j)_V = (\mathcal{L}.\vec{a}_j, \vec{a}_i)_V$, et $\mathcal{L}$ est symétrique.

2- Cas $(\vec{a}_i)$ n'est pas une b.o.n. et $\mathcal{L}$ symétrique : $(\mathcal{L}.\vec{a}_i, \vec{a}_j)_V = \sum_k A_{ki}(\vec{a}_k, \vec{a}_j)_V = \sum_k A_{ki}M_{kj} = (A^T.M)_{ij}$; et de même $(\mathcal{L}.\vec{a}_j, \vec{a}_i)_V = (A^T.M)_{ji} = ((A^T.M)^T)_{ij} = (M.A)_{ij}$ car $M^T = M$. Donc $\mathcal{L}$ est symétrique ssi $A^T.M = M.A$. Et alors $(M.A)^T = A^T.M^T = A^T.M = M.A$ et $M.A$ est symétrique.

En particulier, si $(\vec{a}_i)$ est une b.o.n. on a $M = I$, et on retrouve $A^T = A$.

3- Contre-exemples dans $V = \mathbb{R}^2$, avec $(\vec{E}_1, \vec{E}_2)$ sa base canonique et $(\cdot,\cdot)_V = (\cdot,\cdot)_{\mathbb{R}^2}$ son produit scalaire usuel. On prend les vecteurs $\vec{a}_1 = \vec{E}_1 + \vec{E}_2$ et $\vec{a}_2 = \vec{E}_2$, avec donc $[\vec{a}_1] = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $[\vec{a}_2] = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$), et $P = (\,[\vec{a}_1], [\vec{a}_2]\,) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ la matrice de passage.

31- Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ (qui est bien symétrique). Soit $\mathcal{L}$ l'endomorphisme de matrice $[\mathcal{L}]_{|(\vec{a})} = A$, donc $\mathcal{L}.\vec{a}_1 = \vec{a}_1$ et $\mathcal{L}.\vec{a}_2 = 2\vec{a}_2$ cf. (2.24). On a $(\mathcal{L}.\vec{a}_1, \vec{a}_2)_{\mathbb{R}^2} = (\vec{a}_1, \vec{a}_2)_{\mathbb{R}^2} = 1$, et $(\mathcal{L}.\vec{a}_2, \vec{a}_1)_{\mathbb{R}^2} = (2\vec{a}_2, \vec{a}_1)_{\mathbb{R}^2} = 2$, donc $(\mathcal{L}.\vec{a}_1, \vec{a}_2)_{\mathbb{R}^2} \neq (\mathcal{L}.\vec{a}_2, \vec{a}_1)_{\mathbb{R}^2}$, donc $\mathcal{L}$ n'est pas symétrique bien que sa matrice A soit symétrique.

32- Soit $\mathcal{L}$ donnée par $\mathcal{L}.\vec{E}_1 = \vec{E}_1$ et $\mathcal{L}.\vec{E}_2 = 2\vec{E}_2$, donc $\mathcal{L}$ est symétrique (immédiat). Sa matrice dans la base $(\vec{E}_i)$ est la matrice $E = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ sa matrice dans la base $(\vec{a}_i)$ est la matrice $A = P^{-1}.E.P = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}$ non symétrique. Ici $\mathcal{L}$ est symétrique mais sa matrice dans la base $(\vec{a}_i)$ ne l'est pas.

4- L'endomorphisme $\mathcal{L}$ donné au 31- (dans la base non orthonormée $(\vec{a}_i)$) n'est pas symétrique relativement au produit scalaire usuel $(\cdot,\cdot)_{\mathbb{R}^2}$.

Définissons le produit scalaire $(\cdot,\cdot)_S = (S.,.)_{\mathbb{R}^2}$ qui fait de $(\vec{a}_1, \vec{a}_2)$ une b.o.n., i.e. donné par la matrice symétrique définie positive S t.q. $(S.\vec{a}_i, \vec{a}_j)_S = \delta_{ij}$, donc S t.q. $\delta_{ij} = [\vec{a}_j]^T.S.[\vec{a}_i] = [\vec{E}_j]^T.P^T.S.P.[\vec{E}_i]$. Donc, $P^T.S.P = I$, soit $S = P^{-T}.P^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

On vérifie : $(\mathcal{L}.\vec{a}_1, \vec{a}_2)_S = (\vec{a}_1, \vec{a}_2)_S = \delta_{12} = 0$ et $(\mathcal{L}.\vec{a}_2, \vec{a}_1)_S = (2\vec{a}_2, \vec{a}_1)_S = 2\delta_{12} = 0$: l'endomorphisme $\mathcal{L}$ est bien symétrique relativement au produit scalaire $(\cdot,\cdot)_S$. ▪

2.9 * Image, noyau, injectivité, surjectivité, bijectivité

2.9.1 Image et surjectivité

Soit $\mathcal{L} : V \to W$ une application linéaire où $\dim V = n$ et $\dim W = m$.

Définition 2.19 L'image $\mathrm{Im}(\mathcal{L})$ est le sous-ensemble de W :

$$\mathrm{Im}(\mathcal{L}) = \{\vec{w} \in W \text{ t.q. } \exists \vec{v} \in V,\ \vec{w} = \mathcal{L}.\vec{v}\} \overset{\text{noté}}{=} \{\mathcal{L}.\vec{v} : \vec{v} \in V\} \overset{\text{noté}}{=} \mathcal{L}(V). \tag{2.25}$$

Donc :

$$\vec{w} \in \mathrm{Im}(\mathcal{L}) \iff \exists \vec{v} \in V \quad \text{t.q.} \quad A.\vec{v} = \vec{w}, \tag{2.26}$$

et on dit alors que $\vec{v}$ est un antécédent de $\vec{w}$ par $\mathcal{L}$.

Soit $(\vec{v}_i)_{i=1,...,n}$ une base de V, $(\vec{w}_i)_{i=1,...,m}$ une base de W, et, pour tout $j = 1, ..., n$:

$$\mathcal{L}.\vec{v}_j = \sum_{i=1}^{m} A_{ij}\vec{w}_i, \qquad \text{où donc} \qquad [\mathcal{L}.\vec{v}_j]_{|(\vec{w})} = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{pmatrix}, \qquad [\mathcal{L}]_{|(\vec{v}),(\vec{w})} = [A_{ij}]. \tag{2.27}$$

Proposition 2.20 $\mathrm{Im}(\mathcal{L})$ *est le sous-espace vectoriel de* W *donné par :*

$$\mathrm{Im}(\mathcal{L}) = \mathrm{Vect}\{\mathcal{L}.\vec{v}_1, ..., \mathcal{L}.\vec{v}_n\}, \tag{2.28}$$

i.e. $\mathrm{Im}(\mathcal{L})$ *est engendré par "les vecteurs colonnes" de* $A = [\mathcal{L}]_{|(\vec{v}),(\vec{w})}$, *cf. (2.27).*

Preuve. $\mathrm{Im}(\mathcal{L})$ est stable par addition et par multiplication par un scalaire (immédiat car $\mathcal{L}$ est linéaire) : c'est bien un sous-espace vectoriel de W.

Si $\vec{v} = \sum_{j=1}^n v_j\vec{v}_j \in K^n$, alors $\mathcal{L}.\vec{v} = \sum_{j=1}^n v_j\mathcal{L}.\vec{v}_j$, donc $\mathcal{L}.\vec{v}$ est combinaison linéaire des $\mathcal{L}.\vec{v}_j$. D'où $\mathrm{Im}(\mathcal{L}) \subset \mathrm{Vect}\{\mathcal{L}.\vec{v}_1, ..., \mathcal{L}.\vec{v}_n\}$.

Et si $\vec{w} \in \mathrm{Vect}\{\mathcal{L}.\vec{v}_1, ..., \mathcal{L}.\vec{v}_n\}$, alors il existe n réels w_i t.q. $\vec{w} = \sum_j w_j\mathcal{L}.\vec{v}_j = \mathcal{L}.(\sum_j w_j\vec{v}_j) = \mathcal{L}.\vec{v}$ où on a posé $\vec{v} = \sum_j w_j\vec{v}_j$, donc $\vec{w} \in \mathrm{Im}(\mathcal{L})$. D'où (2.28). ■

Définition 2.21 $\mathcal{L}$ est "surjective" ssi $\mathrm{Im}(\mathcal{L}) = W$.

Proposition 2.22 $\mathcal{L}$ *est surjective ssi* $(\mathcal{L}.\vec{v}_1, ..., \mathcal{L}.\vec{v}_n)$ *est une famille génératrice de* W.

Et alors $\dim(V) \geq \dim(W)$, *soit* $n \geq m$ *(une matrice* A *de représentation de* $\mathcal{L}$ *a plus de colonnes que de lignes).*

Et donc si $n = m$ *et* $\mathcal{L}$ *est surjective, alors* $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ *est une base de* W.

Preuve. Si $\mathcal{L}$ est surjective, alors $\mathrm{Vect}\{\mathcal{L}.\vec{v}_1, ..., \mathcal{L}.\vec{v}_n\} = W$, donc $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ génère bien W. Réciproque immédiate. Et donc nécessairement $n \geq \dim(W) = m$. ■

2.9.2 Noyau et injectivité

Définition 2.23 Le "noyau de $\mathcal{L}$" est le sous-ensemble de V défini par :

$$\mathrm{Ker}\mathcal{L} = \mathrm{Vect}\{\vec{v} \in \mathcal{L} : \mathcal{L}.\vec{v} = \vec{0}\}. \tag{2.29}$$

Proposition 2.24 $\mathrm{Ker}\mathcal{L}$ *est un sous-espace vectoriel de* V.

Preuve. Immédiat. ■

Définition 2.25 $\mathcal{L}$ est injective ssi la relation $\mathcal{L}.\vec{x} = \mathcal{L}.\vec{y}$ implique $\vec{x} = \vec{y}$ (une image a un unique antécédent).

Proposition 2.26 $\mathcal{L}$ *est injective ssi* $\mathrm{Ker}\mathcal{L} = \{\vec{0}\}$, *i.e. ssi :* $\mathcal{L}.\vec{v} = \vec{0} \Leftrightarrow \vec{v} = \vec{0}$.

Et dans ce cas, nécessairement $\dim(V) \leq \dim(W)$, *soit* $n \leq m$ *(une matrice de représentation de* $[\mathcal{L}]$ *a moins de colonnes que de lignes).*

Et donc si $n = m$ *et* $\mathcal{L}$ *est injective, alors* $\mathcal{L}$ *est surjective.*

Preuve. Supposons $\mathcal{L}$ injective. Donc $\mathcal{L}.\vec{x} = \mathcal{L}.\vec{0}$ implique $\vec{x} = \vec{0}$. Donc, comme $\mathcal{L}.\vec{0} = \vec{0}$ (linéarité de $\mathcal{L}$), si $\mathcal{L}.\vec{x} = \vec{0}$ alors $\vec{x} = \vec{0}$, et donc $\mathrm{Ker}\mathcal{L} = \{\vec{0}\}$, cf. (2.29)).

Réciproquement, $\mathcal{L}.\vec{x} = \mathcal{L}.\vec{y}$ équivaut à $\mathcal{L}.(\vec{x} - \vec{y}) = \vec{0}$ (par linéarité de $\mathcal{L}$), donc équivaut à $\vec{x} - \vec{y} \in \mathrm{Ker}\mathcal{L}$. Donc si $\mathrm{Ker}\mathcal{L} = \{\vec{0}\}$ et si $\mathcal{L}.\vec{x} = \mathcal{L}.\vec{y}$ alors $\vec{x} - \vec{y} \in \mathrm{Ker}\mathcal{L} = \{\vec{0}\}$, i.e. $\vec{x} - \vec{y} = \vec{0}$, i.e. $\vec{x} = \vec{y}$. Autrement dit, si $\mathrm{Ker}\mathcal{L} = \{\vec{0}\}$ alors $\mathcal{L}$ est injective.

Et donc, pour $\mathcal{L}$ injective et $(\vec{v}_i)_{i=1,...,k}$ une base de V on a $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ est une famille libre dans W : si α_i sont des scalaires t.q. $\sum_{i=1}^n \alpha_i\mathcal{L}.\vec{v}_i = \vec{0} = \mathcal{L}.(\sum_{i=1}^n \alpha_i\vec{v}_i)$, alors $\sum_{i=1}^n \alpha_i\vec{v}_i = \vec{0}$ (injectivité de $\mathcal{L}$), donc $\alpha_i = 0$ pour tout i car $(\vec{v}_i)$ est une base. Donc $m \geq n$.

Et si $m = n$ alors $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ est une base de W, donc $\mathcal{L}$ est surjective. ■

Cas particulier $V = K^n$ ($= \mathbb{R}^n$ ou $\mathbb{C}^n$) muni de son produit scalaire usuel et de sa base usuelle $(\vec{E}_i)$. Soit :

$$[\mathcal{L}]_{|(\vec{E}^{(n)}),(\vec{w})} = A \tag{2.30}$$

la matrice de $\mathcal{L} : \mathbb{R}^n \to W$ relativement à la base canonique de K^n et à une base $(\vec{w}_i)$ de W. Notons :

$$\vec{\ell}_i = \sum_{j=1}^{n} A_{ij}\vec{E}_j^{(n)} \in V, \qquad \text{donc} \qquad [\vec{\ell}_i]^T = (\, A_{i1} \quad ... \quad A_{in}\,) \qquad \text{et} \qquad A = \begin{pmatrix} [\vec{\ell}_1]^T \\ \vdots \\ [\vec{\ell}_m]^T \end{pmatrix}. \tag{2.31}$$

Ainsi $[\vec{\ell}_i]^T$ le "i-ème vecteur ligne" de A.

Proposition 2.27 *1- Cas $V = \mathbb{R}^n$ muni de son produit scalaire usuel : on a :*

$$\mathrm{Ker}\mathcal{L} = (\mathrm{Vect}\{\vec{\ell}_1, ..., \vec{\ell}_m\})^\perp, \tag{2.32}$$

orthogonal de l'espace vectoriel engendré par les "vecteurs lignes" de A.

2- Cas $V = \mathbb{C}^n$ muni de son produit scalaire usuel : on a :

$$\mathrm{Ker}\mathcal{L} = (\mathrm{Vect}\{\overline{\vec{\ell}_1}, ..., \overline{\vec{\ell}_m}\})^\perp, \tag{2.33}$$

orthogonal de l'espace vectoriel engendré par les conjugués des "vecteurs lignes" de A.

Et donc, si $m = n$ et $\mathcal{L}$ est injective (donc $\mathrm{Ker}\mathcal{L} = \{\vec{0}\}$) alors $(\vec{\ell}_i)$ est une base de $\mathbb{R}^m$ (les vecteurs lignes forment une base de $\mathbb{R}^m$).

Preuve. 1- Dans $\mathbb{R}^n$: $[\mathcal{L}.\vec{x}] = \begin{pmatrix} [\vec{\ell}_1]^T.[\vec{x}] \\ \vdots \\ [\vec{\ell}_m]^T.[\vec{x}] \end{pmatrix} = \begin{pmatrix} (\vec{\ell}_1, \vec{x})_{\mathbb{R}^n} \\ \vdots \\ (\vec{\ell}_m, \vec{x})_{\mathbb{R}^n} \end{pmatrix}$, donc $\mathcal{L}.\vec{x} = \vec{0}$ ssi $(\vec{\ell}_i, \vec{x})_{\mathbb{R}^n} = 0$ pour tout $i = 1, ..., m$ ($\vec{x}$ est orthogonal à "toutes les lignes"), d'où (2.32).

2- Dans $\mathbb{C}^n$: la première composante de $\mathcal{L}.\vec{x}$ est donnée par $(\mathcal{L}.\vec{x})_1 = \sum_{j=1}^n A_{1j}x_j = \sum_{j=1}^n (\vec{\ell}_1)_j x_j$. On a donc $(\mathcal{L}.\vec{x})_1 = (\vec{\ell}_1, \overline{\vec{x}})_{\mathbb{C}^n}$. Idem pour les autres composantes.

Donc $\mathcal{L}.\vec{x} = \begin{pmatrix} (\vec{\ell}_1, \overline{\vec{x}})_{\mathbb{C}^n} \\ \vdots \\ (\vec{\ell}_m, \overline{\vec{x}})_{\mathbb{C}^n} \end{pmatrix}$. Donc $\mathcal{L}.\vec{x} = \vec{0}$ ssi $(\vec{\ell}_i, \overline{\vec{x}})_{\mathbb{C}^n} = 0$ pour tout i, donc ssi $0 = \overline{(\overline{\vec{x}}, \vec{\ell}_i)_{\mathbb{C}^n}} = (\vec{x}, \overline{\vec{\ell}_i})_{\mathbb{C}^n}$ pour tout i, d'où (2.33). ∎

2.9.3 Relation image ↔ noyau : $\mathrm{Ker}(\mathcal{L}^T) = (\mathrm{Im}\mathcal{L})^\perp$

On munit V et W de produits scalaires $(\cdot,\cdot)_V$ et $(\cdot,\cdot)_W$.

Définition 2.28 Soit $\mathcal{L} : V \to W$ une application linéaire. L'application linéaire transposée $\mathcal{L}^T : W \to V$ est définie relativement aux produits scalaires $(\cdot,\cdot)_V$ et $(\cdot,\cdot)_W$ par :

$$(\mathcal{L}^T.\vec{w}, \vec{v})_V = (\vec{w}, \mathcal{L}.\vec{v})_W. \tag{2.34}$$

Cette définition n'est pas intrinsèque : elle dépend des choix des produits scalaires (elle dépend des utilisateurs).

(Pour avoir une définition intrinsèque, il faut considérer les espaces duaux et les applications linéaires adjointes $\mathcal{L}^* : W^* \to V^*$, voir cours algèbre linéaire.)

On définit l'orthogonal d'un sous-espace vectoriel S d'un espace vectoriel E muni d'un produit scalaire $(\cdot,\cdot)_E$ par :

$$S^\perp = \{t \in E : \ \forall s \in S, \ (t,s)_E = 0\}. \tag{2.35}$$

Proposition 2.29 *On a :*

$$\mathrm{Ker}(\mathcal{L}^T) = (\mathrm{Im}\mathcal{L})^\perp \qquad \text{et} \qquad \mathrm{Im}(\mathcal{L}^T) = (\mathrm{Ker}\mathcal{L})^\perp. \tag{2.36}$$

Et donc, si $\mathcal{L}$ est surjective alors $\mathcal{L}^T$ est injective, et si $\mathcal{L}$ est injective alors $\mathcal{L}^T$ est surjective.

Preuve. Soit $\vec{w} \in \mathrm{Ker}(\mathcal{L}^T)$, donc $\mathcal{L}^T.\vec{w} = 0$, donc $(\mathcal{L}^T.\vec{w}, \vec{v})_V = 0$ pour tout $\vec{v} \in V$, donc $(\vec{w}, \mathcal{L}.\vec{v})_W = 0$ pour tout $\vec{v} \in V$, donc $\vec{w} \in (\mathrm{Im}\mathcal{L})^\perp$.

Soit $\vec{w} \in (\mathrm{Im}\mathcal{L})^\perp$, donc $(\vec{w}, \mathcal{L}.\vec{v})_W = 0$ pour tout $\vec{v} \in V$, donc $(\mathcal{L}^T.\vec{w}, \vec{v})_V = 0$ pour tout $\vec{v} \in V$, donc $\mathcal{L}^T.\vec{w} = 0$, donc $\vec{w} \in \mathrm{Ker}(\mathcal{L}^T)$. ∎

2.9.4 Théorème du rang

Théorème 2.30 *(Théorème du rang.) Si $\mathcal{L} : V \to W$ une application linéaire avec* $\dim(V) = n$ *et* $\dim(W) = m$ *:*

$$\dim \mathrm{Ker}(\mathcal{L}) + \dim \mathrm{Im}(\mathcal{L}) = n. \tag{2.37}$$

Preuve. 1- Cas $\mathcal{L} = 0$. Alors $\mathrm{Ker}(\mathcal{L}) = V$ et $\mathrm{Im}(\mathcal{L}) = \{\vec{0}\}$: le théorème est vrai.

2- Cas $\mathcal{L}$ injective (donc $\mathrm{Ker}(\mathcal{L}) = \{\vec{0}\}$). Si $(\vec{v}_i)_{i=1,...,n}$ est une base de V, alors $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ est une famille libre de $\mathrm{Im}(\mathcal{L})$ (car $\mathcal{L}$ injective). Et comme $(\mathcal{L}.\vec{v}_i)_{i=1,...,n}$ est une famille génératrice de $\mathrm{Im}(\mathcal{L})$, on a $\dim(\mathrm{Im}(\mathcal{L})) = n$ avec $\dim(\mathrm{Ker}(\mathcal{L})) = 0$: le théorème est vrai.

3- Autres cas : soit $(\vec{a}_i)_{i=1,...,k}$ une base de $\mathrm{Ker}(\mathcal{L})$ avec $0 < k < n$. On complète la famille libre $(\vec{a}_i)_{i=1,...,k}$ par $n{-}k$ vecteurs $\vec{a}_{k+1}, ..., \vec{a}_n$ tels que $(\vec{a}_i)_{i=1,...,n}$ est une base de V (théorème de la base incomplète).

Et pour tout $\vec{x} = \sum_{i=1}^n x_i\vec{a}_i$ on a $\mathcal{L}.\vec{x} = \vec{0} + \sum_{i=k+1}^n x_i\mathcal{L}.\vec{a}_i$, donc $(\mathcal{L}.\vec{a}_{k+1}, ..., A.\vec{a}_n)$ est une famille génératrice de $\mathrm{Im}(\mathcal{L})$.

Et $(\mathcal{L}.\vec{a}_{k+1}, ..., \mathcal{L}.\vec{a}_n)$ est une famille libre car $\sum_{i=k+1}^n x_i\mathcal{L}.\vec{a}_i = \vec{0} = \mathcal{L}.(\sum_{i=k+1}^n x_i\vec{a}_i)$ donne $\sum_{i=k+1}^n x_i\vec{a}_i \in \mathrm{Ker}(\mathcal{L})$. Donc il existe k réels x_i tels que $\sum_{i=k+1}^n x_i\vec{a}_i = \sum_{i=1}^k x_i\vec{a}_i$. Comme $(\vec{a}_i)$ est une base, on a donc $x_i = 0$ pour tout i : la famille $(\mathcal{L}.\vec{a}_{k+1}, ..., \mathcal{L}.\vec{a}_n)$ est libre.

Donc $\dim(Ker(\mathcal{L})) + \dim(\mathrm{Im}(\mathcal{L})) = k + (n{-}k) = n$: le théorème est vrai. ▪

2.9.5 Bijectivité ou inversibilité

Soit $\mathcal{L} : V \to W$ une application linéaire.

Définition 2.31 $\mathcal{L}$ est dite bijective, ssi $\mathcal{L}$ est injective et surjective.

Exemple 2.32 L'identité $I : V \to V$ est trivialement bijective. ▪

Corollaire 2.33 *Si $\mathcal{L}$ est bijective alors* $\dim(V) = \dim(W)$, *i.e.* $m = n$ *(et nécessairement une matrice représentant $\mathcal{L}$ est une matrice carrée).*

Et $\mathcal{L}$ est alors inversible : $\mathcal{L} : V \to W$ admet une application linéaire inverse $\mathcal{M} : W \to V$ unique telle que $\mathcal{L} \circ \mathcal{M} = I_W$ identité de W, et $\mathcal{M} \circ \mathcal{L} = I_V$ identité de V. Et on note $\mathcal{M} = \mathcal{L}^{-1}$.

Preuve. Si $\mathcal{L}$ est bijective alors $n \geq m$ (surjectivité) et $n \leq m$ (injectivité). D'où $m = n$.

Supposons $\mathcal{L}$ bijective. Donc $\mathcal{L}$ est surjective, donc $\vec{w}_j$ a un antécédent pour tout j : il existe $\vec{b}_j$ t.q. $\mathcal{L}.\vec{b}_j = \vec{w}_j$ pour tout j. Et $(\vec{b}_i)_{i=1,...,n}$ est une base de V car $\mathcal{L}$ est injective.

Définissons alors l'application linéaire $\mathcal{M} : W \to V$ par $\mathcal{M}.\vec{w}_j = \vec{b}_j$. On vérifie immédiatement que $\mathcal{L}(\mathcal{M}(\vec{w}_j) = \mathcal{L}(\vec{b}_j) = \vec{w}_j$, donc $\mathcal{L} \circ \mathcal{M} = I$ (identité dans W), ainsi que $\mathcal{M}(\mathcal{L}(\vec{b}_j) = \mathcal{M}(\vec{w}_j) = \vec{b}_j$ donc $\mathcal{M} \circ \mathcal{L} = I$ (identité dans V), donc $\mathcal{L}$ est inversible d'inverse $\mathcal{M}$. L'unicité est immédiate à vérifier (nécessairement $\mathcal{M}$ doit vérifier $\mathcal{M}.\vec{w}_j = \vec{b}_j$), ou encore on se sert de : l'ensemble des applications linéaires est un anneau (et dans un anneau, si l'inverse existe alors il est unique). ▪

Proposition 2.34 *Soit $A = [\mathcal{L}]_{[(\vec{v}_i),(\vec{w}_i)}$ la matrice représentant $\mathcal{L}$ dans des bases $(\vec{v}_i)$ et $(\vec{w}_i)$.*

$\mathcal{L}$ est bijective ssi A est inversible comme défini en (1.53).

Et alors $[\mathcal{L}^{-1}]_{[(\vec{w}_i),(\vec{v}_i)} = A^{-1}$.

Preuve. Calcul facile. ▪

3 Formes bilinéaires et matrices

3.1 Formes bilinéaires, produits scalaires

3.1.1 Forme bilinéaire

Soit V et W deux espaces vectoriels.

Définition 3.1 Une forme bilinéaire (forme 2-linéaire) $a(\cdot,\cdot) : V \times W \to K$ est une fonction linéaire par rapport à chaque “variable”, i.e. telle que : pour tous $\vec{x}, \vec{x}_1, \vec{x}_2 \in V$, tous $\vec{y}, \vec{y}_1, \vec{y}_2 \in W$, et tous $\alpha, \beta \in K$:

$$\begin{cases} a(\vec{x}_1 + \alpha\vec{x}_2, \vec{y}) = a(\vec{x}_1, \vec{y}) + \alpha a(\vec{x}_2, \vec{y}), \\ a(\vec{x}, \vec{y}_1 + \beta\vec{y}_2) = a(\vec{x}, \vec{y}_1) + \beta a(\vec{x}, \vec{y}_2), \end{cases} \tag{3.1}$$

Proposition 3.2 *Si $a(\cdot,\cdot)$ est bilinéaire sur $V \times W$, alors pour tout $(\vec{x}, \vec{y}) \in V \times W$:*

$$a(\vec{0}, \vec{y}) = a(\vec{x}, \vec{0}) = 0, \tag{3.2}$$

où on a noté $\vec{0}$ les vecteurs nuls $\vec{0}_V$ de V et $\vec{0}_W$ de W.

Preuve. $a(2 \times \vec{0}_V, \vec{y}) = a(\vec{0}_V, \vec{y})$, d'où $2a(\vec{0}_V, \vec{y}) = a(\vec{0}_V, \vec{y})$, d'où $a(\vec{0}_V, \vec{y}) = 0$, idem avec $\vec{0}_W$. ▪

Définition 3.3 Une forme bilinéaire $a(\cdot,\cdot)$ sur $V \times W$ est symétrique ssi $W = V$ et :

$$\forall \vec{x}, \vec{y} \in V, \qquad a(\vec{x}, \vec{y}) = a(\vec{y}, \vec{x}). \tag{3.3}$$

3.1.2 Produit scalaire sur des espaces vectoriels réels

Définition 3.4 Soit V un espace vectoriel réel. Une forme bilinéaire réelle $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ est dite positive (resp. définie positive) ssi :

$$\forall \vec{x} \in V, \quad a(\vec{x}, \vec{x}) \geq 0 \qquad (\text{resp. } \forall \vec{x} \neq \vec{0}, \quad a(\vec{x}, \vec{x}) > 0). \tag{3.4}$$

Définition 3.5 Une forme bilinéaire réelle $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ est dite être un semi-produit scalaire (resp. un produit scalaire) ssi elle est symétrique et positive (resp. symétrique et définie positive). Et on note alors :

$$\forall \vec{x}, \vec{y} \in V, \qquad a(\vec{x}, \vec{y}) = (\vec{x}, \vec{y})_a. \tag{3.5}$$

On note souvent :

$$(\cdot,\cdot)_a = (\cdot,\cdot)_V, \tag{3.6}$$

où donc $(\cdot,\cdot)_V$ est un produit scalaire sur V (une forme bilinéaire symétrique définie positive), notation utilisée quand il n'y a pas d'ambiguïté sur le choix du produit scalaire dans V.

Exemple 3.6 Le produit scalaire usuel $(\cdot,\cdot)_{\mathbb{R}^n}$ est bien un produit scalaire : immédiat. ▪

3.1.3 Forme sesquilinéaire

Soit V un espace vectoriel sur le corps $\mathbb{C}$ des complexes.

Définition 3.7 Une forme sesquilinéaire sur V est une forme $a(\cdot,\cdot) : V \times V \to \mathbb{C}$ qui est :

1- linéaire par rapport à la première variable :

$$\forall \alpha \in \mathbb{C}, \quad \forall \vec{x}, \vec{y}, \vec{z} \in V, \quad a(\vec{x} + \alpha\vec{y}, \vec{z}) = a(\vec{x}, \vec{z}) + \alpha\, a(\vec{y}, \vec{z}), \tag{3.7}$$

2- antilinéaire par rapport à la seconde variable :

$$\forall \beta \in \mathbb{C}, \quad \forall \vec{x}, \vec{y}, \vec{z} \in V, \quad a(\vec{x}, \vec{y} + \beta\vec{z}) = a(\vec{x}, \vec{y}) + \overline{\beta}\, a(\vec{x}, \vec{z}). \tag{3.8}$$

où $\overline{\beta}$ est le complexe conjugué de β.

Remarque 3.8 Attention : suivant les auteurs, une forme sesquilinéaire peut être également définie comme étant antilinéaire par rapport à la première variable et linéaire par rapport à la seconde... ▪

3.1.4 Définition d'un produit scalaire dans $\mathbb{C}^n$

Définition 3.9 Une forme sesquilinéaire sur V est hermitienne ssi :

$$\forall \vec{x}, \vec{y} \in V, \quad a(\vec{x}, \vec{y}) = \overline{a(\vec{y}, \vec{x})}. \tag{3.9}$$

Donc si $a(\cdot, \cdot)$ est sesquilinéaire hermitienne, alors $a(\vec{x}, \vec{x}) \in \mathbb{R}$ pour tout $\vec{x} \in V$.

Définition 3.10 Un semi-produit scalaire (resp. un produit scalaire) sur V est une forme sesquilinéaire hermitienne qui est positive (resp. définie positive) :

$$\forall \vec{x} \in V, \quad a(\vec{x}, \vec{x}) \geq 0, \qquad (\text{resp. } \forall \vec{x} \neq \vec{0}, \quad a(\vec{x}, \vec{x}) > 0) \tag{3.10}$$

et on note alors $a(\cdot, \cdot) \overset{\text{noté}}{=} (\cdot, \cdot)_a$.

Exemple 3.11 Le produit scalaire usuel $(\cdot, \cdot)_{\mathbb{C}^n}$ est bien un produit scalaire : immédiat. ▪

3.1.5 Inégalité de Cauchy-Schwarz et norme associée à un produit scalaire

Définition 3.12 La norme (resp. semi-norme) associée au produit scalaire (resp. semi-produit scalaire) $(\cdot, \cdot)_a$ est la fonction $||.||_a : V \to \mathbb{R}_+$ définie par :

$$||\vec{x}||_a = \sqrt{(\vec{x}, \vec{x})_a}. \tag{3.11}$$

Comme un (semi-)produit scalaire est positif, $||\vec{x}||_a$ est bien un réel positif.

Théorème 3.13 **(Inégalité de Cauchy–Schwarz.)** *Soit $a(\cdot, \cdot)$ un semi-produit scalaire sur V. Alors, pour tout $\vec{x}, \vec{y} \in V$:*

$$|(\vec{x}, \vec{y})_a| \leq ||\vec{x}||_a ||\vec{y}||_a. \tag{3.12}$$

Preuve. 1- Cas $K = \mathbb{R}$: soit $\vec{x}, \vec{y} \in V$ et soit le polynôme donné par, pour tout $t \in \mathbb{R}$:

$$P(t) = (\vec{x} - t\vec{y}, \vec{x} - t\vec{y})_a = ||\vec{y}||_a^2 t^2 - 2(\vec{x}, \vec{y})_a t + ||\vec{x}||_a^2. \tag{3.13}$$

On a $P(t) \geq 0$ pour tout t car un semi-produit scalaire est une forme bilinéaire positive. Donc son discriminant est ≤ 0, soit $4(\vec{x}, \vec{y})_a^2 - 4||\vec{x}||_a^2\, ||\vec{y}||_a^2 \leq 0$, i.e. (3.12).

2- Cas $K = \mathbb{C}$: soit $\vec{x}, \vec{y} \in V$ et, soit $\theta \in \mathbb{R}$ et soit le polynôme donné par, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} P(t) &= (\vec{x} - te^{-i\theta}\vec{y}, \vec{x} - te^{-i\theta}\vec{y})_a = ||\vec{x}||_a^2 - te^{-i\theta}(\vec{y}, \vec{x})_a - te^{+i\theta}(\vec{x}, \vec{y})_a + ||te^{-i\theta}\vec{y}||_a^2 \\ &= ||\vec{y}||_a^2 t^2 - (e^{+i\theta}(\vec{x}, \vec{y})_a + \overline{e^{+i\theta}(\vec{x}, \vec{y})_a})\, t + ||\vec{x}||_a^2, \end{aligned} \tag{3.14}$$

car $(\cdot, \cdot)_a$ est hermitien, donc . Soit $-\theta$ l'argument du complexe $(\vec{x}, \vec{y})_a$, i.e. $(\vec{x}, \vec{y})_a = |(\vec{x}, \vec{y})_a| e^{-i\theta}$. Donc $e^{+i\theta}(\vec{x}, \vec{y})_a = |(\vec{x}, \vec{y})_a| = \overline{e^{+i\theta}(\vec{x}, \vec{y})_a} \in \mathbb{R}$, donc $P(t) = ||\vec{y}||_a^2 t^2 - 2|(\vec{x}, \vec{y})_a|\, t + ||\vec{x}||_a^2$.

On a $P(t) \geq 0$ pour tout t car un semi-produit scalaire est positif. Donc son discriminant est ≤ 0, soit $4|(\vec{x}, \vec{y})_a|^2 - 4||\vec{x}||_a^2 ||\vec{y}||_a^2 \leq 0$, i.e. (3.12). ▪

Proposition 3.14 *Si $a(\cdot, \cdot)$ est un produit scalaire (resp. un semi-produit scalaire) sur V alors $||.||_a$ est bien une norme (resp. une semi-norme) sur V.*

Preuve. $||.||_a$ est positive (car $a(\cdot, \cdot)$ est positive).

On a $||\alpha\vec{x}||_a^2 = a(\alpha\vec{x}, \alpha\vec{x}) = \alpha\overline{\alpha} a(\vec{x}, \vec{x}) = |\alpha|^2 ||\vec{x}||_a^2$, d'où l'homogénéité.

On a $||\vec{x} + \vec{y}||_a^2 = (\vec{x} + \vec{y}, \vec{x} + \vec{y})_V = ||\vec{x}||_a^2 + 2(\vec{x}, \vec{y})_a + ||\vec{y}||_a^2 \leq ||\vec{x}||_a^2 + 2||\vec{x}||_a ||\vec{y}||_a + ||\vec{y}||_a^2 \leq (||\vec{x}||_a + ||\vec{y}||_a)^2$, grâce à Cauchy–Schwarz, d'où l'inégalité triangulaire. ▪

Corollaire 3.15 *Si $a(\cdot, \cdot)$ est un produit scalaire sur V, alors pour $\vec{x}_1, \vec{x}_2 \in V$ on a :*

$$\vec{x}_1 = \vec{x}_2 \quad \Longleftrightarrow \quad \forall \vec{y} \in V, \quad a(\vec{x}_1, \vec{y}) = a(\vec{x}_2, \vec{y}). \tag{3.15}$$

Preuve. $\Rightarrow$: tautologie.

$\Leftarrow$: le membre de gauche équivaut à $a(\vec{x}_1 - \vec{x}_2, \vec{y}) = 0$ pour tout $\vec{y}$, en particulier pour $\vec{y} = \vec{x}_1 - \vec{x}_2$, donc $||\vec{x}_1 - \vec{x}_2||_a = 0$, donc $\vec{x}_1 = \vec{x}_2$. (Faux si $a(\cdot, \cdot)$ est uniquement un semi-produit scalaire.) ▪

3.1.6 Définition d'une b.o.n. relativement à un produit scalaire

Définition 3.16 Si $(\cdot,\cdot)_V$ est un produit scalaire sur V et si $(\vec{e}_i)$ est une base dans V, alors $(\vec{e}_i)$ est une base orthonormée (une b.o.n.) relativement au produit scalaire $(\cdot,\cdot)_V$ ssi, pour tout i,j :

$$(\vec{e}_i,\vec{e}_j)_V=\delta_{ij}. \tag{3.16}$$

Proposition 3.17 *Un espace vectoriel V de dimension finie muni d'un produit scalaire $(\cdot,\cdot)_V$ admet une base orthonormée.*

Preuve. On part d'une base $(\vec{a}_i)$ de V, et on construit alors une b.o.n. $(\vec{e}_i)$ (pour le produit scalaire $(\cdot,\cdot)_V$) à l'aide du procédé de Gram–Schmidt par exemple, cf. (1.76). ▪

3.1.7 * Continuité d'une forme bilinéaire ou sesquilinéaire

On considère deux espaces vectoriels normés $(V,||.||_V)$ et $(W,||.||_W)$.

Définition 3.18 Une forme bilinéaire ou sesquilinéaire $a(\cdot,\cdot):V\times W\rightarrow K$ est dite continue ssi elle est continue par rapport à chaque variable (en tant que forme linéaire) :

$$\exists c>0,\quad \forall(\vec{x},\vec{y})\in V\times W,\quad \text{on a}\quad |a(\vec{x},\vec{y})|\le c\,||\vec{x}||_V||\vec{y}||_W. \tag{3.17}$$

On munit le produit cartésien $Z=V\times W$ de la norme $||.||_Z=||.||_{V\times W}$ définie par :

$$||\vec{z}||_Z=\sup(||\vec{x}||_V,||\vec{y}||_W)\quad\text{quand}\quad \vec{z}=(\vec{x},\vec{y}). \tag{3.18}$$

Soit $f:Z\rightarrow\mathbb{R}$ définie par :

$$f(\vec{z})=a(\vec{x},\vec{y})\quad\text{quand}\quad \vec{z}=(\vec{x},\vec{y})\in Z. \tag{3.19}$$

f n'est pas linéaire.

Proposition 3.19 *Quand $a(\cdot,\cdot)$ est continue (au sens (3.17)), alors l'application $f:Z\rightarrow\mathbb{R}$ est continue sur $(Z,||.||_Z)$. (Donc f est toujours continue en dimension finie.)*

N.B. : penser à la forme bilinéaire = le produit simple dans $\mathbb{R}$ donné par $(x,y)\in\mathbb{R}\times\mathbb{R}\rightarrow xy\in\mathbb{R}$, associée la fonction $f:(x,y)\in\mathbb{R}^2\rightarrow xy\in\mathbb{R}$ dont le graphe $G(f)=\{(x,y,f(x,y))\in\mathbb{R}^3\}$ est la "selle de cheval" : $f(x,y)=0$ si $x=0$ ou $y=0$, $f(x,x)=x^2$ (parabole "vers le haut") et $f(x,-x)=-x^2$ (parabole "vers le bas").

Preuve. Montrons la continuité en un point $\vec{z}_0=(\vec{x}_0,\vec{y}_0)\in Z$. Soit $\vec{z}=(\vec{x},\vec{y})\in Z$. On a :

$$\begin{aligned}|f(\vec{z})-f(\vec{z}_0)|&=|a(\vec{x},\vec{y})-a(\vec{x}_0,\vec{y}_0)|=|a(\vec{x}-\vec{x}_0,\vec{y})+a(\vec{x}_0,\vec{y}-\vec{y}_0)|\\&=|a(\vec{x}-\vec{x}_0,\vec{y}-\vec{y}_0)+a(\vec{x}-\vec{x}_0,\vec{y}_0)+a(\vec{x}_0,\vec{y}-\vec{y}_0)|\\&\le c\,||\vec{x}-\vec{x}_0||_V||\vec{y}-\vec{y}_0||_W+c\,||\vec{x}-\vec{x}_0||_V||\vec{y}_0||_W+c\,||\vec{x}_0||_V||\vec{y}-\vec{y}_0||_W\\&\le c\sup(1,||\vec{y}_0||_W,||\vec{x}_0||_V)\sup(||\vec{x}-\vec{x}_0||_V,||\vec{y}-\vec{y}_0||_W)\\&\le c\sup(1,||\vec{y}_0||_W,||\vec{x}_0||_V)||z-z_0||_Z,\end{aligned}$$

dès que $||\vec{x}-\vec{x}_0||_V\le 1$ et $||\vec{y}-\vec{y}_0||_W\le 1$. Donc quand $\vec{z}\rightarrow\vec{z}_0$, i.e. quand $||\vec{z}-\vec{z}_0||_Z\rightarrow 0$, on a $|f(\vec{z})-f(\vec{z}_0)|\rightarrow 0$, et donc f est continue en $\vec{z}_0$ dans $(Z,||.||_Z)$. Vrai pour tout $\vec{z}_0$. ▪

3.2 Matrices d'une forme bilinéaire

3.2.1 Matrice d'une forme bilinéaire ou sesquilinéaire

Soit V et W deux espaces vectoriels sur K de dimensions respectives m et n. Selon le cas $K=\mathbb{R}$ ou $K=\mathbb{C}$, $a(\cdot,\cdot)$ désignera soit une forme bilinéaire $V\times W\rightarrow\mathbb{R}$ soit une forme sesquilinéaire $V\times W\rightarrow\mathbb{C}$.

Soit $(\vec{v}_i)_{i=1,\dots,m}$ une base de V et soit $(\vec{w}_i)_{i=1,\dots,n}$ une base de W.

Définition 3.20 La matrice $m*n$ (réelle ou complexe suivant le cas) :

$$[a(\cdot,\cdot)]_{|(\vec{v}),(\vec{w})} = A = [A_{ij}] \qquad \text{où} \qquad A_{ij} = a(\vec{v}_i, \vec{w}_j) \quad \forall i,j, \tag{3.20}$$

est appelée la matrice de la forme bilinéaire $a(\cdot,\cdot)$ relativement aux bases $(\vec{v}_i)$ et $(\vec{w}_i)$.

Proposition 3.21 *Soit* $A = [a(\cdot,\cdot)]_{|(\vec{v}),(\vec{w})}$. *Pour* $\vec{x} = \sum_{i=1}^n x_i\vec{v}_i \in V$ *et* $\vec{y} = \sum_{j=1}^n y_j\vec{w}_j \in W$, *on a :*

$$a(\vec{x},\vec{y}) = \sum_{i,j=1}^n x_i A_{ij} \overline{y}_j = [\vec{x}]^T_{|(\vec{v})}.A.[\overline{\vec{y}}]_{|(\vec{w})}. \tag{3.21}$$

(La conjugaison est inutile si $K = \mathbb{R}$.)

Preuve. $a(\vec{x},\vec{y}) = a(\sum_{i=1}^n x_i\vec{v}_i, \sum_{j=1}^n y_j\vec{w}_j) = \sum_{ij} x_i\overline{y}_j a(\vec{v}_i,\vec{w}_j)$. ∎

Exemple 3.22 La matrice du produit scalaire usuel de K^n relativement à la base canonique de K^n est la matrice identité : $[(\cdot,\cdot)_{K^n}] = I$. En effet :

$$(\vec{E}_i^{(n)}, \vec{E}_j^{(n)})_{K^n} = \delta_{ij}, \qquad \forall i,j = 1,...,n. \tag{3.22}$$

∎

3.2.2 Matrice réelle définie positive

Définition 3.23 Une matrice réelle $A \in \mathcal{M}_{nn}(\mathbb{R})$ est positive (resp. définie positive) ssi :

$$\forall \vec{x} \in \mathbb{R}^n, \quad [\vec{x}]^T.A.[\vec{x}] \geq 0 \qquad (\text{resp. } \forall \vec{x} \neq \vec{0}, \quad [\vec{x}]^T.A.[\vec{x}] > 0), \tag{3.23}$$

où $[\vec{x}]$ est la matrice colonne des composantes de $\vec{x}$ dans la base canonique.

Autrement dit, dans $\mathbb{R}^n$ muni de sa base canonique $(\vec{E}_i)$ et de son produit scalaire usuel $(\cdot,\cdot)_{\mathbb{R}^n}$, une matrice A $n*n$ est positive (resp. définie positive) ssi :

$$\forall \vec{x} \in \mathbb{R}^n, \quad (A.\vec{x},\vec{x})_{\mathbb{R}^n} \geq 0 \qquad (\text{resp. } \forall \vec{x} \neq \vec{0}, \quad (A.\vec{x},\vec{x})_{\mathbb{R}^n} > 0), \tag{3.24}$$

où il est sous-entendu que $\vec{x}$ est décomposé sur la base canonique.

Proposition 3.24 *Une matrice A $n*n$ définie positive est inversible.*

Preuve. Si A n'est pas inversible alors $\det(A) = 0$ et les "vecteurs colonnes $\vec{a}_j$" de A ne sont pas indépendants, donc il existe une combinaison linéaire $\sum_j x_j\vec{a}_j = \vec{0}$, où les x_j ne sont pas tous nuls. Notant $[\vec{x}]$ la matrice colonne des x_i (non nulle), on a donc $A.[\vec{x}] = 0$, donc $[\vec{x}]^T.A.[\vec{x}] = 0$ pour $\vec{x} \neq \vec{0}$, donc A n'est pas définie positive. Donc si A est définie positive, elle est inversible. ∎

Exercice 3.25 Montrer que $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est définie positive. Et que $B = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ ne l'est pas.

Réponse. $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = x^2+xy+y^2 = \frac{1}{2}(x^2+2xy+y^2)+\frac{1}{2}(x^2+y^2) \geq \frac{1}{2}(x+y)^2+\frac{1}{2}(x^2+y^2) \geq \frac{1}{2}(x^2+y^2) > 0$ pour tout $\vec{x} = (x,y) \neq \vec{0}$.

$(B.\vec{x},\vec{x})_{\mathbb{R}^2} = x^2 + 3xy + y^2$, donc pour $x = -y = 1$ on a $(B.\vec{x},\vec{x})_{\mathbb{R}^2} = -1$. ∎

Proposition 3.26 *Si A est définie positive, alors il existe $\alpha > 0$ t.q., pour tout $\vec{x} \in \mathbb{R}^n$:*

$$(A.\vec{x},\vec{x})_{\mathbb{R}^n} \geq \alpha ||\vec{x}||^2_{\mathbb{R}^n}, \tag{3.25}$$

et, pour le plus grand α possible, l'égalité à lieu pour un $\vec{x}$ (au moins).

(Et on verra, exercice 5.6, que si de plus A est symétrique, alors le plus grand α possible est la plus petite valeur propre de A.)

Preuve. (3.25) équivaut à $(A.\vec{x},\vec{x})_{\mathbb{R}^n} \geq \alpha$ pour tout $\vec{x} \in S$, où S est la sphère unité de $\mathbb{R}^n$. La forme quadratique f donnée par $f(\vec{x}) = (A.\vec{x},\vec{x})_{\mathbb{R}^n} = \sum_{ij} A_{ij}x_jx_i$ est continue sur $\mathbb{R}^n$ (c'est un polynôme), donc sur S. Comme S est compact, f atteint son minimum sur S : il existe $\vec{x}_{\min} \in S$ t.q. $f(\vec{x}) \geq f(\vec{x}_{\min})$ pour tout $\vec{x} \in S$. Et comme A est définie positive et $\vec{x}_{\min} \neq \vec{0}$ (car $\vec{x}_{\min} \in S$), on a $f(\vec{x}_{\min}) > 0$ (car A est définie positive). On pose $\alpha = f(\vec{x}_{\min})$ et α vérifie (3.25), et de plus cet α est optimal : c'est le plus grand α possible, et il est atteint pour $\vec{x}_{\min}$. ∎

Exercice 3.27 Quel est le plus grand α pour la matrice A de l'exercice 3.25, et pour quels $\vec{x}$ a-t-on l'égalité dans (3.25) ?

Réponse. Pour la matrice de l'exercice 3.25, la meilleure constante possible (la plus grande) est $\alpha = \frac{1}{2}$: on a $(A.\vec{x},\vec{x})_{\mathbb{R}^2} \geq \frac{1}{2}(x^2+y^2) = \frac{1}{2}||\vec{x}||^2_{\mathbb{R}^n}$ pour tout $\vec{x}$, donc $\alpha = \frac{1}{2}$ est possible ; puis prendre $x = -y$ qui donne $(A.\vec{x},\vec{x})_{\mathbb{R}^n} = \frac{1}{2}||\vec{x}||^2_{\mathbb{R}^n}$, donc nécessairement $\alpha \leq \frac{1}{2}$. Donc la plus grande valeur possible de α est $\frac{1}{2}$ et elle est atteinte pour $\vec{x} = (-1,1)$ par exemple, et plus généralement pour tout $\vec{x}$ sur la droite $x+y=0$. ∎

Exercice 3.28 Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice $2*2$ définie positive. Montrer $a,d>0$ et $ad-bc>0$ (déterminant).

Et montrer que la réciproque est fausse en général, et vraie pour les matrices symétriques.

Montrer que si A est symétrique, alors A est définie positive ssi $\mathrm{Tr}(A) > 0$ et $\det(A) > 0$.

Réponse. Soit $\vec{x} \overset{\text{noté}}{=} \begin{pmatrix} x \\ y \end{pmatrix}$. On a $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = ax^2 + (b+c)xy + dy^2 > 0$ pour tout $\vec{x} \neq \vec{0}$.

$\Rightarrow$: supposons A définie positive. En particulier si $x=1$ et $y=0$ on a $0 < (A.\vec{x},\vec{x})_{\mathbb{R}^2} = a$: on a bien $a>0$. De même avec $x=0$ et $y=1$, on a $d>0$.

Puis avec $y \neq 0$ et en posant $z = \frac{x}{y}$ on a $0 < (A.\vec{x},\vec{x})_{\mathbb{R}^2} = y^2(az^2+(b+c)z+d)$ pour tout z. Comme A est définie positive, le polynôme p donné par $p(z) = az^2+(b+c)z+d$ n'a pas de racine réelle : sinon il existe z t.q. $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = 0$ pour $y \neq 0$, impossible car A est définie positive. Donc $(b+c)^2 - 4ad < 0$.

Si de plus $ad - bc \leq 0$, alors $4ad \leq 4bc$ avec $(b+c)^2 < 4ad$. Donc $(b+c)^2 < 4bc$, donc $(b-c)^2 < 0$, absurde. Donc nécessairement $ad - bc > 0$.

$\Leftarrow$: soit $A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$. On a $a = d = 1 > 0$ et $ad - bc = 1 > 0$, mais pour $y = -x = 1$ on a $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = -1 < 0$.

Soit A symétrique (donc $b = c$) avec $a,d>0$ et $ad - b^2 > 0$. Alors pour $y=0$ on a $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = ax^2$ donc >0 pour $\vec{x} \neq \vec{0}$, et pour $y \neq 0$ on a $(A.\vec{x},\vec{x})_{\mathbb{R}^2} = y^2(az^2+2bz+d)$ avec le discriminant du polynôme en z qui vaut $(2b)^2 - 4ad < 0$; donc le polynôme en z n'a pas de racine, donc A est définie positive.

Autre démonstration avec la diagonalisation des matrices symétriques : voir exercice 5.6.

Soit $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$, matrice symétrique. Supposons $a+d>0$ et $ad-b^2>0$. Alors $ad > b^2$, donc a et d sont non nuls de même signe, et $a+d>0$ donne $a>0$ et $d>0$. ∎

Exercice 3.29 Montrer que si $P \in \mathbb{R}^{n^2}$ est une matrice inversible, alors $P^T.P$ est une matrice symétrique définie positive.

Réponse. Symétrie : $(P^T.P)^T = P^T.(P^T)^T = P^T.P$.

Définie positive : $(P^T.P.\vec{x},\vec{x})_{\mathbb{R}^n} = (P.\vec{x},P.\vec{x})_{\mathbb{R}^n} = ||P.\vec{x}||^2_{\mathbb{R}^n} > 0$ pour $\vec{x} \neq \vec{0}$ car P est inversible. ∎

3.2.3 Matrice de masse

Définition 3.30 Soit $(V,(\cdot,\cdot)_V)$ un espace vectoriel réel de dimension finie n muni d'un produit scalaire $(\cdot,\cdot)_V$. Une matrice $M = [M_{ij}] \in \mathbb{R}^{n^2}$ est appelée matrice de masse ssi il existe une base $(\vec{v}_i)_{i=1,\dots,n}$ de V telle que :

$$M_{ij} = (\vec{v}_i,\vec{v}_j)_V, \qquad \forall i,j = 1,...,n. \tag{3.26}$$

M est alors appelée la matrice de masse relativement à la base $(\vec{v}_i)$ et au produit scalaire $(\cdot,\cdot)_V$.

Proposition 3.31 *Une matrice de masse $M = [(\vec{v}_i,\vec{v}_j)_V]$ est une matrice symétrique définie positive.*

Et réciproquement, une matrice M symétrique définie positive est une matrice de masse dans V.

Preuve. Soit $M = [(\vec{v}_i, \vec{v}_j)_V]$ une matrice de masse. Un produit scalaire étant symétrique, on a immédiagement $M_{ij} = M_{ji}$ pour tout i, j, donc M est symétrique. Soit $\vec{x} = \sum_{i=1}^n x_i \vec{v}_i \neq 0$ et $[\vec{x}] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. On a pour tout $\vec{x} \in V$ non nul :

$$0 < ||\vec{x}||_V^2 = (\vec{x}, \vec{x})_V = \sum_{i,j=1}^n x_i x_j (\vec{v}_i, \vec{v}_j)_V = \sum_{i,j=1}^n x_i M_{ij} x_j = [\vec{x}]^T.M.[\vec{x}],$$

et donc M est définie positive.

Réciproque. Soit M symétrique réelle définie positive. Soit $M = L.L^T$ sa décomposition de Cholesky, cf. théorème 1.80. Donc L^T est inversible (car $0 < \det(M) = \det(L)\det(L^T)$). Soit $(\vec{v}_i)$ une b.o.n. dans $(V, (\cdot,\cdot)_V)$ et soit $\vec{w}_j = \sum_{i=1}^n (L^T)_{ij} \vec{v}_j$, où donc L^T est la matrice de passage de la base $(\vec{v}_i)$ à la base $(\vec{w}_i)$. On a $(\vec{w}_i, \vec{w}_j)_V = \sum_{k\ell} (L^T)_{ki} (L^T)_{\ell j} (\vec{v}_k, \vec{v}_\ell)_V = \sum_{k\ell} L_{ik} (L^T)_{\ell j} \delta_{k\ell} = \sum_k L_{ik} (L^T)_{kj} = (L.L^T)_{ij} = M_{ij}$, donc $M = [(\vec{w}_i, \vec{w}_j)_V]$ est une matrice de masse dans V. ∎

Proposition 3.32 *Soit $M = [(\vec{v}_i, \vec{v}_j)_V]$ une matrice de masse sur V. Alors M est également une matrice de masse sur $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$.*

(Ainsi l'étude des matrices de masse sur un espace vectoriel réel est ramené à l'étude des matrices de masse définies sur $\mathbb{R}^n$.)

Preuve. M est réelle, symétrique et définie positive, donc $M = L.L^T$ (factorisation de Cholesky) où L est triangulaire inférieure définie positive. Soit $\vec{w}_i = L^T.\vec{E}_j \in \mathbb{R}^n$ où $(\vec{E}_i)$ est la base canonique. On a $(\vec{w}_i, \vec{w}_j)_{\mathbb{R}^n} = \sum_{k\ell} (L^T)_{ki} (L^T)_{\ell j} (\vec{E}_k, \vec{E}_\ell)_{\mathbb{R}^n} = \sum_{k\ell} L_{ik} (L^T)_{\ell j} \delta_{k\ell} = (L.L^T)_{ij} = M_{ij}$, donc M est la matrice de masse dans $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$ relativement à la base $(\vec{w}_i)$. ∎

3.2.4 Matrice de masse et produit scalaire dans $\mathbb{R}^n$

Soit $(\vec{E}_i)$ la base canonique de $\mathbb{R}^n$. On munit $\mathbb{R}^n$ de son produit scalaire usuel $(\cdot,\cdot)_{\mathbb{R}^n}$.

Proposition 3.33 *$a(\cdot,\cdot)$ est un produit scalaire dans $\mathbb{R}^n$ ssi il existe une matrice $M = [M_{ij}]$ symétrique définie positive telle que :*

$$a(\vec{E}_i, \vec{E}_j) = M_{ij}. \tag{3.27}$$

et on note alors :

$$a(\vec{x}, \vec{y}) = (\vec{x}, \vec{y})_M \qquad (= (M.\vec{x}, \vec{y})_{\mathbb{R}^n}). \tag{3.28}$$

Preuve. Soit $a(\cdot,\cdot)$ un produit scalaire. Alors $M = [a(\vec{E}_i, \vec{E}_j)]$ est une matrice de masse, donc symétrique définie positive, et $a(\vec{x}, \vec{y}) = \sum_{i,j=1}^n x_i y_j M_{ij} = (M.\vec{x}, \vec{y})_{\mathbb{R}^n}$.

Réciproquement, si M est symétrique définie positive, on définit la forme bilinéaire $a(\cdot,\cdot)$ par $a(\vec{E}_i, \vec{E}_j) = M_{ij}$ (donc $a(\vec{x}, \vec{y}) = \sum_{i,j=1}^n x_i y_j M_{ij} = (M.\vec{x}, \vec{y})_{\mathbb{R}^n}$ pour tout $\vec{x}, \vec{y}$). Et $a(\cdot,\cdot)$ ainsi définie est symétrique (car M l'est) et définie positive : $a(\vec{x}, \vec{x}) = (M.\vec{x}, \vec{x})_{\mathbb{R}^n} = [\vec{x}]^T.M.[\vec{x}] > 0$ pour $\vec{x} \neq \vec{0}$ car M est définie positive. ∎

Proposition 3.34 *Soit $(\vec{v}_i)_{i=1,\ldots,n}$ une base de $\mathbb{R}^n$. Alors il existe un unique produit scalaire $a(\cdot,\cdot)$ dans $\mathbb{R}^n$ pour lequel $(\vec{v}_i)$ est une base $(\cdot,\cdot)_M$-orthonormée, à savoir $a(\cdot,\cdot) = (M.,.)_{\mathbb{R}^n} \overset{\text{noté}}{=} (\cdot,\cdot)_M$ où :*

$$M = P^{-T}.P^{-1}, \tag{3.29}$$

où $P = [P_{ij}] = (\, [\vec{v}_1] \quad \ldots \quad [\vec{v}_n] \,)$ est la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{v}_i)$.

Preuve. Soit $P = [P_{ij}] = (\, [\vec{v}_1] \quad \ldots \quad [\vec{v}_n] \,)$ la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{v}_i)$, donc $P.\vec{E}_j = \vec{v}_j$ pour tout j. On veut trouver une matrice M t.q., pour tout i, j :

$$(\vec{E}_i, \vec{E}_j)_{\mathbb{R}^n} = \delta_{ij} = (\vec{v}_i, \vec{v}_j)_M = (M.\vec{v}_i, \vec{v}_j)_{\mathbb{R}^n} = (M.P.\vec{E}_i, P.\vec{E}_j)_{\mathbb{R}^n} = (P^T.M.P.\vec{E}_i, \vec{E}_j)_{\mathbb{R}^n}.$$

Donc $P^T.M.P.\vec{E}_i = \vec{E}_i$ pour tout i, donc $P^T.M.P = I$, soit $M = P^{-T}.P^{-1}$.

Vérifions que $M = P^{-T}.P^{-1}$ convient : $(\vec{v}_i, \vec{v}_j)_M = (M.\vec{v}_i, \vec{v}_j)_{\mathbb{R}^n} = (P^{-T}.P^{-1}.\vec{v}_i, \vec{v}_j)_{\mathbb{R}^n} = (P^{-1}.\vec{v}_i, P^{-1}.\vec{v}_j)_{\mathbb{R}^n} = (\vec{E}_i, \vec{E}_j)_{\mathbb{R}^n} = \delta_{ij}$ pour tout i, j : OK. ∎

Proposition 3.35 *Si M est une matrice de masse, si $(\cdot,\cdot)_M$ est le produit scalaire associé et $||.||_M$ la norme associée (définie par $||\vec{x}||_M = \sqrt{(M.\vec{x},\vec{x})_{\mathbb{R}^n}}$), alors :*

$$||\vec{x}||_{\mathbb{R}^n} = 0 \quad\iff\quad ||\vec{x}||_M = 0 \quad\iff\quad \vec{x} = \vec{0}, \tag{3.30}$$

ou encore :

$$||\vec{x}||_{\mathbb{R}^n} > 0 \quad\iff\quad ||\vec{x}||_M > 0 \quad\iff\quad \vec{x} \neq \vec{0}. \tag{3.31}$$

Preuve. Un produit scalaire est défini positif. ∎

Il est immédiat que si $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ (matrice diagonale), alors :

$$\forall i = 1, ..., n, \quad \lambda_i > 0 \quad\iff\quad \forall \vec{x} \in \mathbb{R}^n - \{\vec{0}\}, \quad (D.\vec{x},\vec{x})_{\mathbb{R}^n} > 0. \tag{3.32}$$

(On a $(D.\vec{x},\vec{x})_{\mathbb{R}^n} = \sum_i \lambda_i x_i^2$.)

En revanche :

Proposition 3.36 *Il existe des matrices diagonales $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ telles que $\lambda_i > 0$ pour tout i (donc définies positives), et il existe des produits scalaires $(\cdot,\cdot)_M$ (associés à de matrices de masse M) telles que $(D.\vec{x},\vec{x})_M < 0$ (bien que $(D.\vec{x},\vec{x})_{\mathbb{R}^n} > 0$ pour tout $\vec{x} \neq \vec{0}$).*

Donc dans (3.24), on ne peut pas remplacer le produit scalaire usuel par un produit scalaire quelconque pour caractériser une matrice définie positive.

Preuve. On prend $D = \mathrm{diag}(1,2) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, donc $\lambda_1 = 1$ et $\lambda_2 = 2$, et $M = \begin{pmatrix} 1 & 1 \\ 1 & \frac{17}{16} \end{pmatrix}$. La matrice M est bien symétrique définie positive (donc matrice d'un produit scalaire). On a $M.D = \begin{pmatrix} 1 & 2 \\ 1 & \frac{17}{8} \end{pmatrix}$. Et on prend $\vec{x} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$ qui donne $(D.\vec{x},\vec{x})_M = (M.D.\vec{x},\vec{x})_{\mathbb{R}^2} = -\frac{1}{2} < 0$, bien que $(D.\vec{x},\vec{x})_{\mathbb{R}^n} = \lambda_1 x_1^2 + \lambda_2 x_2^2 > 0$ car $\lambda_1 = 1$ et $\lambda_2 = 2$ sont > 0.

Remarque : bien que $M.D$ n'est pas définie positive, ses valeurs propres sont tout de même positives : la trace $\lambda_1 + \lambda_2$ et le déterminant $\lambda_1\lambda_2$ sont positifs. Cela donne un exemple de matrice $A = M.D$ qui est non définie positive dont les valeurs propres sont strictement positives ($A = M.D$ est non symétrique). ∎

Exercice 3.37 Faire les calculs ayant permis de trouver le contre-exemple dans la démonstration précédente.

Réponse. On prend $D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, et on cherche $M = \begin{pmatrix} a & 1 \\ 1 & c \end{pmatrix}$ et $\vec{x}$ tels que $(M.D.\vec{x},\vec{x})_{\mathbb{R}^2} < 0$. Comme M symétrique doit être définie positive, nécessairement $a, c > 0$ et $ac - 1 > 0$, cf. exercice 3.28. Puis $M.D = \begin{pmatrix} a & 2 \\ 1 & 2c \end{pmatrix}$ donne $(M.D.\vec{x},\vec{x})_{\mathbb{R}^2} = ax^2 + 3xy + 2cy^2 = y^2(az^2 + 3z + 2c)$ où on a posé $z = \frac{x}{y}$ en supposant $y \neq 0$. Et on veut donc que le polynôme $p(z) = az^2 + 3z + 2c$ ait des racines réelles. Son discriminant est $\Delta = 9 - 8ac$ doit donc être positif, donc $8ac < 9$; avec $ac > 1$ vu précédemment. Prenons $a = 1$ et $c = \frac{17}{16}$. Puis prenons $z = -\frac{3}{2a} = \frac{-3}{2}$ (la demi somme des racines) qui donne $p(\frac{-3}{2}) = -\frac{1}{8}$. On prend $x = -3$ et $y = 2$, ce qui donne : $(M.D.\vec{x},\vec{x})_{\mathbb{R}^2} = -\frac{1}{2}$. ∎

3.3 Changement de base pour une forme bilinéaire : $B = P^T.A.P$

Soit $a(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ une forme bilinéaire. Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases de $\mathbb{R}^n$. Soit, cf.(3.20) :

$$A = [a(\cdot,\cdot)]_{|(\vec{a})}, \qquad B = [a(\cdot,\cdot)]_{|(\vec{b})}, \tag{3.33}$$

les matrices de $a(\cdot,\cdot)$ respectivement dans les bases $(\vec{a}_i)$ est $(\vec{b}_i)$. Soit :

$$P = \begin{pmatrix} [\vec{b}_1]_{|(\vec{a})} & \dots & [\vec{b}_n]_{|(\vec{a})} \end{pmatrix} \tag{3.34}$$

la matrice de changement de base de la base $(\vec{a}_i)$ vers la base $(\vec{b}_i)$.

Proposition 3.38 *On a :*

$$[a(\cdot,\cdot)]_{|(\vec{b})} = P^T.[a(\cdot,\cdot)]_{|(\vec{a})}.P, \qquad \text{soit} \qquad B = P^T.A.P. \tag{3.35}$$

Preuve. $B_{ij} = a(\vec{b}_i, \vec{b}_j) = \sum_{k,\ell=1}^n P_{ki} P_{\ell j} a(\vec{a}_k, \vec{a}_\ell) = \sum_{k,\ell=1}^n (P^T)_{ik} A_{k\ell} P_{\ell j} = (P^T.A.P)_{ij}$. ▪

Exemple 3.39 On reprend les unités de mesure pied et mètre donnés à l'exemple 2.12.

Soit $a(\cdot,\cdot) : (\vec{x}, \vec{y}) \in \mathbb{R} \times \mathbb{R} \to a(\vec{x}, \vec{y}) \in \mathbb{R}$ une forme bilinéaire. Soit $\vec{x}$ donné. Soit x_a (resp. x_b) sa coordonnée en mètre (resp. en pied). Donc $\vec{x} = x_a \vec{a} = x_b \vec{b}$. Soit $\vec{y}$ un second vecteur donné. Soit y_a (resp. y_b) sa coordonnée en mètre (resp. en pied). Donc $\vec{y} = y_a \vec{a} = y_b \vec{b}$.

Formule de changement de base (3.35), ici pour des matrices $1 * 1$ identifiées à des réels : $B = P^T.A.P = P^2.A$, donc $B = (0,3048)^2.A$.

Vérification : $a(\vec{x}, \vec{y}) = [\vec{x}]_{|(\vec{a})}^T.A.[\vec{y}]_{|(\vec{a})} = x_a y_a A$, ainsi que $a(\vec{x}, \vec{y}) = [\vec{x}]_{|(\vec{b})}^T.B.[\vec{y}]_{|(\vec{b})} = x_b y_b B$. Comme $x_b = P^{-1} x_a$, $y_b = P^{-1} y_a$ et $B = P^2.A$, on a bien $x_b y_b B = x_a y_a A = a(\vec{x}, \vec{y})$. ▪

Remarque 3.40 Comparaison avec l'exemple 2.12. La formule de changement de base pour les endomorphismes, $B = P^{-1}.A.P$, est très différente de la formule de changement de base pour les formes bilinéaires, $B = P^T.A.P$. Dans le cas de l'exemple 2.12 on a $B = A$, alors que dans le cas de l'exemple 3.39 (même changement de base) on a $B = (0,3048)^2.A$.

Autrement dit, les matrices des endomorphismes ne se comportent pas comme les matrices des formes bilinéaires. ▪

3.4 Matrices congruentes

Définition 3.41 Deux matrices A et B sont congruentes ssi il existe une matrice P inversible t.q :

$$B = P^T.A.P. \tag{3.36}$$

Proposition 3.42 *Des matrices sont congruentes ssi elles représentent une même application bilinéaire $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ dans deux bases différentes.*

Preuve. Les matrices d'une forme bilinéaire sont toutes congruentes entre elles, cf. (3.35).

Réciproquement supposons $B = P^T.A.P$. Soit $a(\cdot,\cdot)$ définie sur une base $(\vec{a}_i)$ par $[a(\cdot,\cdot)]_{|(\vec{a})} = A$, soit $a(\vec{a}_i, \vec{a}_j) = A_{ij}$ pour tout i,j. Soit $(\vec{b}_i)$ la base définie par $\vec{b}_j = \sum_{i=1}^n P_{ij} \vec{a}_i$ pour tout j. Alors (3.35) donne : $B = [a(\cdot,\cdot)]_{|(\vec{b})}$. ▪

4 Matrice diagonalisable

4.1 Définitions des valeurs propres et vecteurs propres

4.1.1 Valeurs propres et vecteurs propres d'une matrice

Soit $A \in K^{n^2}$ une matrice $n * n$ avec $n \geq 2$. Le problème aux valeurs propres, dit problème spectral, est :

$$\begin{cases} \text{trouver les couples } (\lambda, \vec{v}) \in K \times K^n \text{ t.q. } \vec{v} \neq \vec{0} \text{ et :} \\ A.[\vec{v}] = \lambda[\vec{v}], \end{cases} \tag{4.1}$$

où $[\vec{v}]$ est la matrice colonne des composantes de $\vec{v}$ sur la base canonique de K^n.

Dans la suite, on sous-entendra systématiquement qu'on cherche des couples $(\lambda, \vec{v})$ uniquement pour des $\vec{v} \neq \vec{0}$, même si on omet de l'écrire.

Définition 4.1 Un couple $(\lambda, \vec{v})$ solution de (4.1) est appelé un couple "valeur propre,vecteur propre" (couple "vp-vp") de la matrice A. Et alors λ est appelé valeur propre associée au vecteur propre $\vec{v}$, et $\vec{v}$ est appelé vecteur propre associé à la valeur propre λ.

Exemple 4.2 Notons $(\vec{E}_i)$ la base canonique de K^n. Pour $A = \text{diag}(\lambda_1, ..., \lambda_n)$ (matrice diagonale), les couples $(\lambda_i, \vec{E}_i)$ sont solutions immédiates de (4.1), de même que les couples $(\lambda_i, c\vec{E}_i)$, pour tout $c \neq 0$, par linéarité de A. ▪

4.1.2 Valeurs propres et vecteurs propres d'un endomorphisme

Soit $\mathcal{L} : V \to V$ un endomorphisme sur un espace vectoriel V de dimension n. Le problème aux valeurs propres, dit problème spectral, est :

$$\begin{cases} \text{trouver les couples } (\lambda, \vec{v}) \in K \times V \text{ t.q. } \vec{v} \neq \vec{0} \text{ et :} \\ \mathcal{L}.\vec{v} = \lambda\vec{v}. \end{cases} \tag{4.2}$$

Définition 4.3 Un couple $(\lambda, \vec{v})$ solution de (4.2) est appelé un couple "valeur propre,vecteur propre" (couple "vp-vp") de l'endomorphisme $\mathcal{L}$. Et alors λ est appelé valeur propre associée au vecteur propre $\vec{v}$, et $\vec{v}$ est appelé vecteur propre associé à la valeur propre λ.

Soit $(\vec{a}_i)_{i=1,...,n}$ une base de V, soit $A = [\mathcal{L}]_{|(\vec{a})}$ la matrice de l'endomorphisme $\mathcal{L}$ dans cette base. Soit $[\vec{v}]_{|(\vec{a})}$ la matrice colonne des composantes d'une vecteur $\vec{v}$ dans la base $(\vec{a}_i)$. Donc $[\mathcal{L}.\vec{v}]_{|(\vec{a})} = [\mathcal{L}]_{|(\vec{a})}.[\vec{v}]_{|(\vec{a})}$, et $\mathcal{L}.\vec{v} = \lambda\vec{v}$ implique $[\mathcal{L}]_{|(\vec{a})}.[\vec{v}]_{|(\vec{a})} = \lambda[\vec{v}]_{|(\vec{a})}$, soit $A.[\vec{v}]_{|(\vec{a})} = \lambda[\vec{v}]_{|(\vec{a})}$.

Donc trouver les couples vp-vp $(\lambda, \vec{v})$ de $\mathcal{L}$ c'est trouver les couples vp-vp $(\lambda, \vec{x})$ de $A = [\mathcal{L}]_{|(\vec{a})}$.

4.2 Exemple illustré : l'ellipse

Soit $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^n$ l'endomorphisme de $\mathbb{R}^n$ défini sur la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$ par :

$$\mathcal{L}.\vec{E}_j = \sum_{i=1}^n A_{ij}\vec{E}_i, \qquad \text{où donc} \qquad A = [A_{ij}] = [\mathcal{L}]_{|(\vec{E})}, \tag{4.3}$$

les composantes de $\mathcal{L}.\vec{E}_j$ dans la base $(\vec{E}_i)$ étant les A_{ij}.

Soit S la sphère unité de $\mathbb{R}^n$:

$$S = \{\vec{x} = \sum_{i=1}^n x_i\vec{E}_i : \sum_{i=1}^n x_i^2 = 1\}. \tag{4.4}$$

On visualise la transformation $\mathcal{L}$ en regardant l'ensemble image $\mathcal{L}(S)$ (image de S) :

$$\mathcal{L}(S) = \{\vec{y} = \mathcal{L}.\vec{x},\ \vec{x} \in S\}. \tag{4.5}$$

Si A est diagonale, $A = D = \text{diag}(\lambda_1, ..., \lambda_n)$ avec $\lambda_i > 0$ pour tout i, alors $\mathcal{L}(S)$ est facile à représenter : c'est l'ellipsoïde usuel :

$$\mathcal{L}(S) = \{\vec{y} = \sum_{i=1}^n y_i\vec{E}_i : \sum_{i=1}^n \frac{y_i^2}{\lambda_i^2} = 1\}. \tag{4.6}$$

En effet $\mathcal{L}.\vec{E}_j = \lambda_j\vec{E}_j$ pour tout j donne $\vec{y} = \mathcal{L}.\vec{x} = \sum_{i=1}^n x_i\lambda_i\vec{E}_i = \sum_{i=1}^n y_i\vec{E}_i$, où $y_i = \lambda_i x_i$ pour tout i, et donc $\sum_{i=1}^n \frac{y_i^2}{\lambda_i^2} = 1$ quand $\vec{x} \in S$.

Exemple 4.4 Dans $\mathbb{R}^2$ et $A = \text{diag}(1, 2)$, $\mathcal{L}(S)$ est l'ellipse $x_1^2 + \frac{x_2^2}{2} = 1$ de petit axe l'axe des x_1 et de grand axe l'axe des x_2. ▪

Si A n'est pas diagonale, $\mathcal{L}(S)$ n'est pas immédiat à visualiser. Sauf quand A symétrique définie positive, car alors A est diagonalisable avec des valeurs propres $\lambda_i > 0$ associés à une b.o.n. $(\vec{v}_i)$ de vecteurs propres (voir la suite). En appliquant les formules de changement de base des endomorphismes, de la base $(\vec{E}_i)$ vers la base de diagonalisation $(\vec{v}_i)$ (voir la suite), on a :

$$\mathcal{L}.\vec{v}_j = \lambda_j\vec{v}_j, \qquad D = P^{-1}.A.P, \qquad D = [\mathcal{L}]_{|(\vec{v})}, \tag{4.7}$$

où $D = \text{diag}(\lambda_1, \lambda_2)$, $P = \begin{pmatrix} [\vec{v}_1] & [\vec{v}_2] \end{pmatrix}$ est la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{v}_i)$. Les matrices $A = [\mathcal{L}]_{|(\vec{E})}$ et $D = [\mathcal{L}]_{|(\vec{v})}$ sont deux représentations de la même l'application linéaire $\mathcal{L}$, mais dans deux bases différentes. La représentation D de $\mathcal{L}$ dans la base $(\vec{v}_i)$ permet de visualiser l'ellipsoïde $\mathcal{L}(S)$: d'axes alignés suivant les $\vec{v}_i$.

Exemple 4.5 Soit $A = \begin{pmatrix} \frac{5}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{7}{4} \end{pmatrix}$ représentant un endomorphisme de $\mathbb{R}^2$ dans la base canonique. Cette matrice symétrique est diagonalisable, ses valeurs propres sont 1 et 2, et des vecteurs propres associés sont $\vec{v}_1 = \begin{pmatrix} \cos\frac{\pi}{3} \\ \sin\frac{\pi}{3} \end{pmatrix}$ et $\vec{v}_2 = \begin{pmatrix} -\sin\frac{\pi}{3} \\ \cos\frac{\pi}{3} \end{pmatrix}$ (la base canonique initiale a été tournée de $\frac{\pi}{3}$). Dans cette nouvelle base $(\vec{v}_1, \vec{v}_2)$, l'endomorphisme $\mathcal{L}$ est représenté par la matrice diagonale $D = P^{-1}.A.P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$: on "voit bien" l'ellipse $\{\vec{y} = y_1\vec{v}_1 + y_2\vec{v}_2 : y_1^2 + \frac{y_2^2}{2^2} = 1\}$ de petit axe le long de $\vec{v}_1$ et de grand axe le long de $\vec{v}_2$. ▪

4.3 Espaces propres

Définition 4.6 Si λ est valeur propre de la matrice A (resp. de l'endomorphisme $\mathcal{L}$) alors le noyau :

$$\mathrm{Ker}(A-\lambda I) = \{\vec{x} \in K^n : A.\vec{x} = \lambda\vec{x}\} \qquad (\text{resp.} \quad \mathrm{Ker}(\mathcal{L}-\lambda I) = \{\vec{x} \in V : \mathcal{L}.\vec{x} = \lambda\vec{x}\}) \tag{4.8}$$

est appelé l'espace propre de A associé à λ (resp. de $\mathcal{L}$ associé à λ).

Proposition 4.7 *Dans cette proposition A signifie matrice de K^n (resp. endomorphisme sur V) .*
1- Un sous-espace propre $\mathrm{Ker}(A-\lambda I)$ est un sous-espace vectoriel de K^n (resp. de V).
2- $\vec{v}$ est vecteur propre associé à λ ssi $\vec{v} \in \mathrm{Ker}(A-\lambda I) - \{\vec{0}\}$.
3- λ est une valeur propre ssi $\mathrm{Ker}(A-\lambda I)$ est un sous-espace vectoriel de dimension ≥ 1 (il n'est pas réduit à $\{\vec{0}\}$).
4- Un espace propre est stable par A, i.e. : si $\vec{v} \in \mathrm{Ker}(A-\lambda I)$, alors $A.\vec{v} \in \mathrm{Ker}(A-\lambda I)$.
5- Si λ_i et λ_j sont deux valeurs propres distinctes alors $\mathrm{Ker}(A-\lambda_i I) \cap \mathrm{Ker}(A-\lambda_j I) = \{\vec{0}\}$.

Preuve. 1- $\mathrm{Ker}(A-\lambda I)$ est un sous-espace vectoriel de K^n : si $A.\vec{x} = \lambda\vec{x}$ et $A.\vec{y} = \lambda\vec{y}$ alors pour tout $\alpha \in K$ alors $A.(\vec{x}+\alpha\vec{y}) = \lambda(\vec{x}+\alpha\vec{y})$.
2- $\vec{v}$ est vecteur propre associé à λ ssi $A.\vec{v} - \lambda\vec{v} = \vec{0}$ et $\vec{v} \neq \vec{0}$.
3- Si λ est valeur propre alors il existe $\vec{x} \neq \vec{0}$ t.q. $\vec{x} \in \mathrm{Ker}(A-\lambda I)$, donc $\mathrm{Ker}(A-\lambda I)$ n'est pas réduit à $\{\vec{0}\}$. Il est donc nécessairement de dimension ≥ 1 (c'est un sous-espace vectoriel).
4- Si $\vec{v} \in \mathrm{Ker}(A-\lambda I)$ alors $A.(A.\vec{v}) = A.(\lambda\vec{v}) = \lambda(A.\vec{v})$, donc $A.\vec{v} \in \mathrm{Ker}(A-\lambda I)$.
5- Si $A.\vec{v} = \lambda_i\vec{v}$ et $A.\vec{v} = \lambda_j\vec{v}$ alors $(\lambda_i - \lambda_j)\vec{v} = \vec{0}$, et si $\lambda_i \neq \lambda_j$ alors nécessairement $\vec{v} = \vec{0}$. ▪

4.4 Polynôme caractéristique

Définition 4.8 Soit $A \in K^{n^2}$ une matrice. Le polynôme caractéristique de A est le polynôme de degré n défini par, pour tout $\lambda \in K$:

$$p(\lambda) = \det(A-\lambda I), \tag{4.9}$$

déterminant de la matrice $A - \lambda I$.

Définition 4.9 Soit $\mathcal{L}$ un endomorphisme sur V. Soit $A = [\mathcal{L}]_{|(\vec{a})}$ la matrice de $\mathcal{L}$ dans une base $(\vec{a}_i)$ de V. Le polynôme caractéristique de $\mathcal{L}$ est le polynôme défini en (4.9).

Proposition 4.10 *Soit $\mathcal{L}$ un endomorphisme sur V. Son polynôme caractéristique est indépendant de la base dans laquelle $\mathcal{L}$ est représenté.*

Preuve. Soit $(\vec{b}_i)$ une autre base et soit $B = [\mathcal{L}]_{|(\vec{b})}$ la matrice de $\mathcal{L}$ dans une base $(\vec{b}_i)$. Soit P la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$. On a $B = P^{-1}.A.P$, donc :

$$\begin{aligned} \det(B-\lambda I) &= \det(P^{-1}.A.P - \lambda I) = \det(P^{-1}.(A-\lambda I).P) = \det(P^{-1})\det(P)\det(A-\lambda I) \\ &= \det(A-\lambda I). \end{aligned}$$

Donc $p(\lambda)$ donné en (4.9) ne dépend pas de la matrice représentant $\mathcal{L}$ ▪

Corollaire. Dans la suite on représente $\mathcal{L} : V \to V$ par une de ses matrices $A \in K^{n^2}$.

Notons $\lambda_i \in \mathbb{C}$ les racines complexes de p :

$$p(\lambda) = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i). \tag{4.10}$$

Notons $\mu_i \in \mathbb{C}$ les racines de p deux à deux distinctes et m_i la multiplicité de la racine μ_i :

$$n = \sum_{i=1}^{n_v} m_i, \qquad p(\lambda) = (-1)^n \prod_{i=1}^{n_v} (\lambda - \mu_i)^{m_i}, \qquad \mu_i \neq \mu_j \quad \forall i \neq j. \tag{4.11}$$

Définition 4.11 La multiplicité d'une racine du polynôme caractéristique, soit m_i dans (4.11), est appelée la multiplicité de la valeur propre μ_i et des valeurs propres λ_j t.q. $\lambda_j = \mu_i$.

Et on utilise alors l'expression "la valeur propre λ_j est comptée autant de fois que sa multiplicité" pour passer de (4.11) à (4.10).

Exemple 4.12 Soit $\mu_1 \neq \mu_2$ et $A = \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_1 & 0 \\ 0 & 0 & \mu_2 \end{pmatrix}$. La matrice A a deux valeurs propres distinctes : μ_1 de multiplicité 2 et μ_2 de multiplicité 1. On pose $\lambda_1 = \mu_1 = \lambda_2$ et $\lambda_3 = \mu_2$. Et $A = \text{diag}(\mu_1, \mu_1, \mu_2) = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$. Ici $\mu_1 = \lambda_1$ a été compté deux fois (sa multiplicité est 2). ▪

Proposition 4.13 *λ est valeur propre ssi $A - \lambda I$ n'est pas inversible, i.e. ssi λ est racine du polynôme caractéristique p, ou encore ssi* $\dim(\text{Ker}(A - \lambda I)) \geq 1$ *:*

$$\lambda \text{ valeur propre} \quad \Longleftrightarrow \quad p(\lambda) = 0 \quad \Longleftrightarrow \quad \dim(\text{Ker}(A - \lambda I)) \geq 1. \tag{4.12}$$

Preuve. Première équivalence : $\Rightarrow$. Si λ est valeur propre alors il existe $\vec{x}$ non nul tel que $(A - \lambda I)\vec{x} = \vec{0}$, donc $A - \lambda I$ n'est pas inversible, donc $\det(A - \lambda I) = 0$.

$\Leftarrow$. Si λ n'est pas valeur propre, alors si $\vec{x}$ vérifie $(A - \lambda I).\vec{x} = \vec{0}$ on a nécessairement $\vec{x} = \vec{0}$ (sinon λ serait valeur propre), donc $\det(A - \lambda I) \neq 0$ (sinon les vecteurs colonnes seraient liés).

Deuxième équivalence : voir proposition précédente 4.7. ▪

Exercice 4.14 Montrer que si T est une matrice triangulaire supérieure, alors ses valeurs propres sont données par les scalaires sur la diagonale.

Et donc en particulier si $T = D$ est diagonale, ses valeurs propres sont les scalaires sur la diagonale.

Réponse. Soit λ_i les scalaires sur la diagonale de T. Alors $p(\lambda) = \det(T - \lambda I) = \prod_{i=1}^{n} (\lambda - \lambda_i)$ (calcul immédiat). Donc les valeurs propres sont données par les λ_i. ▪

Exercice 4.15 Montrer que la trace $\text{Tr}(A)$ et le déterminant $\det(A)$ sont donnés dans $\mathbb{C}$ par :

$$\text{Tr}(A) = \sum_{i=1}^{n} \lambda_i, \qquad \det(A) = \prod_{i=1}^{n} \lambda_i. \tag{4.13}$$

Réponse. $p(\lambda) = \det(A - \lambda I) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + ... + a_0$ obtenu par exemple en développant par rapport à la première colonne. D'où $a_0 = p(0) = \det(A)$, et $a_{n-1} = \sum_{i=1}^{n} a_{ii}$ car dans le calcul de a_{n-1} interviennent exactement les $n-1$ éléments de la diagonale de $A - \lambda I$ (coefficient de λ^{n-1}). ▪

Exercice 4.16 Montrer que si λ est valeur propre de A alors λ est aussi valeur propre de A^T avec la même multiplicité.

Réponse. Comme $I^T = I$ on a $\det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I) = p(\lambda)$: donc A et A^T ont le même polynôme caractéristique. ▪

4.5 Matrice diagonalisable : $D = P^{-1}.A.P$

4.5.1 Définition et exemples

Définition 4.17 On dit que A est diagonalisable dans K ssi il existe n couples $(\lambda_i, \vec{v}_i) \in K \times K^n$ solutions de (4.1) t.q. $(\vec{v}_i)_{i=1,\dots,n}$ est une base de K^n :

$$\exists\lambda_1, ..., \lambda_n \in K^n, \quad \exists(\vec{v}_i)_{i=1,\dots,n} \text{ base de } K^n, \quad \text{t.q.} \quad A.\vec{v}_i = \lambda_i\vec{v}_i \quad \forall i = 1, ..., n, \tag{4.14}$$

au sens $A.[\vec{v}_j] = \lambda_j[\vec{v}_j]$ où on a utilisé implicitement la base canonique.

Exemple 4.18 Une matrice diagonale $A = \text{diag}(\lambda_1, ..., \lambda_n)$ est diagonalisable (c'est la moindre des choses...) : ses valeurs propres sont les scalaires λ_i lus sur la diagonale, et une base de diagonalisation est la base canonique (immédiat). En particulier la matrice identité, cas particulier pour lequel toutes les bases sont des bases de diagonalisation. ▪

Exemple 4.19 Les matrices des projections sont diagonalisables :

Soit $(\vec{v}_i)_{i=1,\dots,n}$ est une base de $\mathbb{R}^n$ et soit l'application linéaire de projection sur le s.e.v. $\text{Vect}\{\vec{v}_1, ..., \vec{v}_k\}$ parallèlement à $\text{Vect}\{\vec{v}_{k+1}, ..., \vec{v}_n\}$. Sa matrice A dans cette base est donnée par $A.\vec{v}_j = \vec{v}_j$ pour tout $j \le k$ et $A.\vec{v}_j = \vec{0}$ pour tout $j > k$. Donc $A = \text{diag}(\lambda_i)$ où les λ_i valent 1 pour $j \le k$, et 0 pour $j > k$. ▪

Exemple 4.20 La matrice antisymétrique $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est diagonalisable dans $\mathbb{C}$, puisque $p(\lambda) = \lambda^2 + 1$ donne $\lambda = \pm i$. Deux vecteurs propres associés sont $\vec{v}_+ = \begin{pmatrix} 1 \\ i \end{pmatrix}$ et $\vec{v}_- = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. ▪

Exemple 4.21 La matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (premier bloc de Jordan) n'est pas diagonalisable. En effet λ est valeur propre ssi $0 = \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2$. Et donc $\lambda = 1$ est la seule valeur propre. Et l'espace propre associé est $\text{Ker}(A - I) = \text{Vect}\{\vec{E}_1\}$ de dimension 1. Et $\text{Vect}\{\vec{E}_1\} \subsetneq \mathbb{R}^2$: on ne peut donc pas former une base de $\mathbb{R}^2$ avec des vecteurs propres.

(D'ailleurs si A était diagonalisable on aurait $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{diag}(1, 1) = I$, voir plus loin, ce qui n'est pas le cas.) ▪

Exercice 4.22 Montrer que la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ est diagonalisable dans $\mathbb{R}$: calculer ses valeurs propres et donner des vecteurs propres associés.

Réponse. $p(\lambda) = \det(A - \lambda I) = \lambda^2 - 7\lambda = \lambda(\lambda - 7)$. Donc 0 et 7 sont valeurs propres.

Le vecteur propre $\vec{v}_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ associé à $\lambda_1 = 0$ est obtenu à l'aide de $(A - \lambda_1 I).\vec{v}_1 = \vec{0}$, soit $x + 2y = 0$ (la seconde équation est redondante), et par exemple $\vec{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ convient.

Le vecteur propre $\vec{v}_2 = \begin{pmatrix} x \\ y \end{pmatrix}$ associé à $\lambda_2 = 7$ est obtenu à l'aide de $(A - \lambda_2 I).\vec{v}_2 = \vec{0}$, soit $-6x + 2y = 0$ (la seconde équation est redondante), et par exemple $\vec{v}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ convient.

Noter que, relativement au produit scalaire usuel, ces vecteurs ne sont pas orthogonaux de $\mathbb{R}^2$ (la matrice A n'est pas symétrique, voir plus loin). ▪

Exercice 4.23 Diagonaliser la matrice $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$.

Réponse. Les valeurs propres sont 1 et 6, les vecteurs propres sont proportionnels à $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ et à $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. ▪

4.5.2 Formule $D = P^{-1}.A.P$

Proposition 4.24 *Toute matrice diagonalisable est semblable à une matrice diagonale. Donc, si :*

1- $A \in K^{n^2}$ est une matrice diagonalisable dans K,

2- $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ est la matrice diagonale des valeurs propres,
3- $(\vec{v}_i)$ est une base de K^n de vecteurs propres, avec donc $A.\vec{v}_j = \lambda_j \vec{v}_j$ pour tout j, cf. (4.14),
4- P est la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{v}_i)$, i.e. :

$$\vec{v}_j = \sum_{i=1}^{n} P_{ij} \vec{E}_i, \qquad [\vec{v}_j] = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \tag{4.15}$$

alors : la relation $A.\vec{v}_j = \lambda_j \vec{v}_j$ pour tout $j = 1, ..., n$ s'écrit :

$$A.P = P.D, \tag{4.16}$$

soit :

$$D = P^{-1}.A.P, \tag{4.17}$$

ou encore :

$$A = P.D.P^{-1}. \tag{4.18}$$

(Autrement dit : un endomorphisme est diagonalisable s'il existe une base dans laquelle la matrice le représentant est diagonale.)

Preuve. On a :

$$\begin{aligned} A.P &= A.(\,[\vec{v}_1] \quad ... \quad [\vec{v}_n]\,) \\ &= (\,A.[\vec{v}_1] \quad ... \quad A.[\vec{v}_n]\,) \\ &= (\,\lambda_1[\vec{v}_1] \quad ... \quad \lambda_n[\vec{v}_n]\,). \end{aligned} \tag{4.19}$$

Et :

$$\begin{aligned} P.D &= P.\big(\,[\lambda_1 \vec{E}_1] \quad ... \quad [\lambda_n \vec{E}_n]\,\big) \\ &= \big(\,P.[\lambda_1 \vec{E}_1] \quad ... \quad P.[\lambda_n \vec{E}_n]\,\big) \\ &= \big(\,\lambda_1 P.[\vec{E}_1] \quad ... \quad \lambda_n P.[\vec{E}_n]\,\big) \\ &= (\,\lambda_1[\vec{v}_1] \quad ... \quad \lambda_n[\vec{v}_n]\,). \end{aligned} \tag{4.20}$$

On a bien $A.P = P.D$ n'est autre que l'écriture matricielle par colonnes des n égalités $A.\vec{v}_j = \lambda_j \vec{v}_j$ de définition des couples vp-vp : on a (4.16).

De plus, $(\vec{v}_i)$ étant une base, P est inversible. D'où P^{-1} existe et (4.16) s'écrit également (4.17) ou (4.18). ▪

Exemple 4.25 Soit $(\vec{v}_n)$ une base. Soit Π la projection sur $\mathrm{Vect}\{\vec{v}_1, ..., \vec{v}_{n-1}\}$ parallèlement à $\vec{v}_n$. Soit $A = [\Pi]_{|(\vec{E})}$ la matrice de l'application linéaire Π dans la base canonique. Alors $D = P^{-1}.A.P = \mathrm{diag}(1, ..., 1, 0) = [\Pi]_{|(\vec{v})}$ est la matrice diagonale représentant la projection Π dans la base $(\vec{v}_i)$. ▪

Exercice 4.26 Diagonaliser la matrice $T = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

Réponse. Les valeurs propres sont $\lambda_1 = 1$ et $\lambda_2 = 3$. Des vecteurs propres associés sont $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Et ici $P = (\,[\vec{v}_1] \quad [\vec{v}_2]\,) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, et $D = \mathrm{diag}(1, 3) = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$, et on vérifie (4.16) : $T.P = P.D = \begin{pmatrix} 1 & 3 \\ 0 & 3 \end{pmatrix}$, ou encore $D = P^{-1}.T.P$. ▪

Exercice 4.27 Montrer que si A est une matrice réelle diagonalisable, alors A^T est diagonalisable, et donner une base de vecteurs propres.

Réponse. $A = P.D.P^{-1}$ donne $A^T = P^{-T}.D.P^T = R.D.R^{-1}$ où on a posé $R = P^{-T}$. Donc $A^T.R = R.D$: donc A^T a même valeurs propres que A (même matrice diagonale), et sa j-ème valeur propre est associée au j-ème "vecteur colonne" de $R = P^{-T}$.

Cas particulier $A = A^T$: on verra qu'on peut choisir P t.q. $P^T = P^{-1}$, et donc $R = P$ convient : A et $A^T = A$ ont mêmes vecteurs propres ! ▪

Exercice 4.28 Montrer que A diagonalisable est non inversible ssi 0 est valeur propre de A.

Réponse. $\det(A) = \det(P.D.P^{-1}) = \det(P)\det(D)\det(P^{-1}) = \det(D) = \prod_{i=1}^n(\lambda_i)$. ▪

Exercice 4.29 Montrer que si $A = P.D.P^{-1}$ alors $A^k = P.D^k.P^{-1}$ pour tout $k \in \mathbb{N}$.

Réponse. Vrai pour $k = 0$ ($I = I$), pour $k = 1$ (hypothèse), $A^2 = P.D.P^{-1}.P.D.P^{-1} = P.D^2.P^{-1}$, récurrence immédiate. ▪

4.5.3 Caractérisation $\dim(\mathrm{Ker}(A - \mu_i I)) = m_i$

Proposition 4.30 *Soit A matrice $n * n$ et soit $\mu_1, ..., \mu_{n_v}$ ses valeurs propres 2 à 2 distinctes de multiplicité m_i. Alors :*

$$\begin{aligned} A \text{ diagonalisable} \quad &\iff \quad \forall i = 1, ..., n_v, \quad \dim(\mathrm{Ker}(A - \mu_i I)) = m_i \\ &\iff \quad \bigoplus_{i=1}^{n_v} \mathrm{Ker}(A - \mu_i I) = K^n. \end{aligned} \tag{4.21}$$

En particulier, si A possède n valeurs propres 2 à 2 distinctes (donc toutes de multiplicité $= 1$), alors A est diagonalisable.

Preuve. Notons $F \stackrel{\text{def}}{=} \bigoplus_{i=1}^{n_v} \mathrm{Ker}(A - \mu_i I)$: c'est bien une somme directe, cf. proposition 4.7, point 5.

Supposons $F = K^n$: on prend une base dans chaque $\mathrm{Ker}(A - \mu_i I)$ et on obtient une base de K^n de vecteurs propres. Donc A est diagonalisable et $\dim(\mathrm{Ker}(A - \mu_i I)) = m_i$ pour tout i.

Supposons $\dim(\mathrm{Ker}(A - \mu_i I)) = m_i$ pour tout i : on prend une base dans chaque $\mathrm{Ker}(A - \mu_i I)$ et on obtient une famille libre de K^n de vecteurs propres constituée de n vecteurs, donc A est diagonalisable et $F = K^n$. ▪

Exercice 4.31 Vérifier que si A est une matrice réelle, si $\lambda \in \mathbb{C} - \mathbb{R}$ est valeur propre, alors $\overline{\lambda}$ est également valeur propre, et les vecteurs propres associés sont conjugués. Le vérifier sur l'exemple 4.20.

Réponse. Comme A est réelle on a $\overline{A} = A$. Donc $A.\vec{v} = \lambda\vec{v}$ donne $\overline{A.\vec{v}} = A.\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}$. ▪

Exercice 4.32 Montrer qu'on a toujours $\dim(\mathrm{Ker}(A - \mu_i I)) \le m_i$.

Réponse. Sinon $\dim(\mathrm{Ker}(A-\mu_i I)) \ge m_i+1$. Soit $(\vec{w}_j)_{j=1,...,m_i+1}$ une famille libre dans $\mathrm{Ker}(A-\mu_i I)$. On la complète par une famille libre $(\vec{w}_{m_i+2}, ..., \vec{w}_n)$ pour former une base de K^n. Soit $P = ([\vec{w}_1] \ \ ... \ \ [\vec{w}_n])$ la matrice de passage. Alors $p(\lambda) = \det(A - \lambda I) = \det(P^{-1}.A.P - \lambda I) = \det(P^{-1})\det(A.P - \lambda P) = \det(P^{-1})\det((\mu_i - \lambda)[\vec{w}_1], ..., (\mu_i - \lambda)[\vec{w}_{m_i+1}], ...) = (\mu_i - \lambda)^{m_i+1}\det(P^{-1})\det([\vec{w}_1], ..., [\vec{w}_{m_i+1}], ...)$, donc μ_i est de multiplicité $\ge m_i + 1$: contraire à la définition de m_i. ▪

4.6 Matrices qui commutent avec une matrice diagonale

Proposition 4.33 *Si $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ avec $\lambda_i \ne \lambda_j$ pour tout $i \ne j$ (les valeurs propres sont distinctes 2 à 2, i.e. toutes de multiplicité $= 1$), alors une matrice B commute avec D ssi B est une matrice diagonale :*

$$B.D = D.B \quad \iff \quad B = \mathrm{diag}(\alpha_1, ..., \alpha_n). \tag{4.22}$$

Si $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ est diagonale avec les λ_i de multiplicité $m_i \ge 1$ rangées de telle sorte que les valeurs propres identiques sont adjacentes, notant n_v le nombre de valeurs propres distinctes 2 à 2, alors une matrice B commute avec D ssi B est une matrice diagonale par blocs :

$$B.D = D.B \quad \iff \quad B = \mathrm{diag}(B_1, ..., B_{n_v}). \tag{4.23}$$

où les matrices B_i sont des matrices carrées $m_i * m_i$.

Plus généralement, si A une matrice diagonalisable, avec la notation $A = P.D.P^{-1}$ de la proposition 4.24, si les λ_i pour $i = 1, ..., n$ sont les valeurs propres de A de multiplicité $m_i \ge 1$ rangées dans D de telle sorte que les valeurs propres identiques sont adjacentes, alors, B commute avec A ssi $B = P.D_B.P^{-1}$ où $D_B = \mathrm{diag}(B_1, ..., B_{n_v})$ est une matrice diagonale par blocs :

$$B.A = A.B \quad \iff \quad B = P.\mathrm{diag}(B_1, ..., B_{n_v}).P^{-1}, \tag{4.24}$$

où les matrices B_i sont des matrices carrées $m_i * m_i$ (donc B est "diagonalisable par blocs" avec la matrice de passage P de A).

Preuve. Traitons le cas général (4.24), les cas particuliers (4.22) et (4.23) s'en déduisant. Soit $(\vec{v}_i)$ une base de diagonalisation associée aux vp λ_i. Donc $A.P = P.D$ où $P = (\,[\vec{v}_1] \;\; ... \;\; [\vec{v}_n]\,)$ est la matrice de passage "contenant les vecteurs propres", et $D = \operatorname{diag}(\lambda_1, ..., \lambda_n)$.

D'une part on a $(A.B).\vec{v}_i = A.(B.\vec{v}_i)$, et d'autre part on a $(A.B).\vec{v}_i = (B.A).\vec{v}_i = B.(A.\vec{v}_i) = B.(\lambda_i \vec{v}_i) = \lambda_i B.\vec{v}_i$.

Donc $A.(B.\vec{v}_i) = \lambda_i(B.\vec{v}_i)$, soit $(A - \lambda_i I).(B.\vec{v}_i) = \vec{0}$, i.e. $B.\vec{v}_i \in \operatorname{Ker}(A - \lambda_i I)$.

Notons μ_i pour $i = 1, ..., n_v$ les valeurs propres distinctes de$A$ comptées autant de fois que leur multiplicité m_i.

Considérons le cas $i = 1$: ici $\operatorname{Ker}(A - \mu_1 I) = \operatorname{Vect}\{\vec{v}_1, ..., \vec{v}_{m_1}\}$. Et comme $B.\vec{v}_1 \in \operatorname{Ker}(A - \mu_1 I)$ pour $j = 1, ..., m_1$, il existe m_1 réels c_{i1} pour $i = 1, ..., m_1$ t.q. $B.[\vec{v}_1] = \sum_{i=1}^{m_1} c_{i1} \vec{v}_i$.

Alors on a obtenu $B.[\vec{v}_1] \overset{\text{noté}}{=} P.\vec{v}_1$, où $\vec{v}_1$ a ses m_1 premières composantes $= c_{i1}$ pour $i = 1, ..., m_1$ et les autres nulles.

De même pour $B.[\vec{v}_j] = \sum_{i=1}^{m_1} c_{ij} \vec{v}_i \overset{\text{noté}}{=} P.\vec{v}_j$ pour $j = 1, ..., m_1$, où $\vec{v}_j$ a ses m_1 premières composantes $= c_{ij}$ pour $i = 1, ..., m_1$ et les autres nulles.

Notons $B_1 = [c_{ij}]_{\substack{i=1,...,m_1 \\ j=1,...,m_1}}$. On a obtenu $B.(\,[\vec{v}_1] \;\; ... \;\; [\vec{v}_{m_1}]\,) = P.\begin{pmatrix} B_1 \\ 0 \end{pmatrix}$ où $\begin{pmatrix} B_1 \\ 0 \end{pmatrix} = (\,[\vec{v}_1] \;\; ... \;\; [\vec{v}_{m_1}]\,)$.

On procède de même avec les valeurs propres suivantes.

On a obtenu $B.P = P.D_B$ où $D_B = \operatorname{diag}(B_1, ..., B_n)$ est la matrice diagonale par blocs cherchée.

Cas particulier $A = D$ diagonale : ici la base de diagonalisation est la base canonique, et donc $P = I$. Et donc B est diagonale par blocs, et en particulier diagonale si les valeurs propres sont 2 à 2 distinctes. ∎

5 Une matrice symétrique réelle est diagonalisable

Rappel : $\mathbb{C}$ est un corps "algébriquement clos" : tout polynôme complexe de degré $n \geq 1$ admet au moins une racine complexe. Alors que $\mathbb{R}$ n'est pas "algébriquement clos" : le polynôme $x^2 + 1$ n'a pas de racine réelle. On plonge $\mathbb{R}$ dans $\mathbb{C}$ pour avoir l'existence des racines du polynôme caractéristique (dans $\mathbb{C}$).

5.1 Les valeurs propres d'une matrice réelle symétrique sont réelles

Les matrices de $\mathbb{C}^{n^2}$ transconjuguées ont été définies en (1.46) par $A^* = \overline{A^T}$ et les matrices hermitiennes sont les matrices complexes vérifiant $A^* = A$, cf. (1.51).

Proposition 5.1 *Si A est hermitienne alors ses valeurs propres sont toutes réelles.*
En particulier, les valeurs propres d'une matrice symétrique réelle sont toutes réelles.

Preuve. Soit $\lambda \in \mathbb{C}$ une racine du polynôme caractéristique $p(\lambda) = \det(A - \lambda I)$. Soit $\vec{v} \in \mathbb{C}^n - \{\vec{0}\}$ le vecteur propre associé : on a $A.\vec{v} = \lambda \vec{v}$ avec $\vec{v} \neq \vec{0}$. Donc :

$(A.\vec{v}, \vec{v})_{\mathbb{C}^n} = (\lambda \vec{v}, \vec{v})_{\mathbb{C}^n} = \lambda (\vec{v}, \vec{v})_{\mathbb{C}^n}$,

et, ayant A hermitienne et $(\cdot, \cdot)_{\mathbb{C}^n}$ sesquilinéaire :

$(A.\vec{v}, \vec{v})_{\mathbb{C}^n} = (\vec{v}, A.\vec{v})_{\mathbb{C}^n} = (\vec{v}, \lambda \vec{v})_{\mathbb{C}^n} = \overline{\lambda} (\vec{v}, \vec{v})_{\mathbb{C}^n}$,

d'où $\lambda = \overline{\lambda}$: toute valeur propre de A hermitienne est réelle.

Et une matrice symétrique réelle est hermitienne. ∎

5.2 Théorème de diagonalisation des matrices symétriques réelles

Soit $(\vec{E}_i)$ la base canonique de $\mathbb{R}^n$. On munit $\mathbb{R}^n$ de son produit scalaire usuel $(\cdot, \cdot)_{\mathbb{R}^n}$.

Théorème 5.2 (diagonalisation des matrices symétriques réelles.) *Soit $A \in \mathbb{R}^{n^2}$ une matrice symétrique réelle. Alors A est diagonalisable dans une b.o.n. de $(\mathbb{R}^n, (\cdot, \cdot)_{\mathbb{R}^n})$:*

$$\exists \lambda_1, ..., \lambda_n \in \mathbb{R}, \quad \exists (\vec{v}_i)_{i=1,...,n} \text{ b.o.n de } \mathbb{R}^n, \quad A.\vec{v}_j = \lambda_j \vec{v}_j \quad \forall j = 1, ..., n. \tag{5.1}$$

Notant $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ la matrice diagonale des valeurs propres, et $P = (\,[\vec{v}_1] \;\; ... \;\; [\vec{v}_n]\,)$ la matrice de passage de la base canonique à la base $(\vec{v}_i)$ (stocke dans sa j-ème colonne les composantes *de $\vec{v}_j$ dans la base canonique), (5.1) s'écrit également sous la forme :*

$$A.P = P.D, \qquad \text{avec} \qquad P^T.P = I, \tag{5.2}$$

soit :

$$D = P^{-1}.A.P, \qquad \text{avec} \qquad P^{-1} = P^T. \tag{5.3}$$

Preuve. Démonstration de (5.1) par récurrence. Pour $n = 1$, A est une matrice réduite à un réel, donc $Ax = \lambda x$, et $(\lambda, 1) \in \mathbb{R} \times \mathbb{R}^1$ est un couple vp-vp qui convient.

Supposons que ce soit vrai pour toute matrice symétrique $n * n$.

Soit A une matrice symétrique $(n+1)*(n+1)$. Donc $\det(A-\lambda I)$ est une polynôme de degré $n+1$, et donc admet au moins une racine λ_{n+1} dans $\mathbb{C}$. Et A étant symétrique, cette racine est réelle, cf. proposition 5.1. Soit $\vec{w}_{n+1} \in \mathrm{Ker}(A - \lambda_{n+1} I)$ vecteur propre associé t.q. $||\vec{w}_{n+1}||_{\mathbb{R}^{n+1}} = 1$.

Soit $\vec{w}_1, \dots, \vec{w}_n \in \mathbb{R}^{n+1}$ choisis tels que $(\vec{w}_1, \dots, \vec{w}_n, \vec{w}_{n+1})$ soit une base orthonormée de $\mathbb{R}^{n+1}$ (par exemple à l'aide de Gram–Schmidt). Et notons $G = (\,[\vec{w}_1] \quad \dots \quad [\vec{w}_{n+1}]\,)$ la matrice $(n+1) * (n+1)$ dont les colonnes sont les composantes des vecteurs $\vec{w}_j$ dans la base canonique. Donc $G^T.G = I$ (par construction de la base $(\vec{w}_1, \dots, \vec{w}_n, \vec{w}_{n+1})$). Et on a :

$$\begin{aligned} G^T.A.G &= \begin{pmatrix} [\vec{w}_1]^T \\ \vdots \\ [\vec{w}_{n+1}]^T \end{pmatrix} . (\,[A.\vec{w}_1] \quad \dots \quad [A.\vec{w}_{n+1}]\,) \\ &= \begin{pmatrix} [\vec{w}_1]^T.A.[\vec{w}_1] & \dots & [\vec{w}_1]^T.A.[\vec{w}_{n+1}] \\ \vdots & & \vdots \\ [\vec{w}_{n+1}]^T.A.[\vec{w}_1] & \dots & [\vec{w}_{n+1}]^T.A.[\vec{w}_{n+1}] \end{pmatrix}, \end{aligned}$$

avec $\vec{w}_{n+1}^T.A.\vec{w}_{n+1} = \lambda_{n+1}\vec{w}_{n+1}^T.\vec{w}_{n+1} = \lambda_{n+1}$, et pour $i \neq 0$, $\vec{w}_i^T.A.\vec{w}_{n+1} = \lambda_{n+1}\vec{w}_i^T.\vec{w}_{n+1} = \lambda_{n+1}.0 = 0 = \vec{w}_{n+1}^T.A.\vec{w}_j$, la dernière égalité par symétrie de A. Donc :

$$G^T.A.G = \begin{pmatrix} & & & 0 \\ & B & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & \lambda_{n+1} \end{pmatrix}, \tag{5.4}$$

où B est la matrice $n * n$ symétrique $= [\vec{v}_i^T.A.\vec{v}_j]_{\substack{1 \le i \le n \\ 1 \le j \le n}}$. Par hypothèse de récurrence, B est diagonalisable dans une b.o.n. $(\vec{u}_1, \dots, \vec{u}_n)$ de $\mathbb{R}^n$. On plonge $\mathbb{R}^n$ dans $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ et on prolonge les $\vec{u}_i$ en les vecteurs $\vec{v}_i = (0, \vec{u}_i)$. Et il est immédiat que la base $(\vec{v}_1, ..., \vec{v}_n, \vec{v}_{n+1})$ est une base orthonormée de diagonalisation de A.

Enfin, notant P la matrice dont les colonnes sont données par les composantes de $\vec{v}_j$, on a $A.P = P.D$, qui est une écriture condensée de $A.\vec{v}_k = \lambda_k \vec{v}_k$ (égalité des colonnes des matrices $A.P$ et $P.D$), i.e. on a (5.2). ▪

Exercice 5.3 Diagonaliser $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.

Réponse. $\det(M - \lambda I) = \det \begin{pmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} = \lambda^2 - 6\lambda + 8 = (\lambda - 2)(\lambda - 4)$. Valeurs propres $\lambda_1 = 2$, $\lambda_2 = 4$. Deux vecteurs propres associés $\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; on vérifie que $(\vec{v}_1, \vec{v}_2)_{\mathbb{R}^2} = 0$. Ici $D = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ et comme $||\vec{v}_1|| = ||\vec{v}_2|| = \sqrt{2}$, on a $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ matrice orthonormée : on vérifie $P^T.P = I$. ▪

Exercice 5.4 Diagonaliser $M = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Réponse. $\det(M - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = \lambda^2 - 3\lambda + 1$. Discrimiant $\Delta = 9 - 4 = 5$, racines $\lambda_\pm = \dfrac{3 \pm \sqrt{5}}{2}$. Deux vecteurs propres associés $\vec{v}_\pm = \begin{pmatrix} 1 \\ \frac{1 \pm \sqrt{5}}{2} \end{pmatrix}$. ▪

Exercice 5.5 Si A est symétrique, si λ_1 et λ_2 sont deux valeurs propres distinctes, montrer que les espaces propres sont orthogonaux.

Réponse. Soit $\vec{e}_i$ un vecteur propre associé à λ_i, $i = 1, 2$. On a $(A.\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = \lambda_i(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n}$ et $(A.\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = (\vec{e}_i, A.\vec{e}_j)_{\mathbb{R}^n} = \lambda_j(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n}$ car A est symétrique. Donc si $\lambda_i \neq \lambda_j$ on a nécessairement $(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = 0$. ▪

Exercice 5.6 Montrer que si A est symétrique définie positive alors $(A.\vec{x}, \vec{x})_{\mathbb{R}^n} \geq \lambda_{\min}||\vec{x}||^2_{\mathbb{R}^n}$ pour tout $\vec{x} \in \mathbb{R}^n$, où $\lambda_{\min} = \min(\lambda_1, ..., \lambda_n)$ est la plus petite des valeurs propres.

Réponse. On a $D = P^{-1}.A.P$ avec $P^T = P^{-1}$, donc $A = P.D.P^{-1} = P.D.P^T$, donc $(A.\vec{x}, \vec{x})_{\mathbb{R}^n} = (P.D.P^T.\vec{x}, \vec{x})_{\mathbb{R}^n} = (D.\vec{y}, \vec{y})_{\mathbb{R}^n}$ où on a posé $\vec{y} = P^T.\vec{x}$. Et $(D.\vec{y}, \vec{y})_{\mathbb{R}^n} = \sum_{i=1}^n \lambda_i y_i^2$ quand $\vec{y} = \sum_{i=1}^n y_i \vec{E}_i$, donc $(D.\vec{y}, \vec{y})_{\mathbb{R}^n} \geq \min_{i=1,...,n}(\lambda_i) \sum_{i=1}^n y_i^2 = \lambda_{\min}||\vec{y}||^2_{\mathbb{R}^n}$. Comme $P^T.P = I$ on a Avec $||\vec{y}||^2 = ||\vec{x}||^2_{\mathbb{R}^n}$, d'où $(A.\vec{x}, \vec{x})_{\mathbb{R}^n} \geq \lambda_{\min}||\vec{x}||^2_{\mathbb{R}^n}$, ce pour tout $\vec{x} \in \mathbb{R}^n$. ▪

Remarque 5.7 S'il existe Q inversible telle que $D = Q^T.A.Q$ est diagonale, alors $A = Q^{-T}.D.Q^{-1}$ est symétrique (immédiat), et donc est diagonalisable. Mais il n'y a aucune raison d'avoir $Q^T = Q^{-1}$, i.e. que $Q^T.Q = I$, i.e. que les "vecteurs constituant les colonnes de Q" soient orthonormés : prendre $A = I$, $Q = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = Q^T$ qui donne $D = Q^T.I.Q = Q^T.Q = Q^2 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \neq I$ (les "colonnes" ne sont pas toutes normées). ▪

Réciproque :

Proposition 5.8 *Si $A \in \mathbb{R}^{n^2}$ est diagonalisable et si on peut choisir des vecteurs propres tels qu'ils forment une b.o.n. de $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$, alors A est symétrique.*

Preuve. Notons $(\vec{v}_i)$ une b.o.n. de vecteurs propres, et $P = ([\vec{v}_1] \quad ... \quad [\vec{v}_n])$ la matrice qui stocke dans sa j-ème colonne les composantes de $\vec{v}_j$ dans la base canonique : donc $A = P.D.P^{-1} = P.D.P^T$, donc $A^T = (P^T)^T.D^T.P^T = P.D.P^T$, donc $A^T = A$. ▪

5.3 Diagonalisation d'un endomorphisme symétrique

Soit V un espace vectoriel sur $\mathbb{R}$, muni d'un produit scalaire $(\cdot,\cdot)_V$, de dimension n.

Corollaire 5.9 *Soit $\mathcal{L}$ un endomorphisme symétrique dans $(V, (\cdot,\cdot)_V)$. Alors $\mathcal{L}$ est diagonalisable dans une b.o.n. de $(V, (\cdot,\cdot)_V)$:*

$$\exists \lambda_1, ..., \lambda_n \in \mathbb{R}, \qquad \exists (\vec{v}_1, ..., \vec{v}_n) \text{ b.o.n. de } (V, (\cdot,\cdot)_V), \qquad \mathcal{L}.\vec{v}_j = \lambda_j \vec{v}_j \quad \forall j = 1, ..., n. \tag{5.5}$$

Preuve. Soit $(\vec{a}_i)$ une b.o.n. de V. Soit $A = [\mathcal{L}]_{|(\vec{a})} \in \mathbb{R}^{n^2}$. La matrice A est symétrique car $\mathcal{L}$ est symétrique et $(\vec{a}_i)$ est une b.o.n. (proposition 2.18 page 25).

1- Passage dans $\mathbb{R}^n$. A, symétrique réelle, est diagonalisable dans une b.o.n. $(\vec{e}_i)$ de $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$:

$$D = P^{-1}.A.P \qquad \text{avec} \qquad P^T.P = I, \tag{5.6}$$

où $D = \text{diag}(\lambda_1, ..., \lambda_n)$ est la matrice diagonale des valeurs propres, et où $P = ([\vec{e}_1] \quad ... \quad [\vec{e}_n])$ stocke dans sa j-ème colonne les composantes de $\vec{e}_j$ dans la base canonique $(\vec{E}_i)$: pour tout j :

$$\vec{e}_j = \sum_i P_{ij} \vec{E}_i, \qquad [\vec{e}_j]_{|(\vec{E})} = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \qquad A.[\vec{e}_j]_{|(\vec{E})} = \lambda_j [\vec{e}_j]_{|(\vec{E})}. \tag{5.7}$$

2- Retour dans V : soit $\vec{v}_j$ la fonction de V dont les composantes dans la b.o.n. $(\vec{a}_i)$ sont celles de $\vec{e}_j$ dans la base canonique de $\mathbb{R}^n$: pour tout j :

$$\vec{v}_j = \sum_i P_{ij} \vec{a}_i, \qquad \text{donc} \qquad [\vec{v}_j]_{|(\vec{a})} = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix} \qquad \text{et} \qquad A.[\vec{v}_j]_{|(\vec{a})} = \lambda_j [\vec{v}_j]_{|(\vec{a})}. \tag{5.8}$$

Vérifions que $(\vec{v}_j)$ est une b.o.n. dans $(V,(\cdot,\cdot)_V)$:

$$(\vec{v}_j,\vec{v}_j)_V=\sum_{k\ell}P_{ki}P_{\ell j}(\vec{a}_k,\vec{a}_\ell)_V=\sum_{k\ell}P_{ki}P_{\ell j}\delta_{k\ell}=\sum_k P_{ki}P_{kj}=(P^T.P)_{ij}=\delta_{ij}. \tag{5.9}$$

Et :

$$\begin{aligned}\mathcal{L}.\vec{v}_j=&\sum_i P_{ij}\mathcal{L}.\vec{a}_i=\sum_{ik}P_{ij}A_{ki}\vec{a}_k=\sum_{ik\ell}P_{ij}A_{ki}(P^{-1})_{\ell k}\vec{v}_\ell\\ =&\sum_\ell(P^{-1}.A.P)_{\ell j}\vec{v}_\ell=\sum_\ell D_{\ell j}\vec{v}_\ell=\lambda_j\vec{v}_j,\end{aligned} \tag{5.10}$$

donc $[\mathcal{L}]_{|(\vec{v})}=D=\text{diag}(\lambda_i)$ où $(\vec{v}_i)$ est une b.o.n. de diagonalisation de $\mathcal{L}$ dans $(V,(\cdot,\cdot)_V)$. ∎

Remarque 5.10 Diagonalisation à partir d'une base quelconque : voir le paragraphe 6 (valeurs propres généralisées) et le théorème 6.8 ; et le paragraphe 6.6 pour un exemple dans $V=P_1$. ∎

5.4 Matrice symétrique : base orthogonale commune de diagonalisation

Pour A symétrique réelle, avec les notations du théorème 5.2, on a obtenu, pour $i,j=1,...,n$:

$$(A.\vec{v}_j,\vec{v}_i)_{\mathbb{R}^n}=\lambda_j(\vec{v}_j,\vec{v}_i)_{\mathbb{R}^n}=\lambda_j\delta_{ij}, \tag{5.11}$$

$(\vec{v}_i)$ étant une b.o.n. de diagonalisation associée aux valeurs propres λ_i.

En particulier, si A est une matrice symétrique réelle définie positive, alors :

$$(\cdot,\cdot)_A:\begin{cases}\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}\\ \quad(\vec{x},\vec{y})\longrightarrow(\vec{x},\vec{y})_A=(A.\vec{x},\vec{y})_{\mathbb{R}^n},\end{cases} \tag{5.12}$$

est un produit scalaire sur $\mathbb{R}^n$. Et (5.11) se lit, pour tout i,j :

$$(\vec{v}_j,\vec{v}_i)_A=\lambda_j\delta_{ij}, \tag{5.13}$$

donc $(\vec{v}_i)$ est une base orthogonale dans $(\mathbb{R}^n,(\cdot,\cdot)_A)$ (dans $\mathbb{R}^n$ muni du produit scalaire $(\cdot,\cdot)_A$).

Et donc $(\vec{v}_i)$ est une base orthogonale à la fois pour les produits scalaires $(\cdot,\cdot)_{\mathbb{R}^n}$ et $(\cdot,\cdot)_A$: c'est une base orthogonale commune de diagonalisation.

Remarque 5.11 Les λ_j étant tous non nuls pour A définie positive, on a :

$$((A.\frac{\vec{v}_j}{\sqrt{\lambda_j}},\frac{\vec{v}_i}{\sqrt{\lambda_i}})_{\mathbb{R}^n}=)\qquad(\frac{\vec{v}_j}{\sqrt{\lambda_j}},\frac{\vec{v}_i}{\sqrt{\lambda_i}})_A=\delta_{ij}\qquad(=(\vec{v}_j,\vec{v}_i)_{\mathbb{R}^n}).$$

(C'est nul si $i\neq j$ et cela vaut 1 pour $i=j$.) Et donc $(\frac{\vec{v}_i}{\sqrt{\lambda_i}})$ est une b.o.n. dans $(\mathbb{R}^n,(\cdot,\cdot)_A)$. ∎

5.5 Diagonalisation simultanée de matrices symétriques qui commutent

Proposition 5.12 *Soit A et M sont deux matrices réelles symétriques qui commutent :*

$$A^T=A,\qquad M^T=M,\qquad A.M=M.A. \tag{5.14}$$

Alors $A.M$ est symétrique réelle et A et M sont diagonalisables dans une même b.o.n. $(\vec{v}_i)$ de $(\mathbb{R}^n,(\cdot,\cdot)_{\mathbb{R}^n})$: il existe n réels λ_i, n réels μ_i et une b.o.n. $(\vec{v}_i)$ de $\mathbb{R}^n$ tels que pour tout $i=1,...,n$:

$$A.\vec{v}_i=\lambda_i\vec{v}_i,\quad M.\vec{v}_i=\mu_i\vec{v}_i,\quad \textit{et}\quad(\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n}=0\quad\forall i,j. \tag{5.15}$$

Preuve. Comme $(A.M)^T=M^T.A^T=M.A=A.M$, la matrice $A.M$ est bien symétrique.

Montrons que A et M sont diagonalisable dans une même b.o.n..

1- Traitons donc le cas $A=D=\text{diag}(\lambda_1,...,\lambda_n)$ et M symétrique t.q. $D.M=M.D$. Comme $M.D$ est symétrique, $M.D$ est diagonalisable.

Quitte à changer l'ordre des λ_i, on les prend t.q. $D=\text{diag}(\mu_1 I_1,...,\mu_{n_v}I_{n_v})$, où n_v est le nombre de valeurs propres distinctes, les μ_i étant les valeurs propres 2 à 2 distinctes de multiplicité m_i, et I_{m_i} étant la matrice identité de $\mathbb{R}^{m_i}$. (Une base de diagonalisation de D est la base canonique.)

Comme M commute avec D, on utilise la proposition 4.33 : on a $M = \mathrm{diag}(M_1, ..., M_{n_v})$ matrice diagonale par blocs où M_i est une matrice carrée $m_i \times m_i$.

Et donc $D.M = \mathrm{diag}(\mu_1 M_1, ..., \mu_{n_v} M_{n_v})$ (immédiat), et si $(\vec{b}_i)$ est une b.o.n. de diagonalisation de M associée aux valeurs propres μ_i, celle-ci est constituée de b.o.n. de $\mathbb{R}^{m_i}$ de diagonalisation des M_i complétées par des zéros pour se plonger dans $\mathbb{R}^n$.

Comme M est symétrique, toutes les matrices M_i sont symétriques (immédiat), donc diagonalisables. Donc si α_{ij} pour $j = 1, ..., m_i$ sont les valeurs propres de M_i, les μ_i (valeurs propres de M) sont des α_{ij}, et les valeurs propres de $D.M$ sont les $\mu_i \alpha_{ij}$.

Et $(\vec{b}_i)$ est aussi une b.o.n. de diagonalisation de D, car toute b.o.n. dans $\mathbb{R}^{m_i}$ est une b.o.n. de diagonalisation de I_i (en particulier vrai pour la b.o.n. de diagonalisation de M_i).

Donc M et D et $M.D = D.M$ sont diagonalisables dans une même b.o.n..

2- Cas général A symétrique réelle. A étant symétrique réelle, A est diagonalisable dans une b.o.n. $(\vec{v}_i)$, avec $D = P^{-1}.A.P$, avec les notations du théorème 5.2, avec donc $P^{-1} = P^T$. Effectuons un changement de b.o.n. pour se placer dans la b.o.n. $(\vec{v}_i)$ de diagonalisation de A.

La matrice M devient $N = P^{-1}.M.P = P^T.M.P$, et N est également symétrique (car M l'est), et on a $D.N = P^{-1}.A.P.P^{-1}.M.P = P^{-1}.A.M.P$ et $N.D = P^{-1}.M.P.P^{-1}.A.P = P^{-1}.M.A.P$.

Comme $M.A = A.M$ on a $N.D = D.N$, et on s'est ramené au cas de $A = D$ matrice diagonale : les matrices D, N et $N.D$ sont diagonalisables dans une même b.o.n. $(\vec{f}_i)$ qui est la base de diagonalisation de N : $D.\vec{f}_i = \lambda_i \vec{f}_i$ et $N.\vec{f}_i = \mu_i \vec{f}_i$. Donc par changement de base inverse, A, M et $A.M$ sont diagonalisables dans la b.o.n. $(\vec{v}_i = P.\vec{f}_i)$ (expression dans la base initiale qui est la base canonique de $\mathbb{R}^n$) : $P.D.P^{-1}.\vec{v}_i = \lambda_i P.P^{-1}.\vec{v}_i$ et $P.N.P^{-1}.\vec{v}_i = \mu_i P.P^{-1}.\vec{v}_i$, soit $A.\vec{v}_i = \lambda_i \vec{v}_i$ et $M.\vec{v}_i = \mu_i \vec{v}_i$. ▪

Exercice 5.13 $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $M = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$: $A.M = M.A$, et une base commune de diagonalisation est donnée par $\vec{v}_1 = \vec{v}_1$ et $(\vec{v}_2, \vec{v}_3)$ base de diagonalisation de $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. ▪

5.6 Application : résolution du système différentiel $\vec{\alpha}\,' + M.\vec{\alpha} = \vec{f}$

On souhaite résoudre le système différentiel de Cauchy (avec condition initiale) : trouver $\vec{\alpha}$: $[0, T] \to \mathbb{R}^n$ t.q. :

$$\frac{d\vec{\alpha}}{dt}(t) + M.\vec{\alpha}(t) = \vec{f}(t), \qquad \vec{\alpha}(0) = \vec{\alpha}_0, \tag{5.16}$$

où M est une matrice symétrique, $\vec{f} : \mathbb{R} \to \mathbb{R}^n$ est une fonction continue, et $\vec{\alpha}_0 \in \mathbb{R}^n$ est un vecteur donné (condition initiale).

Notation implicite : on s'est placé dans la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$ avec :

$$\vec{\alpha}(t) = \sum_{i=1}^n \alpha_i(t) \vec{E}_i, \quad [\vec{\alpha}(t)] = \begin{pmatrix} \alpha_1(t) \\ \vdots \\ \alpha_n(t) \end{pmatrix}, \qquad f(t) = \sum_{i=1}^n f_i(t) \vec{E}_i, \quad [\vec{f}(t)] = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \tag{5.17}$$

les matrices colonnes $[\vec{\alpha}(t)]$ et $[\vec{f}(t)]$ donnant les composantes dans la base canonique. Et (5.16) a le sens :

$$\frac{d[\vec{\alpha}]}{dt}(t) + M.[\vec{\alpha}(t)] = [\vec{f}(t)], \qquad [\vec{\alpha}(0)] = [\vec{\alpha}_0], \tag{5.18}$$

les inconnues étant les n fonctions $\alpha_i : [0, T] \to \mathbb{R}$. Ces fonction sont liées les unes aux autres par (5.18) :

$$\left\{ \begin{array}{l} \dfrac{d\alpha_1}{dt}(t) + M_{11}\alpha_1(t) + ... + M_{1n}\alpha_n(t) = f_1(t), \\ \vdots \\ \dfrac{d\alpha_n}{dt}(t) + M_{n1}\alpha_1(t) + ... + M_{nn}\alpha_n(t) = f_n(t), \end{array} \right\} \quad \text{et} \quad \forall i = 1, ..., n, \quad \alpha_i(0) = (\vec{\alpha}_0)_i. \tag{5.19}$$

On dit que c'est un système différentiel couplé.

Une des méthodes classiques pour résoudre ce problème consiste à trouver une b.o.n. de diagonalisation de M, et à travailler dans cette base pour “découpler le problème (5.19)” :

Démarche : M étant symétrique soit $(\vec{v}_i)$ une b.o.n. de diagonalisation. Notons β_i les composantes de $\vec{\alpha}$ sur la b.o.n $(\vec{v}_i)$ de diagonalisation :

$$\vec{\alpha}=\sum_i \beta_i\vec{v}_i, \qquad [\vec{\alpha}]_{|(\vec{v})}=\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} \overset{\text{noté}}{=} [\vec{\beta}]. \tag{5.20}$$

Formules de changement de base, avec $P=\big([\vec{v}_1]_{|\vec{E}},...,[\vec{v}_n]_{|\vec{E}}\big)$ la matrice de changement de base de $(\vec{E}_i)$ vers $(\vec{v}_i)$:

$$[\vec{\beta}]=P^{-1}.[\vec{\alpha}], \tag{5.21}$$

Le problème (5.18) s'écrit, puisque $[\vec{\alpha}]=P.[\vec{\beta}]$: trouver $[\beta]$ t.q. :

$$P.\frac{d[\vec{\beta}]}{dt}(t)+M.P.[\vec{\beta}](t)=[\vec{f}](t), \qquad [\vec{\beta}](0)=P^{-1}.[\vec{\alpha}_0]. \tag{5.22}$$

D'où, multipliant par P^T, ayant $P^T.P=I$ et $P^T.M.P=D$:

$$\frac{d[\vec{\beta}]}{dt}(t)+D.[\vec{\beta}](t)=P^T.[\vec{f}](t), \qquad [\vec{\beta}](0)=P^T.[\vec{\alpha}_0]. \tag{5.23}$$

On a obtenu un système différentiel totalement découplé : pour tout $i=1,...,n$:

$$\beta_i{}'(t)+\lambda_i\beta_i(t)=(P^T.[\vec{f}(t)])_i, \qquad \beta_i(0)=(P^T.[\vec{\alpha}_0])_i. \tag{5.24}$$

Ces n équations différentielles indépendantes entre elles peuvent être intégrées de manière classique (méthode de variation de la constante).

D'où les α_i données par le calcul simple $[\vec{\alpha}]=P.[\vec{\beta}]$.

5.7 Application : recherche de b.o.n. dans $(P_1,(\cdot,\cdot)_{L^2})$

5.7.1 Rappel : l'espace $L^2([a,b])$

$L^2([a,b];\mathbb{R}) \overset{\text{noté}}{=} L^2([a,b])$ est l'ensemble des fonctions réelles d'énergie finie :

$$L^2([a,b])=\{f:[a,b]\to\mathbb{R},\ \int_a^b f(x)^2\,dx<\infty\}. \tag{5.25}$$

$L^2([a,b])$ est un espace vectoriel qu'on munit usuellement du produit scalaire et de la norme associées définis par :

$$(f,g)_{L^2}=\int_a^b f(x)g(x)\,dx, \qquad ||f||^2_{L^2}=(f,f)_{L^2}=\int_a^b f(x)^2\,dx. \tag{5.26}$$

Remarque 5.14 $L^2([a,b])$ est un espace vectoriel de dimension infinie qui n'a pas de base canonique : $L^2([a,b])$ ne s'écrit pas sous la forme d'un produit cartésien. Et donc $L^2([a,b])$ n'a pas de produit scalaire usuel. ▪

5.7.2 L'espace P_1

On se donne $N+1$ points $x_i\in\mathbb{R}$ pour $i=0,...,N$ où $x_0=a$, $x_N=b$ et $x_{i-1}<x_i$ pour tout i. Et on “maille” alors l'intervalle $[a,b]$ en les N intervalles :

$$[a,b]=\bigcup_{i=1}^N [x_{i-1},x_i]. \tag{5.27}$$

On note P_1 l'ensemble des fonctions continues sur $[a,b]$ et affines sur chaque $[x_{i-1},x_i]$:

$$P_1=\{f_h\in C^0([a,b];\mathbb{R})\ :\forall i=1,...,N,\ f_{h|[x_{i-1},x_i]} \text{ est affine}\}. \tag{5.28}$$

P_1 est sous-espace vectoriel de $L^2([a,b])$: immédiat.

Et P_1 est un espace vectoriel de dimension $N{+}1$, une fonction $f \in P_1$ est complètement et uniquement déterminée si on connaît les $N{+}1$ valeurs $f(x_i)$. La base $(\varphi_i)_{i=0,...,N}$ usuelle de P_1 est donnée par les fonctions chapeau $\varphi_i \in P_1$ définies par (voir cours d'éléments finis) :

$$\varphi_i \in P_1 \qquad \text{t.q.} \qquad \forall i,j=0,...,N, \quad \varphi_i(x_j)=\delta_{ij}. \tag{5.29}$$

Faire un dessin. Ainsi une fonction $f_h \in P_1$ s'écrit de manière unique :

$$f_h = \sum_{i=0}^{N} f_h(x_i)\varphi_i, \tag{5.30}$$

où donc les $c_i = f_h(x_i)$ sont les composantes de f sur la base (φ_i). Autrement dit, pour tout $x \in [a,b]$:

$$f_h(x) = \sum_{i=0}^{N} f_h(x_i)\varphi_i(x). \tag{5.31}$$

Remarque 5.15 La base (φ_i) est très simple, mais n'a rien de canonique (l'espace P_1 n'est pas un produit cartésien $K \times ... \times K$ où K est un corps). ▪

Remarque 5.16 La base (φ_i) n'est pas orthogonale pour le produit scalaire $(\cdot,\cdot)_{L^2}$. Par exemple $(\varphi_1,\varphi_2)_{L^2} > 0$ puisque $\varphi_1\varphi_2 > 0$ sur $]x_1,x_2[\neq \emptyset$, et $\varphi_1\varphi_2$ est nulle ailleurs. ▪

5.7.3 Approximation P_1

Soit $f \in L^2([a,b])$. Le but est d'approximer f par une fonction $f_h \in P_1$, et ce "le mieux possible au sens de l'énergie", i.e. t.q. :

$$||f-f_h||_{L^2} = \min_{g_h \in P_1} ||f-g_h||_{L^2}, \tag{5.32}$$

f_h réalisant le minimum. On cherche donc $f_h \in P_1$ t.q. $f - f_h \perp V_h$ (orthogonalité au sens de $(\cdot,\cdot)_{L^2}$), soit :

$$\forall g_h \in P_1, \qquad (f-f_h, g_h)_{L^2} = 0, \tag{5.33}$$

soit :

$$\forall g_h \in P_1, \qquad (f_h, g_h)_{L^2} = (f, g_h)_{L^2}. \tag{5.34}$$

L'approximation f_h de f permet de stocker uniquement $N{+}1$ valeurs de f ; et on calcule avec f_h au lieu de f.

5.7.4 Le problème matriciel dans $\mathbb{R}^n$

Disposant de la base $(\varphi_i)_{i=0,...,N}$ de l'espace vectoriel P_1, le problème (5.34) équivaut à (système de $n = N{+}1$ lignes) :

$$\forall i=0,...,N, \qquad (f_h,\varphi_i)_{L^2} = (f,\varphi_i)_{L^2}, \tag{5.35}$$

puisqu'un produit scalaire est linéaire par rapport à la deuxième variable. Et trouver f_h c'est trouver ses $n = N{+}1$ composantes c_j sur la base (φ_i) :

$$f_h = \sum_{j=0}^{N} c_j\,\varphi_j, \qquad [f_h]_{|(\varphi_i)} = \begin{pmatrix} c_0 \\ \vdots \\ c_N \end{pmatrix}, \tag{5.36}$$

où donc $[f_h]_{|(\varphi_i)}$ est le "vecteur colonne" donne les composantes relativement à la base (φ_i) de P_1.

On obtient un système de n lignes à $n = N{+}1$ inconnues :

$$\forall i=0,...,N, \qquad \sum_{j=0}^{N} c_j\,(\varphi_j,\varphi_i)_{L^2} = (f,\varphi_i)_{L^2}. \tag{5.37}$$

Autrement écrit, les c_i sont solutions du problème matriciel :

$$M.[\vec{c}] = [\vec{b}], \tag{5.38}$$

où :

$$M=[M_{ij}]=[(\varphi_j,\varphi_i)_{L^2}],\qquad [\vec{b}]=\begin{pmatrix} b_0=(f,\varphi_0)_{L^2}\\ \vdots \\ b_N=(f,\varphi_n)_{L^2}\end{pmatrix},\qquad [\vec{c}]=\begin{pmatrix} c_0\\ \vdots\\ c_N\end{pmatrix}. \tag{5.39}$$

M est matrice de masse, donc inversible, donc (5.38) a une unique solution, trouvée par exemple avec une méthode comme Cholesky.

5.7.5 Diagonalisation dans $\mathbb{R}^n$

M est symétrique réelle, donc est diagonalisable dans une b.o.n. de $\mathbb{R}^n$: notons λ_i les valeurs propres et $(\vec{v}_i)$ la b.o.n. de $(\mathbb{R}^n,(\cdot,\cdot)_{\mathbb{R}^n})$ de vecteurs propres associés :

$$\forall j=0,...,N,\quad M.\vec{v}_j=\lambda_j\vec{v}_j,\qquad D=P^T.M.P,\qquad P^T.P=I, \tag{5.40}$$

où $D=\mathrm{diag}(\lambda_1,...,\lambda_n)$ est la matrice diagonale des valeurs propres, et $P=[P_{ij}]=(\,[\vec{v}_1],...,[\vec{v}_n]\,)$ est la matrice stockant dans sa j-ème colonne les composantes de $\vec{v}_j$ dans la base canonique :

$$\vec{v}_j=\sum_i P_{ij}\vec{E}_i,\qquad [\vec{v}_j]=\begin{pmatrix}P_{1j}\\ \vdots\\ P_{nj}\end{pmatrix}. \tag{5.41}$$

5.7.6 Retour dans P_1 : b.o.n dans $(P_1,(\cdot,\cdot)_{L^2})$

Dans P_1, soit χ_j la fonction de P_1 dont les composantes dans la base (φ_i) de P_1 sont celles de $\vec{v}_j$ dans la base canonique de $\mathbb{R}^n$, cf. (5.41) :

$$\chi_j=\sum_i P_{ij}\varphi_i,\qquad [\chi_j]_{|(\varphi_i)}=[\vec{v}_j]=\begin{pmatrix}P_{1j}\\ \vdots\\ P_{nj}\end{pmatrix}. \tag{5.42}$$

On vérifie immédiatement que (χ_i) est une base orthogonale de $(P_1,(\cdot,\cdot)_{L^2})$:

$$(\chi_i,\chi_j)_{L^2}=\sum_{k\ell}P_{ki}P_{\ell j}(\varphi_k,\varphi_\ell)_{L^2}=\sum_{k\ell}P_{ki}P_{\ell j}M_{k\ell}=(P^T.M.P)_{ij}=D_{ij}=\lambda_i\delta_{ij}, \tag{5.43}$$

Posons $\psi_i=\dfrac{\chi_i}{\sqrt{\lambda_i}}$ (les λ_i sont tous >0 car M est définie positive) : on a donc :

$$(\psi_i,\psi_j)_{L^2}=\delta_{ij}. \tag{5.44}$$

(C'est bien nul si $i\neq j$ et cela vaut bien 1 si $i=j$.) Donc $(\psi_i)_{i=0,...,N}$ est une b.o.n. de $(P_1,(\cdot,\cdot)_{L^2})$.

5.7.7 Calcul avec la b.o.n. de P_1 trouvée

Il est plus plus simple de chercher f_h, solution de (5.34), à l'aide de la b.o.n. calculée (ψ_i) : si on note α_i les composantes de f_h dans la b.o.n. (ψ_i) :

$$f_h=\sum_{j=0}^N\alpha_j\psi_j,\qquad [f_h]_{|(\psi_i)}=\begin{pmatrix}\alpha_0\\ \vdots\\ \alpha_n\end{pmatrix}, \tag{5.45}$$

alors (5.34) équivaut à :

$$\forall i=0,...,N,\qquad (f_h,\psi_i)_{L^2}=(f,\psi_i)_{L^2}, \tag{5.46}$$

soit avec (5.45), ayant $(\psi_i,\psi_j)=\delta_{ij}$:

$$\alpha_i=(f,\psi_i)_{L^2},\quad \forall i=0,...,N. \tag{5.47}$$

Et donc les α_i sont donnés explicitement (contrairement aux c_i dans (5.38) qui sont donnés implicitement : nécessite la résolution du système matriciel (5.38)).

5.8 Loi d'inertie de Sylvester

5.8.1 Forme quadratique

Définition 5.17 Une forme quadratique est une application $Q : \mathbb{R}^n \to \mathbb{R}$ telle qu'il existe une forme bilinéaire symétrique $a(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ vérifiant, pour tout $\vec{x} \in \mathbb{R}^n$:

$$Q(\vec{x}) = a(\vec{x},\vec{x}). \tag{5.48}$$

Proposition 5.18 *Soit Q la forme quadratique donnée par (5.48) : on a l'identité de polarisation : pour tout $\vec{x},\vec{y}$:*

$$a(\vec{x},\vec{y}) = \frac{1}{4}(Q(\vec{x}+\vec{y}) - Q(\vec{x}-\vec{y})). \tag{5.49}$$

Donc il existe une unique forme bilinéaire symétrique $a(\cdot,\cdot)$ telle que $Q(\vec{x}) = a(\vec{x},\vec{x})$.

Preuve. Soit Q définie par (5.48). 0n a, pour tout $\vec{x},\vec{y}$, $a(\vec{x}+\vec{y},\vec{x}+\vec{y}) = a(\vec{x},\vec{x})+a(\vec{x},\vec{y})+a(\vec{y},\vec{x})+a(\vec{y},\vec{y}) = a(\vec{x},\vec{x})+2a(\vec{x},\vec{y})+a(\vec{y},\vec{y})$ par symétrie de $a(\cdot,\cdot)$, et $a(\vec{x}-\vec{y},\vec{x}-\vec{y}) = a(\vec{x},\vec{x})-2a(\vec{x},\vec{y})+a(\vec{y},\vec{y})$ par symétrie de $a(\cdot,\cdot)$. Donc, pour tout $\vec{x},\vec{y}$:

$$\frac{1}{4}(Q(\vec{x}+\vec{y}) - Q(\vec{x}-\vec{y})) = \frac{1}{4}(a(\vec{x}+\vec{y},\vec{x}+\vec{y}) - a(\vec{x}-\vec{y},\vec{x}-\vec{y})) = \frac{1}{4}(4a(\vec{x},\vec{y})) = a(\vec{x},\vec{y}), \tag{5.50}$$

et $a(\cdot,\cdot)$ vérifie bien (5.49). Et si $b(\cdot,\cdot)$ est une autre forme bilinéaire symétrique, alors de même $b(\cdot,\cdot)$ satisfait l'identité de polarisation, donc $b(\cdot,\cdot) = a(\cdot,\cdot)$. ∎

5.8.2 Matrice d'une forme quadratique

Calculs dans une base : on considère la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$. Soit $A = [a(\cdot,\cdot)]_{|(\vec{E})}$ la matrice de $a(\cdot,\cdot)$ dans cette base :

$$A = [A_{ij}] = [a(\cdot,\cdot)]_{|(\vec{a})}, \qquad \text{où} \qquad A_{ij} = a(\vec{E}_i,\vec{E}_j) \quad \forall i,j. \tag{5.51}$$

Définition 5.19 La matrice de Q dans la base $(\vec{E}_i)$ est la matrice de $a(\cdot,\cdot)$ dans la base $(\vec{E}_i)$. Et on note :

$$[Q]_{|(\vec{E})} \stackrel{\text{def}}{=} [a(\cdot,\cdot)]_{|(\vec{E})} \qquad (= A). \tag{5.52}$$

On obtient immédiatement avec (5.51) :

$$\vec{x} = \sum_{i=1}^{n} x_i\vec{E}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i,j=1}^{n} x_ix_ja(\vec{E}_i,\vec{E}_j) = \sum_{i,j=1}^{n} x_ix_jA_{ij} = [\vec{x}]^T.A.[\vec{x}], \tag{5.53}$$

et Q est polynôme de degré 2 en les variables $x_1,...,x_n$ (une forme quadratique).

5.8.3 Signature

Soit $a(\cdot,\cdot)$ une forme bilinéaire symétrique réelle. $A = [a(\cdot,\cdot)]_{|(\vec{E})}$ est symétrique réelle, donc est diagonalisable.

Définition 5.20 La signature de $a(\cdot,\cdot)$ (forme bilinéaire symétrique) est le couple $(s_+,s_-) \in \mathbb{N}^2$ où s_+ est le nombre de valeurs propres > 0 et s_- est le nombre de valeurs propres < 0 de $A = [a(\cdot,\cdot)]_{|(\vec{E})}$.

La signature de Q est la signature de $a(\cdot,\cdot)$.

Et le rang de Q est l'entier $r = s_+ + s_-$.

Proposition 5.21 *Soit $(\vec{v}_i)$ une b.o.n. de diagonalisation de $A = [a(\cdot,\cdot)]_{|(\vec{E})}$. Quitte à réordonner cette base, on suppose $\lambda_1,...,\lambda_{s_+} > 0$ et $\lambda_{s_++1},...,\lambda_{s_++s_-} < 0$ (avec donc $\lambda_i = 0$ pour tout*

$i > s_+ + s_-$). Alors dans la b.o.n. $(\vec{v}_i)$ la forme quadratique Q prend la "forme diagonale" :

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{v}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i=1}^{s_+} \lambda_i x_i^2 - \sum_{i=s_+ +1}^{s_+ + s_-} (-\lambda_i) x_i^2. \tag{5.54}$$

Et dans la base orthogonale $(\vec{w}_i) = (\frac{\vec{v}_i}{\sqrt{\lambda_i}})$ on a :

$$\vec{x} = \sum_{i=1}^{n} y_i \vec{w}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i=1}^{s_+} y_i^2 - \sum_{i=s_+ =1}^{s_+ + s_-} y_i^2. \tag{5.55}$$

Ainsi :

$$[a(\cdot,\cdot)]_{|(\vec{w})} = \begin{pmatrix} I_{s_+} & 0 & 0 \\ 0 & I_{s_-} & 0 \\ 0 & 0 & 0 \end{pmatrix} = Z^T.[a(\cdot,\cdot)]_{|(\vec{E})}.Z, \tag{5.56}$$

où $Z = (\,[\vec{w}_1] \quad ... \quad [\vec{w}_n]\,)$ est la matrice de passage de la base canonique $(\vec{E}_i)$ à la base $(\vec{w}_i)$.

Preuve. $A = P.D.P^{-1}$ où $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$, $P = (\,[\vec{v}_1] \quad ... \quad [\vec{v}_n]\,)$ avec $P^T.P = I$, et où on a réordonné les valeurs propres pour avoir $\lambda_i > 0$ pour $i \in [1, s_+]_{\mathbb{N}}$, $\lambda_i < 0$ pour $i \in [s_+ +1, s_+ + s_-]_{\mathbb{N}}$, et $\lambda_i = 0$ sinon. Et $\vec{x} = \sum_{i=1}^n x_i \vec{v}_i$ donne $a(\vec{x}, \vec{x}) = \sum_{i,j=1}^n x_i x_j a(\vec{v}_i, \vec{v}_j) = \sum_{i,j=1}^n x_i x_j \lambda_i \delta_{ij} = \sum_{i=1}^n x_i^2 \lambda_i$, d'où (5.54), d'où (5.55), d'où (5.56). ∎

Remarque 5.22 Pour les matrices hermitiennes, il n'y a pas de signature : comme $i^2 = -1$, une forme quadratique peut se décomposer comme un somme de carrés. ∎

5.8.4 "Expressions diagonales" d'une forme quadratique

Définition 5.23 La forme quadratique Q est sous forme "d'expression diagonale" relativement à une base $(\vec{e}_i)$ ssi :

$$\vec{x} = \sum_{i=1}^{n} y_i \vec{e}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i=1}^{n} \alpha_i y_i^2, \tag{5.57}$$

où les α_i sont des réels.

Exemple 5.24 Dans une b.o.n. $(\vec{v}_i)$ de diagonalisation de $A = [a(\cdot,\cdot)]_{|(\vec{E})}$, Q a l'expression diagonale donnée par (5.54). ∎

Exercice 5.25 Dans $\mathbb{R}^2$ muni de sa base canonique. On note $\vec{x} = x_1 \vec{E}_1 + x_2 \vec{E}_2$. Soit :

$$Q\vec{x}) = x_1^2 + 2b x_1 x_2 + c x_2^2. \tag{5.58}$$

1- Calculer la signature de Q. Décrire la base dans laquelle Q est donnée par (5.54) :

$$Q(\vec{x}) = \lambda_1 y_1^2 + \lambda_2 y_2^2. \tag{5.59}$$

2- À l'aide de la "méthode de réduction de Gauss", mettre Q sous la forme :

$$Q(\vec{x}) = X_1^2 + \beta X_2^2, \tag{5.60}$$

où X_1 et X_2 sont fonctions de x_1 et x_2. Donner la base dans laquelle dans laquelle les coordonnées de $\vec{x}$ sont X_1 et X_2. Donner la matrice de $a(\cdot,\cdot)$ dans cette base. Vérifier que la signature de Q est $(2,0)$ si $\beta > 0$, est $(1,0)$ si $\beta = 0$, et est $(1,1)$ si $\beta < 0$.

3- Donner deux bases dans lesquelles Q à une "expression diagonale".

Réponse. 1- $A = [Q]_{|(\vec{E})} = \begin{pmatrix} 1 & b \\ b & c \end{pmatrix}$. Valeurs propres : $p(\lambda) = \lambda^2 - (1+c)\lambda + c - b^2$. Discriminant $\Delta = (1+c)^2 + 4(b^2 - c)$ (donc $= 4b^2 + (1-c)^2 \geq 0$ car le discriminant est positif, les valeurs propres d'une matrice symétrique réelle étant réelles). Et les valeurs propres sont données par $\lambda = \frac{(1+c) \pm \sqrt{(1+c)^2 + 4(b^2-c)}}{2}$. Donc :

si $c = b^2$ alors une racine est nulle et l'autre vaut $\frac{1}{2}(1 + c + |1+c|) = 1 + b^2 > 0$ et la signature vaut $(1,0)$,

si $c > b^2$ alors les deux racines sont de signe opposées et la signature vaut $(1,-1)$
si $c < b^2$ alors les deux racines sont de même signe et la signature vaut $(2,0)$.

Une b.o.n. $(\vec{v}_1,\vec{v}_2)$ de vecteurs propres associés donne la matrice de passage $P = ([\vec{v}_1] \; ... \; [\vec{v}_2])$. Dans cette b.o.n. la matrice de A est la matrice $D = P^T.A.P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, donc si $\vec{x} = y_1\vec{v}_1 + y_2\vec{v}_2$ on a $Q(\vec{x}) = \lambda_1 y_1^2 + \lambda_2 y^2$.

2- Méthode de réduction de Gauss : on a (termes contenant x_1) : $x_1^2 + 2bx_1x_2 = (x_1+bx_2)^2 - b^2x_2^2$. D'où $Q(\vec{x}) = (x_1+bx_2)^2 + (c-b^2)x_2 = X_1^2 + \beta X_2^2$, où $X_1 = x_1 + bx_2$, où $X_2 = x_2$, et où $\beta = (c-b^2)$.

Donc $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = P^{-1}.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Donc $P = \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix}$. Soit $\vec{e}_1 = \vec{E}_1$ et $\vec{e}_2 = -b\vec{E}_1 + \vec{E}_2$: c'est dans la base $(\vec{e}_1,\vec{e}_2)$ que Q à la forme (5.60), où donc $\vec{x} = X_1\vec{e}_2 + X_2\vec{e}_2$.

La matrice $B = [Q]_{|(\vec{e})} = [a(\cdot,\cdot)]_{|(\vec{e})}$ est donnée par $B = P^T.A.P = \begin{pmatrix} 1 & 0 \\ 0 & c-b^2 \end{pmatrix}$.

Et si $c = b^2$ alors $Q(\vec{x}) = X_1^2$: ce cas correspond bien à la signature $(1,0)$. Si $c > b^2$ alors $\beta > 0$: ce cas correspond bien à la signature $(2,0)$. Si $c < b^2$ alors $\beta < 0$: ce cas correspond bien à la signature $(1,-1)$.

3- Donc dans la b.o.n. de diagonalisation $(\vec{v}_1,\vec{v}_2)$ la forme quadratique Q à l'expression diagonale $Q(\vec{x}) = \lambda_1 y_1^2 + \lambda_2 y^2$, et dans la base $(\vec{e}_1,\vec{e}_2)$, qui n'est pas orthogonale quand $b \neq 0$, la forme quadratique Q à l'expression diagonale $Q(\vec{x}) = X_1^2 + \beta X_2^2$. ▪

5.8.5 Conservation de la signature (Sylvester)

Théorème 5.26 *(Loi d'inertie de Sylvester.) On considère (5.55) exprimée dans la base $(\vec{w}_i)$.*

Soit $(\vec{f}_i)$ une autre base de $\mathbb{R}^n$ telle que :

$$\vec{x} = \sum_{i=1}^{n} y_i\vec{u}_i \quad \Longrightarrow \quad Q(\vec{y}) = \sum_{i=1}^{t_+} y_i^2 - \sum_{i=t_+=1}^{t_++t_-} y_i^2, \tag{5.61}$$

où $t_+, t_- \in \mathbb{N}$. Alors $t_+ = s_+$ et $t_- = s_-$.

Et donc la signature (s_+, s_-) de la forme quadratique est indépendante de la base dans laquelle Q prend la forme (5.61). (Voir aussi § 6.7 et § 6.7.2).

Preuve. Par définition de la base $\vec{e}_i$ on a $Q(\vec{e}_i) > 0$ pour tout $i \in [1, s_+]_{\mathbb{N}}$. Par définition de la base $\vec{u}_i$ on a $Q(\vec{u}_i) \le 0$ pour tout $i \in [t_++1, n]_{\mathbb{N}}$. Montrons que la famille $(\vec{e}_1, ..., \vec{e}_{s_+}, \vec{u}_{t_++1}, ..., \vec{u}_n)$ est libre. Si $\sum_{i=1}^{s_+} \alpha_i\vec{e}_i + \sum_{i=t_++1}^{n} \beta_i\vec{u}_i = \vec{0}$, alors notons $\vec{z} \stackrel{\text{def}}{=} \sum_{i=1}^{s_+} \alpha_i\vec{e}_i$, et donc on a aussi $\vec{z} = -\sum_{i=t_++1}^{n} \beta_i\vec{u}_i$. Donc d'une part $Q(\vec{z}) = \sum_{i=1}^{s_+} \alpha_i^2 Q(\vec{e}_i) \ge 0$ (par définition de s_+) et d'autre part $Q(\vec{z}) = \sum_{i=t_++1}^{n} \beta_i^2 Q(\vec{u}_i) \le 0$ (par définition de t_+). Donc $Q(\vec{z}) = 0$. Comme les $Q(\vec{e}_i) > 0$ pour $i = 1, ..., s_+$, on a donc $\alpha_i = 0$ pour $i = 1, ..., s_+$. Dons $\vec{z} = 0$. Donc $\sum_{i=t_++1}^{n} \beta_i\vec{u}_i = \vec{0}$. Donc $\beta_i = 0$ pour $i = t_+ + 1, ..., n$ (car $(\vec{u}_i)$ est une base).

Donc $s_+ + (n - t_+) \le n$. Donc $s_+ \le t_+$. Quitte à inverser les rôles de $(\vec{e}_i)$ et $(\vec{u}_i)$ on a donc aussi $t_+ \le s_+$. Donc $s_+ = t_+$. De même (considérer $-Q$), $s_- = t_-$. ▪

5.9 Racine carrée d'une matrice symétrique définie positive

5.9.1 Racines carrées d'une matrice symétrique

Définition 5.27 Soit $A \in \mathbb{R}^{n^2}$. S'il existe une matrice S telle que $A = S^2$, on dit que S est une racine carrée de la matrice A (Square root of A).

Exemple 5.28 Soit $S = \begin{pmatrix} 1 & a \\ 0 & -1 \end{pmatrix}$ où $a \in \mathbb{R}$, alors $S^2 = I$, et donc S est une racine carrée de I pour tout $a \in \mathbb{R}$. ▪

Exemple 5.29 La matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'a pas de racine carrée. En effet si $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors $S^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix}$, et donc $S^2 = A$ donne

$ab+bd = 1 = b(a+d)$, donc $a+d \neq 0$, et
$ca+dc = 0 = c(a+d)$ donc $c = 0$, et
$a^2+bc = 0 = a^2$ donc $a = 0$, et
$cb+d^2 = 0 = d^2$ donc $d = 0$, donc $1 = b(0+0) = 0$: absurde. ▪

Exemple 5.30 Donner un exemple de matrice $S \neq 0$ telle que $S^2 = 0$.

Réponse. $S = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$ ou encore $\begin{pmatrix} a & b \\ -\frac{a^2}{b} & -a \end{pmatrix}$ pour $b \neq 0$ (nécessairement $\det(S) = 0$). ▪

Exemple 5.31 La matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ n'a pas de racine carrée réelle, mais elle a des racines carrées complexes, à savoir $S = \begin{pmatrix} 1 & 0 \\ 0 & \pm i \end{pmatrix}$. ▪

Proposition 5.32 *Soit $A \in \mathbb{R}^{n^2}$. S'il existe une matrice $S \in \mathbb{R}^{n^2}$ symétrique réelle vérifiant $A = S^2$, alors A est symétrique positive.*

Et si S est réelle symétrique inversible alors A est réelle symétrique définie positive.

Preuve. Ici $S^T = S$ donc $A = S^2 = S^T.S$. Donc $A^T = A$ et A est symétrique. Et $(A.\vec{x},\vec{x})_{\mathbb{R}^n} = (S^T.S.\vec{x},\vec{x})_{\mathbb{R}^n} = (S.\vec{x},S.\vec{x})_{\mathbb{R}^n} = ||S.\vec{x}||^2_{\mathbb{R}^n} \geq 0$ pour tout $\vec{x}$, et A est bien positive.

Et si de plus S est inversible, alors $(A.\vec{x},\vec{x})_{\mathbb{R}^n} = ||S.\vec{x}||^2_{\mathbb{R}^n} > 0$ pour $\vec{x} \neq \vec{0}$ (car S injective), donc A est définie positive. ▪

Proposition 5.33 *Soit M une matrice réelle symétrique positive. Alors il existe une unique matrice réelle symétrique positive S telle que $M = S^2$.*

En particulier, si $D = \mathrm{diag}(\lambda_1,...,\lambda_n)$ est positive, alors :

$$\sqrt{D} \stackrel{\text{def}}{=} \mathrm{diag}(\sqrt{\lambda_1},...,\sqrt{\lambda_n}) \tag{5.62}$$

est la seule racine carrée de D qui est symétrique positive.

Et disposant de la formule de diagonalisation (5.3) ($M = P.D.P^{-1}$ avec $P^{-1} = P^T$), on a :

$$M = S^2, \qquad \text{où} \qquad S = P.\sqrt{D}.P^{-1}. \tag{5.63}$$

Preuve. Comme M est définie positive, on a $\lambda_i > 0$ pour tout i. D'où $\sqrt{D} = \mathrm{diag}(\sqrt{\lambda_1},...,\sqrt{\lambda_n})$ est bien une matrice réelle, et $(\sqrt{D})^2 = D$ (trivial). Et, ayant $P^{-1} = P^T$:

$$M = P.\sqrt{D}.\sqrt{D}.P^T = P.\sqrt{D}.P^T.P.\sqrt{D}.P^T, \tag{5.64}$$

d'où (5.63). Et $S^T = (P.\sqrt{D}.P^T)^T = P.\sqrt{D}.P^T = S$, et S est bien symétrique.

Unicité : si $S^2 = M$ alors $S^3 = S.M = M.S$: les matrices symétriques M et S commutent. Donc les matrices sont diagonalisables dans une même b.o.n., cf. proposition 5.12 : on a $M = P.D.P^{-1}$ et $S = P.d.P^{-1}$ où D et d sont diagonales. Donc $d = P^{-1}.S.P$ donne $d^2 = P^{-1}.M.P = D$. Comme S est positive on a d positive, et comme d et D sont diagonales, on a $d = \sqrt{D}$ et (5.63). ▪

5.9.2 Racine carrée d'une matrice de masse

On a vu qu'une matrice de masse réelle était une matrice symétrique définie positive.

Soit $(\vec{b}_i)$ une base de $\mathbb{R}^n$. Soit B_{ij} les composantes de $\vec{b}_j$ dans la base canonique :

$$\vec{b}_j = \sum_{i=1}^n B_{ij}\vec{E}_i = \begin{pmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{pmatrix}_{|(\vec{E})}, \qquad [\vec{b}_j] = \begin{pmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{pmatrix}, \qquad B = [B_{ij}] = \begin{pmatrix} [\vec{b}_1] & \dots & [\vec{b}_n] \end{pmatrix}. \tag{5.65}$$

Soit $M = [M_{ij}] \in \mathbb{R}^{n^2}$ la matrice de masse correspondante :

$$M_{ij} = (\vec{b}_i,\vec{b}_j)_{\mathbb{R}^n} = [\vec{b}_i]^T.[\vec{b}_j], \qquad \forall i,j = 1,...,n. \tag{5.66}$$

Proposition 5.34 *Soit B la matrice de passage de la base $(\vec{E}_i)$ à la base $(\vec{b}_i)$. On a :*

$$M = B^T.B, \tag{5.67}$$

Preuve. $(B^T.B)_{ij} = [\vec{b}_i]^T.[\vec{b}_j]$, donc (5.66) donne (5.67). ∎

Proposition 5.35 *Soit M donnée par (5.67). De (5.3) ($M = P.D.P^T$ car $P^{-1} = P^T$) on déduit :*

$$B = \sqrt{D}.P^T, \tag{5.68}$$

ainsi que :

$$M = S^2, \qquad \text{où} \qquad S = P.\sqrt{D}.P^T = P.B, \tag{5.69}$$

où S est la matrice symétrique racine carrée de M, cf. (5.63).

Ainsi les colonnes de S (ou les lignes de S par symétrie) donnent une seconde base $(\vec{s}_i)$ pour laquelle $M = [(\vec{s}_i, \vec{s}_j)_{\mathbb{R}^n}]$: M est aussi la matrice de masse relativement à la base $(\vec{s}_i)$.

Preuve. (5.68) et (5.69) sont immédiates. Et $M = S^T.S$, donc M est bien une matrice de masse relativement à la base $(\vec{s}_i)$, où $\vec{s}_j$ est le j-ème "vecteur colonne" de S. ∎

Exercice 5.36 Soit $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Donner les matrices B et S définies en (5.68) et (5.69), et les bases associées, et vérifier que $[\vec{s}_j] = P.[\vec{b}_j]$.

Réponse. On dispose de l'exercice 5.3. Donc $\sqrt{D} = \operatorname{diag}(\sqrt{2}, 2)$ et $B = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 2 \end{pmatrix} . \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ \sqrt{2} & \sqrt{2} \end{pmatrix}$. D'où $\vec{b}_1 = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$ et $\vec{b}_2 = \begin{pmatrix} -1 \\ \sqrt{2} \end{pmatrix}$. On vérifie que $B^T.B = \begin{pmatrix} 1 & \sqrt{2} \\ -1 & \sqrt{2} \end{pmatrix} . \begin{pmatrix} 1 & -1 \\ \sqrt{2} & \sqrt{2} \end{pmatrix} = M$.

Et $S = P.B = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} . \begin{pmatrix} 1 & -1 \\ \sqrt{2} & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 1+\frac{1}{\sqrt{2}} & 1-\frac{1}{\sqrt{2}} \\ 1-\frac{1}{\sqrt{2}} & 1+\frac{1}{\sqrt{2}} \end{pmatrix}$ est bien symétrique et vérifie $S^2 = M$. D'où $\vec{s}_1 = \begin{pmatrix} 1+\frac{1}{\sqrt{2}} \\ 1-\frac{1}{\sqrt{2}} \end{pmatrix}$ et $\vec{s}_2 = \begin{pmatrix} 1-\frac{1}{\sqrt{2}} \\ 1+\frac{1}{\sqrt{2}} \end{pmatrix}$. Et $P.\vec{b}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} . \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} 1+\frac{1}{\sqrt{2}} \\ 1-\frac{1}{\sqrt{2}} \end{pmatrix} = \vec{s}_1$, idem avec $\vec{b}_2$. ∎

6 Valeurs et vecteurs propres généralisés

On se place dans le cadre réel pour simplifier les écritures.

6.1 Le problème aux valeurs propres généralisé

Le problème aux valeurs propres généralisé est, étant donné R et M deux matrices $n*n$ avec M symétrique définie positive : trouver les couples $(\lambda, \vec{x}) \in \mathbb{R} \times \mathbb{R}^n$ tels que $\vec{x} \neq 0$ et :

$$\begin{cases} \text{trouver les couples } (\lambda, \vec{x}) \in \mathbb{R} \times \mathbb{R}^n - \{\vec{0}\} \text{ t.q. :} \\ R.\vec{x} = \lambda M.\vec{x}. \end{cases} \tag{6.1}$$

N.B. : on évitera de ramener le problème (6.1) au problème classique $A.\vec{x} = \lambda\vec{x}$ avec $A = M^{-1}.R$ car A n'est pas symétrique en général, même quand M et R le sont.

Définition 6.1 Un couple (valeur propre généralisé,vecteur propre généralisé) relativement à la matrice M est un couple $(\lambda, \vec{x}) \in \mathbb{R} \times \mathbb{R}^n - \{\vec{0}\}$ solution de (6.1).

Proposition 6.2 *λ est valeur propre généralisée ssi* $\det(R - \lambda M) = 0$.

Preuve. Immédiat (on rappelle que M est une matrice symétrique définie positive et p est bien un polynôme de degré n). ∎

Définition 6.3 Et le polynôme $p(\lambda) = \det(R - \lambda M)$ de degré n est appelé le polynôme caractéristique généralisé. Et si λ est racine de p de multiplicité m, alors λ est une valeur propre généralisé de multiplicité m.

Définition 6.4 La matrice R est M-diagonalisable ssi il existe n couples solutions $(\lambda_i, \vec{e}_i) \in \mathbb{R}\times\mathbb{R}^n$ de (6.1) tels que $(\vec{e}_i)_{i=1,\dots,n}$ est une base de $\mathbb{R}^n$.

Exemple 6.5 La matrice M est trivialement M-diagonalisable, avec $\lambda = 1$ la seule valeur propre généralisée de multiplicité n, toute base de $\mathbb{R}^n$ étant base de vecteurs propres généralisés.

Cela montre que l'ordre de multiplicité d'une valeur propre classique est en général différent de l'ordre de multiplicité d'une valeur propre généralisé (quand $M \neq I$). En effet, $M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ est telle que $\mu_1 = 1$ et $\mu_2 = 2$ sont les valeurs propres classiques de M de multiplicité 1, alors que $\lambda_1 = 1$ est valeur propre généralisée double de M. ▪

Exercice 6.6 Soit $R = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ et $M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Calculer les couples $(\lambda, \vec{e})$ (valeurs propres , vecteurs propres) généralisés de R, ainsi que les couples $(\mu, \vec{v})$ (valeurs propres,vecteurs propres) classiques de R.

Réponse. V.p. généralisées : on cherche $(\lambda, \vec{e}) \in \mathbb{R}\times\mathbb{R}^2$ tq $(R-\lambda M)\vec{e} = 0$ et $\vec{e} \neq 0$. La matrice $R-\lambda M$ est alors non inversible, i.e. son déterminant est nul : $\det \begin{pmatrix} 1-\lambda & 2 \\ 2 & 5-2\lambda \end{pmatrix} = 0$, soit $2\lambda^2 - 7\lambda + 1 = 0$. D'où $\lambda_1 = \frac{1}{4}(7-\sqrt{41}) \simeq 0.15$ et $\lambda_2 = \frac{1}{4}(7+\sqrt{41}) \simeq 3.35$ sont les valeurs propres généralisées. D'où deux vecteurs propres généralisées donnés par $\vec{e}_j = \begin{pmatrix} x_{1j} \\ x_{2j} \end{pmatrix}$ avec $(1-\lambda_j)x_{1j} + 2x_{2j} = 0$. Par exemple $\vec{e}_j = \begin{pmatrix} -2 \\ 1-\lambda_j \end{pmatrix}$.

V.p. classiques : on cherche $(\mu, \vec{v}) \in \mathbb{R}\times\mathbb{R}^2$ tq $(R-\mu I)\vec{v} = 0$ et $\vec{v} \neq 0$. La matrice $R-\mu I$ est alors non inversible, i.e. son déterminant est nul : $\det \begin{pmatrix} 1-\mu & 2 \\ 2 & 5-\mu \end{pmatrix} = 0$, soit $\mu^2 - 6\mu + 1 = 0$. D'où $\mu_1 = 3-2\sqrt{2} \simeq 0.17$ et $\mu_2 = 3+2\sqrt{2} \simeq 5.83$ sont les valeurs propres classiques. D'où deux vecteurs propres classiques donnés par $\vec{v}_j = \begin{pmatrix} y_{1j} \\ y_{2j} \end{pmatrix}$ avec $(1-\mu_j)y_{1j} + 2y_{2j} = 0$. Par exemple $\vec{v}_j = \begin{pmatrix} -2 \\ 1-\mu_j \end{pmatrix}$. ▪

Exercice 6.7 1- Montrer que si R et M sont symétriques et que $R.M$ est symétrique, alors $R.M = M.R$ (i.e. R et M commutent).

2- Montrer que si R et M sont symétriques et commutent alors $R.M$ est symétrique.

3- Montrer que si M commute avec R, alors M commute avec R^2.

4- Montrer que si R et M sont des matrices symétriques, que $R.M$ est symétrique et que M est définie positive, alors $M^{-1}.R$ est symétrique.

Réponse. 1- $R.M$ symétrique s'écrit $(R.M)^T = R.M$, soit $M^T.R^T = R.M$, d'où $M.R = R.M$ puisque $M = M^T$ et $R = R^T$.

2- $(R.M)^T = M^T.R^T = M.R = R.M$.

3- $M.R^2 = (M.R).R = (R.M).R = R.(M.R) = R.(R.M) = R^2.M$.

4- $M.R = R.M$ donne $M.R.M^{-1} = R$ donne $R.M^{-1} = M^{-1}.R$. ▪

6.2 Théorème de diagonalisation généralisée de matrices symétriques

On disposera donc du produit scalaire $(\cdot,\cdot)_M$ donné par :

$$(\vec{x}, \vec{y})_M \stackrel{\text{def}}{=} (M.\vec{x}, \vec{y})_{\mathbb{R}^n}. \tag{6.2}$$

Théorème 6.8 *Si R et M sont des matrices symétriques réelles, et si M est définie positive, alors le problème (6.1) admet n valeurs propres généralisées $\lambda_1, \dots, \lambda_n$ associées à n vecteurs propres généralisés $\vec{e}_1, \dots, \vec{e}_n$ qu'on peut choisir formant une base orthonormale de $(\mathbb{R}^n, (\cdot,\cdot)_M)$. I.e., il existe $(\vec{e}_1, \dots, \vec{e}_n)$ une base de $\mathbb{R}^n$ telle que :*

$$\forall i, j = 1, \dots, n, \qquad R.\vec{e}_j = \lambda_j M.\vec{e}_j, \qquad (\vec{e}_j, \vec{e}_i)_M = \delta_{ij}. \tag{6.3}$$

Notant P la matrice dont les colonnes j sont les composantes dans la base canonique des vecteurs propres généralisés $\vec{e}_j$, et notant $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ la matrice diagonale contenant les valeurs propres généralisées, on peut réécrire (6.3) sous la forme :

$$R.P = M.P.D, \qquad P^T.M.P = I. \tag{6.4}$$

En particulier on a $(M.P)^{-1} = P^T$ et donc (cas R et M symétriques, M définie positive) :

$$D = P^T.R.P, \qquad P^T.M.P = I, \tag{6.5}$$

à comparer avec (5.3).

Preuve. Noter qu'il ne sert à rien de réécrire ce problème sous la forme $(M^{-1}R).\vec{x} = \lambda\vec{x}$, car bien que M^{-1} et R soient symétriques, il n'y a aucune raison pour que le produit $M^{-1}.R$ soit une matrice symétrique (exemple $M^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ pour lesquelles $M^{-1}.R = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$.

M étant une matrice symétrique, on peut utiliser sa décomposition de Cholesky, et on introduit le vecteur $\vec{y}$:

$$M = L.L^T, \quad \vec{y} = L^T.\vec{x}, \quad M.\vec{x} = L.\vec{y}, \quad R.\vec{x} = R.L^{-T}.\vec{y}, \tag{6.6}$$

avec L matrice triangulaire inférieure (inversible puisque M l'est). Le problème spectral (6.1) se réécrit en $\vec{y}$ dans $\mathbb{R}^n$: trouver les $(\lambda, \vec{y}) \in \mathbb{R} \times \mathbb{R}^n$ tels que :

$$S.\vec{y} = \lambda\vec{y} \qquad \text{où} \qquad S = L^{-1}.R.L^{-T}. \tag{6.7}$$

Et S étant trivialement symétrique réelle, S est diagonalisable (classique) dans une b.o.n. de $\mathbb{R}^n$: il existe n couples vp-vp $(\lambda_j, \vec{y}_j) \in \mathbb{R} \times \mathbb{R}^n$ de S t.q. $(\vec{y}_j)_{i=1,\dots,n}$ est une b.o.n. de $\mathbb{R}^n$.

D'où posant $\vec{e}_j = L^{-T}\vec{y}_j$, on a obtenu n vecteurs propres généralisés tels que :

$$R.\vec{e}_j = \lambda_j M.\vec{e}_j, \qquad \vec{e}_i^{\,T}.M.\vec{e}_j = \delta_{ij}, \tag{6.8}$$

$\vec{e}_i^{\,T}.M.\vec{e}_j = \vec{y}_i^{\,T}.L^{-1}.M.L^{-T}.\vec{y}_j = \vec{y}_j^{\,T}.\vec{y}_j = \delta_{ij}$. la dernière égalité car D'où la décomposition spectrale cherchée du problème spectral généralisé (6.3).

(Méthode directe pour voir que les valeurs propres classiques de S sont les valeurs propres généralisées de R : on a $\det(S - \lambda I) = \det(L^{-1}.R.L^{-T} - \lambda I) = \det(L^{-1}.(R - \lambda M).L^{-T}) = \det(L^{-1})\det(R - \lambda M)\det(L^{-T})$.) ∎

Corollaire 6.9 *1- Si R est symétrique positive, alors ses vp généralisées λ_j sont toutes ≥ 0.*

1'- Si R est symétrique négative, alors ses vp généralisées λ_j sont toutes ≤ 0.

2- Si R est symétrique définie positive, alors ses vp généralisées λ_j sont toutes > 0.

2'- Si R est symétrique définie négative, alors ses vp généralisées λ_j sont toutes < 0.

3- Si R admet 0 comme valeur propre classique de multiplicité m, alors R admet 0 comme valeur propre généralisée de même multiplicité m. (C'est faux pour les autres valeurs propres, voir exemple 6.5).

Preuve. 1- Si R est symétrique positive, alors S donnée en (6.7) est également positive car pour $\vec{y}$ on a $(S.\vec{y}, \vec{y})_{\mathbb{R}^n} = (L^{-1}.R.L^{-T}.\vec{y}, \vec{y})_{\mathbb{R}^n} = (R.\vec{x}, \vec{x}) \geq 0$ où $\vec{x} = L^{-T}.\vec{y}$. Donc les λ_j sont ≥ 0.

1'- Si $-R$ est symétrique positive, le problème s'écrivant $(-R).\vec{x} = (-\lambda)M.\vec{x}$, le 1- donne : les $-\lambda_j$ sont ≥ 0.

2- Si R est symétrique définie positive, alors S est également définie positive car pour $\vec{y} \neq \vec{0}$ on a $(S.\vec{y}, \vec{y})_{\mathbb{R}^n} = (L^{-1}.R.L^{-T}.\vec{y}, \vec{y})_{\mathbb{R}^n} = (R.\vec{x}, \vec{x}) > 0$ où $\vec{x} = L^{-T}.\vec{y}$ car $L^{-T}.\vec{y} \neq \vec{0}$ (la matrice L^{-T} est inversible). Donc les λ_j sont > 0.

2'- : adapter 1'-.

3- Si R admet 0 comme valeur propre classique de multiplicité m, alors $\dim(\mathrm{Ker}(R)) = m$ (car R est diagonalisable classiquement), soit $m = \dim(\mathrm{Ker}(R - 0I)) = \dim(\mathrm{Ker}(R - 0M))$, donc 0 est également valeur propre généralisée de multiplicité m (car R est diagonalisable généralisée). ∎

Corollaire 6.10 *Soit R une matrice symétrique et N une matrice symétrique définie positive, alors la matrice $N.R$ (non symétrique en général) est diagonalisable.*

Preuve. On pose $M = N^{-1}$ et $N.R.\vec{x} = \lambda\vec{x}$ s'écrit $R.\vec{x} = \lambda M.\vec{x}$: on applique le théorème 6.8. ∎

6.3 Diagonalisation d'un endomorphisme à partir d'une base quelconque

On suppose qu'on connaît une base $(\vec{a}_i)$ dans V mais que cette base n'est pas orthonormée dans $(V, (\cdot,\cdot)_V)$. Exemple : la base des fonctions chapeau dans P_1 (si $(\vec{a}_i)$ est une b.o.n, voir le § 5.3).

Soit $\mathcal{L}$ un endomorphisme symétrique : $(\mathcal{L}.\vec{x}, \vec{y})_V = (\mathcal{L}.\vec{y}, \vec{x})_V$ pour tout $\vec{x}, \vec{y}$. En particulier $(\mathcal{L}.\vec{a}_i, \vec{a}_j)_V = (\mathcal{L}.\vec{a}_j, \vec{a}_i)_V$ pour tout i, j.

On souhaite trouver les vecteurs propres $\vec{v}_j$ de $\mathcal{L}$ dans la base $(\vec{a}_i)$, i.e. posant $\vec{v}_j = \sum_{i=1}^n P_{ij}\vec{a}_i$, les composantes réelles P_{ij} sont à déterminer.

Trouver les $(\lambda,\vec{v})\in\mathbb{R}\times V$ t.q. $\mathcal{L}.\vec{v}=\lambda\vec{v}$ équivaut à : trouver les $(\lambda,\vec{v})\in\mathbb{R}\times V$ t.q. :

$$\forall i=1,...,n,\qquad (\mathcal{L}.\vec{v},\vec{a}_i)_V=\lambda(\vec{v},\vec{a}_i)_V, \tag{6.9}$$

(n équations) soit, notant $\vec{v}=\sum_{j=1}^n c_j\vec{a}_j$, trouver les $(\lambda,c_1,...,c_n)\in\mathbb{R}\times\mathbb{R}^n$ t.q. :

$$\forall i=1,...,n,\qquad \sum_{j=1}^n c_j(\mathcal{L}.\vec{a}_j,\vec{a}_i)_V=\lambda\sum_{j=1}^n c_j(\vec{a}_j,\vec{a}_i)_V, \tag{6.10}$$

(n+1 inconnues les c_j et λ) soit :

$$\forall i=1,...,n,\qquad \sum_{j=1}^n R_{ij}c_j=\lambda\sum_{j=1}^n M_{ij}c_j,\qquad \text{où}\qquad \begin{cases} R=[R_{ij}]=[(\mathcal{L}.\vec{a}_j,\vec{a}_i)_V,\\ M=[M_{ij}]=[(\vec{a}_j,\vec{a}_i)_V].\end{cases} \tag{6.11}$$

Donc :

$$R.[\vec{c}]=\lambda M.[\vec{c}],\qquad \text{où}\qquad [\vec{c}]=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix} \tag{6.12}$$

est le problème à résoudre.

1- (6.12) est un problème aux valeurs propres généralisé dans $\mathbb{R}^n$ où R est symétrique car $\mathcal{L}$ l'est et M est symétrique définie positive (matrice de masse).

Notons $\vec{e}_j\in\mathbb{R}^n$ les vecteurs propres généralisés associés aux valeurs propres généraliséesé λ_i trouvés pour (6.12) : avec $(\vec{E}_i)$ la base canonique de $\mathbb{R}^n$, pour tout $j=1,...,n$:

$$\vec{e}_j=\sum_{i=1}^n P_{ij}\vec{E}_i,\qquad [\vec{e}_j]=\begin{pmatrix}P_{1j}\\ \vdots\\ P_{nj}\end{pmatrix},\qquad R.[\vec{e}_j]=\lambda_j M.[\vec{e}_j],\qquad P^T.M.P=I, \tag{6.13}$$

où donc :

$$D=\mathrm{diag}(\lambda_1,...,\lambda_n),\quad P=(\,[\vec{e}_1]\quad ...\quad [\vec{e}_n]\,). \tag{6.14}$$

2- Retour dans V (on remonte les calculs) : notons $\vec{v}_j\in V$ les vecteurs définis par :

$$\vec{v}_j=\sum_{i=1}^n P_{ij}\vec{a}_i,\qquad \text{donc}\qquad [\vec{v}_j]_{|(\vec{a})}=\begin{pmatrix}P_{1j}\\ \vdots\\ P_{nj}\end{pmatrix}\qquad \text{et}\qquad R.[\vec{v}_j]_{|(\vec{a})}=\lambda_j M.[\vec{v}_j]_{|(\vec{a})}. \tag{6.15}$$

Donc, pour tout i,j :

$$\begin{aligned}(\mathcal{L}.\vec{v}_j,\vec{a}_i)_V &= \sum_{k=1}^n P_{kj}(\mathcal{L}.\vec{a}_k,\vec{a}_i)_V=\sum_{k=1}^n P_{kj}R_{ik}=(R.P)_{ij}=(R.[\vec{e}_j])_i=(\lambda_j M.[\vec{e}_j])_i\\ &=\lambda_j\sum_{k=1}^n M_{ik}P_{kj}=\lambda_j\sum_{k=1}^n P_{kj}(\vec{a}_k,\vec{a}_i)_V=\lambda_j(\sum_{k=1}^n P_{kj}\vec{a}_k,\vec{a}_i)_V=\lambda_j(\vec{v}_j,\vec{a}_i)_V,\end{aligned} \tag{6.16}$$

donc $\mathcal{L}.\vec{v}_j=\lambda_\ell\vec{v}_j$ pour tout j (car $(\vec{a}_i)$ est une base) : les $\vec{v}_j$ sont des vecteurs propres de $\mathcal{L}$ associés aux valeurs propres λ_j. Et ils vérifient, pour tout i,j :

$$(\vec{v}_i,\vec{v}_j)_V=\sum_{k\ell}P_{ki}P_{\ell j}(\vec{a}_k,\vec{a}_\ell)_V=\sum_{k\ell}P_{ki}P_{\ell j}M_{k\ell}=(P^T.M.P)_{ij}=\delta_{ij}, \tag{6.17}$$

donc $(\vec{v}_i)$ est une b.o.n., dans $(V,(\cdot,\cdot)_V)$, de diagonalisation de $\mathcal{L}$.

6.4 Base commune de diagonalisation

Les solutions de (6.1) vérifient :

$$(R.\vec{x},\vec{y})_{\mathbb{R}^n}=\lambda(\vec{x},\vec{y})_M,\qquad \forall\vec{y}\in\mathbb{R}^n. \tag{6.18}$$

(Le problème (4.1) s'écrivait $(R.\vec{x},\vec{y})_{\mathbb{R}^n}=\lambda(\vec{x},\vec{y})_{\mathbb{R}^n}$, pour tout $\vec{y}\in\mathbb{R}^n$).

Et en projetant (6.3) sur les vecteurs $\vec{e}_i$ à l'aide du produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$, on obtient :

$$(R.\vec{e}_j, \vec{e}_i)_{\mathbb{R}^n} = \lambda_j (M.\vec{e}_j, \vec{e}_i)_{\mathbb{R}^n} = \lambda_j \delta_{ij}. \tag{6.19}$$

En particulier, si R est également (symétrique) définie positive, alors $(\vec{e}_i)$ est un base orthogonale à la fois pour les produits scalaires $(\cdot,\cdot)_M$ et $(\cdot,\cdot)_R$. En effet, (6.19) se réécrit :

$$(\frac{\vec{e}_i}{\sqrt{\lambda_i}}, \frac{\vec{e}_j}{\sqrt{\lambda_j}})_R = (\vec{e}_j, \vec{e}_i)_M = \delta_{ij}.$$

I.e., $(\vec{e}_i)$ est une base orthogonale commune de diagonalisation pour ces deux produits scalaires.

Exercice 6.11 Soit $R = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ et $M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Vérifier que les vecteurs propres généralisés sont M-orthogonaux et R-orthogonaux, et qu'ici ils ne sont pas orthogonaux pour le produit scalaire usuelle de $\mathbb{R}^n$.

Réponse. Voir exercice 6.6. On vérifie que $(\vec{e}_1, \vec{e}_2)_{\mathbb{R}^n} = 4 + (1-\lambda_1)(1-\lambda_2) = 2 \neq 0$ alors que $(R\vec{e}_1, \vec{e}_2)_{\mathbb{R}^n} = \lambda_1 (M\vec{e}_1, \vec{e}_2)_{\mathbb{R}^n} = \lambda_1(4 + 2(1-\lambda_1)(1-\lambda_2)) = 0$, i.e. $\vec{e}_1$ et $\vec{e}_2$ sont R et M orthogonaux. ∎

Exercice 6.12 Vérifier dans l'exercice précédent que la matrice $R.M$ n'est pas symétrique (relativement au produit scalaire $(\cdot,\cdot)_{\mathbb{R}^n}$).

Réponse. $R.M = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}$ n'est pas symétrique : $(R.M)^T \neq R.M$. ∎

6.5 Application : résolution du système différentiel $M.\vec{\alpha}\,' + R.\vec{\alpha} = \vec{f}$

But : trouver $\vec{\alpha} : [0,T] \to \mathbb{R}^n$ t.q. :

$$M.\frac{d\vec{\alpha}}{dt}(t) + R.\vec{\alpha}(t) = \vec{f}(t), \qquad \vec{\alpha}(0) = \vec{\alpha}_0, \tag{6.20}$$

où M est une matrice symétrique définie positive, R est une matrice symétrique, $\vec{f} : \mathbb{R} \to \mathbb{R}^n$ est une fonction continue, et $\vec{\alpha}_0 \in \mathbb{R}^n$ est un vecteur donné (condition initiale).

Le cas $M = I$ est le cas particulier d'un système différentiel sous forme normale, cf. § 5.6.

On reprend la démarche du § 5.6 : notons α_i les composantes de $\vec{\alpha}$ sur la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$, et β_i les composantes de $\vec{\alpha}$ sur la b.o.n $(\vec{v}_i)$ de diagonalisation généralisée :

$$\vec{\alpha} = \sum_i \alpha_i \vec{E}_i = \sum_i \beta_i \vec{v}_i, \qquad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \overset{\text{noté}}{=} [\vec{\alpha}], \qquad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \overset{\text{noté}}{=} [\vec{\beta}], \qquad [\vec{\beta}] = P^{-1}.[\vec{\alpha}]. \tag{6.21}$$

Le problème (6.20) est une écriture abusive du problème matriciel : trouver les α_i t.q. :

$$M.\frac{d[\vec{\alpha}]}{dt}(t) + R.[\vec{\alpha}](t) = [\vec{f}](t), \qquad [\vec{\alpha}(0)] = [\vec{\alpha}_0]. \tag{6.22}$$

Donc, puisque $[\vec{\alpha}] = P.[\vec{\beta}]$:

$$M.P.\frac{d[\vec{\beta}]}{dt}(t) + R.P.[\vec{\beta}](t) = [\vec{f}](t), \tag{6.23}$$

D'où, multipliant par P^T, avec $P^T.M.P = I$ et $P^T.R.P = D$:

$$\frac{d[\vec{\beta}]}{dt}(t) + D.[\vec{\beta}](t) = P^T.[\vec{f}](t), \tag{6.24}$$

On a obtenu un système différentiel totalement découplé : pour tout $i = 1, ..., n$:

$$\beta_i\,'(t) + \lambda_i \beta_i(t) = (P^T.[\vec{f}(t)])_i, \tag{6.25}$$

Résolution facile usuelle, solutions données à une constantes près. D'où les α_i donnés par :

$$[\vec{\alpha}] = P.[\vec{\beta}], \qquad [\vec{\alpha}(0)] = [\vec{\alpha}_0] = P.[\vec{\beta}(0)], \tag{6.26}$$

ce qui fixe les "constantes" des solutions β_i données "à une constantes près".

6.6 Application : recherche de b.o.n. dans $(P_1, (\cdot,\cdot)_{H^1})$

Un problème aux valeurs propres formulé dans une base non orthonormée se ramène à un problème aux valeurs propres généralisé.

6.6.1 Problème initial

Exemple : Exemple de problème aux valeurs propres dans P_1 : "trouver les fréquences propres de vibration d'un ressort attaché à ses extrémités". Le problème initial est, pour $a < b$:

$$\begin{cases} \text{trouver } (\lambda, u) \in \mathbb{R} \times H_0^1(]a,b[) \text{ tel que :} \\ -u'' = \lambda u, \end{cases} \tag{6.27}$$

où :

$$H_0^1(]a,b[) = \{u \in L^2([a,b]) : u' \in L^2([a,b]),\ u(a) = u(b) = 0\} \tag{6.28}$$

est l'ensemble des fonctions dérivables dans $L^2([a,b])$ et nulles au bord (cf. cours d'éléments finis pour les détails). On muni $H_0^1(]a,b[)$ du produit scalaire et de la norme associée :

$$(u,v)_{H_0^1} = (u',v')_{L^2}, \qquad ||u||_{H^1} = ||u'||_{L^2}, \tag{6.29}$$

qui fait de $H_0^1(]a,b[)$ un Hilbert (voir cours d'éléments finis).

6.6.2 Problème approché dans P_1

On reprend l'espace P_1 donné en (5.28). Et on note :

$$P_{10} = \{f_h \in P_1 : f_h(a) = f_h(b) = 0\}, \tag{6.30}$$

où P_{10} est le sous-espace de P_1 des fonctions nulles au bord, avec ici $\dim(P_{10}) = N-1 = n$. Et une base de P_{10} est $(\varphi_i)_{i=1,\dots,N-1}$ (des fonctions chapeaux nulles au bord).

On ne peut pas chercher directement une solution $u \in P_1$ de (6.27) : dériver deux fois une fonction P_1 donne des masses de Dirac (on passe au sens des distributions).

On prend la formulation "intégrée par parties" : on multiplie (6.27) par v puis on intègre par parties :

$$\begin{cases} \text{trouver } (\lambda, u) \in \mathbb{R} \times H_0^1(]a,b[) \text{ tel que :} \\ \displaystyle\int_a^b u'(x)v'(x)\,dx = \lambda \int_a^b u(x)v(x)\,dx, \qquad \forall v \in H_0^1(]a,b[), \end{cases} \tag{6.31}$$

les termes de bord ayant disparus car v est supposé nul au bord. Cette formulation (6.31) est adaptée à notre recherche de solutions approchées : on pose le problème (notre problème approché) :

$$\begin{cases} \text{trouver } (\lambda_h, u_h) \in \mathbb{R} \times P_{10} \text{ tel que :} \\ \displaystyle\int_a^b u_h'(x)v_h'(x)\,dx = \lambda_h \int_a^b u_h(x)v_h(x)\,dx, \qquad \forall v_h \in P_{10}. \end{cases} \tag{6.32}$$

Pour la suite (simplification des notations et mise en évidence des "produits scalaires"), on pose :

$$a(u,v) = \int_a^b u'(x)v'(x)\,dx \qquad (= (u',v')_{L^2}). \tag{6.33}$$

Donc ici $a(\cdot,\cdot) = (\cdot,\cdot)_{H_0^1}$ est le produit scalaire de $H_0^1(]a,b[)$. Et (6.32) se réécrit :

$$\begin{cases} \text{trouver } (\lambda_h, u_h) \in \mathbb{R} \times P_{10} \text{ tel que :} \\ a(u_h, v_h) = \lambda(u_h, v_h)_{L^2}, \qquad \forall v_h \in P_{10}, \end{cases} \tag{6.34}$$

soit :

$$\begin{cases} \text{trouver } (\lambda_h, u_h) \in \mathbb{R} \times P_{10} \text{ tel que :} \\ a(u_h, \varphi_i) = \lambda(u_h, \varphi_i)_{L^2}, \qquad \forall i = 1, \dots, n. \end{cases} \tag{6.35}$$

Une solution u_h sera connue si on trouve les $n = N-1$ composantes de u_h sur la base $(\varphi_i)_{i=1,\dots,n}$. Notons c_j ces composantes :

$$u_h = \sum_{j=0}^{N} c_j\varphi_j = \sum_{j=1}^{N-1} c_j\varphi_j = \sum_{j=1}^{n} c_j\varphi_j, \tag{6.36}$$

puisque $u_h(a) = u_h(b) = 0$ donne $c_0 = c_N = 0$.

6.6.3 Problème plongé dans $\mathbb{R}^n$

Donc, avec (6.36), (6.35) s'écrit, avec $n = N-1$:

$$\begin{cases} \text{trouver } (\lambda_h, (c_j)_{j=1,\dots,n}) \in \mathbb{R} \times \mathbb{R}^n \text{ tq :} \\ \displaystyle\sum_{j=1}^n c_j\, a(\varphi_j, \varphi_i) = \lambda_h \sum_{j=1}^n c_j\, (\varphi_j, \varphi_i)_{L^2}, \quad \forall i = 1,..,n, \end{cases} \tag{6.37}$$

qui est le système de n équations (les lignes i) et n inconnues (les c_j) :

$$R.[\vec{c}] = \lambda_h M.[\vec{c}], \qquad [\vec{c}] = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \qquad R = [a(\varphi_j, \varphi_i)], \qquad M = [(\varphi_j, \varphi_i)_{L^2}], \tag{6.38}$$

où $\vec{c} = \sum_{i=1}^n c_i \vec{E}_i$ est le vecteur inconnu de composantes les c_i.

Ici M n'est pas la matrice identité (ni une matrice diagonale) : la base (φ_i) n'est pas orthogonale : le problème (6.38) est un problème aux valeurs propres généralisées.

Ayant R et M symétriques avec M définie positive (matrice de masse), le problème (6.38) admet des solutions $(\lambda_i, \vec{e}_i) \in \mathbb{R} \times \mathbb{R}^n$ telles que :

$$R.\vec{e}_j = \lambda_j M.\vec{e}_j \qquad \text{et} \qquad (M.\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = \delta_{ij}, \tag{6.39}$$

où donc $(\vec{e}_i)_{i=1,\dots,n}$ est une b.o.n. de $(\mathbb{R}^n, (\cdot,\cdot)_M)$. Notons comme précédemment P la matrice de passage de la base canonique à la base $(\vec{e}_i)$, i.e. pour tout j :

$$\vec{e}_j = \sum_i P_{ij} \vec{E}_i, \qquad [\vec{e}_j] = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \qquad P^T.M.P = I. \tag{6.40}$$

6.6.4 Retour dans P_1

Définissons les fonctions affines par morceaux $\chi_j \in P_{10}$ par, avec les P_{ij} donnés en (6.40) :

$$\chi_j = \sum_i P_{ij} \varphi_i, \qquad [\chi_j]_{|(\varphi_i)} = [\vec{e}_j] = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}. \tag{6.41}$$

La matrice P étant inversible, la famille $(\chi_i)_{i=1,\dots,n}$ est bien une base de P_{10} (famille de $n = \dim P_{10}$ éléments et famille libre car P est inversible).

Comme $(\vec{e}_j)$ est une b.o.n. de $(\mathbb{R}^n, (\cdot,\cdot)_M)$, on en déduit que (χ_i) est une b.o.n. dans $(P_{10}, (\cdot,\cdot)_{L^2})$ car :

$$(\chi_i, \chi_j)_{L^2} = (\sum_k P_{ki}\varphi_k, \sum_\ell P_{\ell j}\varphi_\ell)_{L^2} = \sum_{k\ell} P_{ki} P_{\ell j} M_{k\ell} = (P^T.M.P)_{ij} = \delta_{ij}. \tag{6.42}$$

Et donc que :

$$a(\chi_i, \chi_j) = (\sum_k P_{ki} a\varphi_k, \sum_\ell P_{\ell j}\varphi_\ell)_{L^2} = \sum_{k\ell} P_{ki} P_{\ell j} R_{k\ell} = (P^T.R.P)_{ij} = \lambda_i \delta_{ij}. \tag{6.43}$$

Donc (χ_i) est une base orthogonale dans $(H_0^1(]a,b[), a(\cdot,\cdot))$: c'est une base commune de diagonalisation dans $H_0^1(]a,b[)$ pour les produits scalaires $(\cdot,\cdot)_{L^2}$ et $a(\cdot,\cdot) = (\cdot,\cdot)_{H_0^1}$.

Et $(\frac{\chi_i}{\lambda_i})$ est une b.o.n. dans $(H_0^1(]a,b[), a(\cdot,\cdot))$.

6.7 Loi d'inertie de Sylvester, suite

6.7.1 Signature de matrices congruentes

Les matrices congruentes ont été définies à la définition 3.41 page 36.

La signature d'une matrice $A = [a(\cdot,\cdot)] \in \mathbb{R}^{n^2}$ a été définie à la définition 5.20 page 51.

Corollaire 6.13 *Soit A et B deux matrices symétriques réelles congruentes (« $B = P^T.A.P$ »). Alors A et B ont même signature (même nombre de valeurs propres > 0 et < 0).*

Preuve. Soit A matrice symétrique réelle. On note λ_i ses valeurs propres et $(\vec{v}_i)$ la b.o.n. de diagonalisation. Soit $a(\cdot,\cdot)$ la forme bilinéaire de matrice A dans la base canonique $(\vec{E}_i)$: donc $a(\cdot,\cdot)$ est définie par $[a(\cdot,\cdot)]_{|(\vec{E})} = A$, soit $a(\vec{E}_i, \vec{E}_j) = A_{ij}$ pour tout i,j. Et $a(\cdot,\cdot)$ est symétrique car A l'est. Donc notant Q la forme quadratique associée à $a(\cdot,\cdot)$, on a (5.54) (expression de Q dans la base $(\vec{v}_i)$ de diagonalisation) :

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{v}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i,j=1}^{n} x_i x_j a(\vec{v}_i, \vec{v}_j) = \sum_{i=1}^{n} \lambda_i x_i^2, \tag{6.44}$$

puisque $a(\vec{v}_i, \vec{v}_j) = (A.\vec{v}_i, \vec{v}_j)_{\mathbb{R}^n} = (\lambda_i \vec{v}_i, \vec{v}_j)_{\mathbb{R}^n} = \lambda_i \delta_{ij}$.

Soit B congruente à A. Donc $B = Z^T.A.Z$ où $Z = (\,[\vec{z}_1] \quad ... \quad [\vec{z}_n]\,)$ est une matrice inversible (donc $\vec{z}_j = \sum_{i=1}^n Z_{ij} \vec{E}_i$ pour tout j). Donc $B = [a(\cdot,\cdot)]_{|(\vec{z})}$ est la matrice de $a(\cdot,\cdot)$ dans la base $(\vec{z}_i)$ (formule de changement de base pour les formes bilinéaires), i.e. $B_{ij} = a(\vec{z}_i, \vec{z}_j) = [\vec{z}_i]^T.A.[\vec{z}_j]$ pour tout i,j.

Donc B est symétrique réelle ; donc diagonalisable dans une b.o.n. $(\vec{p}_i)$ de $\mathbb{R}^n$: $B = P.D.P^{-1}$ où $D = \mathrm{diag}(\mu_i)$, $P = (\,[\vec{p}_1] \quad ... \quad [\vec{p}_n]\,)$, avec $P^T.P = I$, $B.[\vec{p}_j] = \mu_j [\vec{p}_j]$ et $[\vec{p}_i]^T.B.[\vec{p}_j] = \mu_j \delta_{ij}$ pour tout i,j. Notons $\vec{w}_j$ les vecteurs t.q. $[\vec{w}_j]_{|(\vec{z})} = [\vec{p}_j]$ (donc $\vec{w}_j = \sum_{i=1}^n P_{ij} \vec{z}_i$). Alors pour tout i,j :

$$a(\vec{w}_i, \vec{w}_j) = \sum_{k,\ell=1}^{n} P_{ki} P_{\ell j} a(\vec{z}_k, \vec{z}_\ell) = \sum_{k,\ell=1}^{n} (P^T)_{ik} B_{k\ell} P_{\ell j} = (P^T.B.P)_{ij} = D_{ij}\delta_{ij} = \mu_i \delta_{ij}.$$

Donc

$$\vec{x} = \sum_{i=1}^{n} y_i \vec{w}_i \quad \Longrightarrow \quad Q(\vec{x}) = \sum_{i,j=1}^{n} y_i y_j a(\vec{w}_i, \vec{w}_j) = \sum_{i=1}^{n} \mu_i y_i^2. \tag{6.45}$$

On compare avec (6.44) : la loi d'inertie de Sylvester donne le résultat : la signature est une caractéristique de Q (indépendante de la représentation dans une base) : même nombre de valeurs propres > 0 et < 0. ∎

6.7.2 Valeurs propres vs valeurs propres généralisées : "signature" conservée

Soit A une matrice symétrique réelle de signature (s_+, s_-), donc avec s_+ valeurs propres classiques > 0 et s_- valeurs propres classiques < 0.

Corollaire 6.14 *Soit M une matrice symétrique définie positive. Alors A est s_+ valeurs propres M-généralisées > 0 et s_- valeurs propres M-généralisées < 0.*

Preuve. Soit $M = L.L^T$ la décomposition de Cholesky de M. Polynôme caractéristique généralisé : $p(\mu) = \det(A - \mu M) = \det(A - \mu L.L^T) = \det(L)\det(L^{-1}.A.L^{-T} - \mu I)\det(L^T)$, donc $p(\mu) = \det(L)^2 \det(B - \mu I)$ où on a posé $B = L^{-1}.A.L^{-T}$. La matrice $B = L^{-1}.A.L^{-T}$ est congruente à A (poser $P = L^{-T}$ qui donne $B = P^T.A.P$). Donc B a même signature que A, cf. corollaire 6.13. Et les valeurs propres classiques de B sont les valeurs propres généralisées de A : même nombre de valeurs propres classiques > 0 (resp. < 0) de A que de valeurs propres classiques > 0 (resp. < 0) de B qui sont les valeurs propres généralisées de A. ∎

7 Une matrice complexe hermitienne est diagonalisable

Les matrices hermitiennes ont été définies en (1.51) : ce sont les matrices complexes $n * n$ vérifiant $A^* = A$ où $A^* = \overline{A^T}$.

Théorème 7.1 *Soit $A \in \mathbb{C}^{n^2}$ une matrice hermitienne. Alors ses valeurs propres sont toutes réelles, et A est diagonalisable dans une base $(\cdot,\cdot)_{\mathbb{C}^n}$-orthonormée : i.e. il existe n couples $(\lambda_j, \vec{v}_j) \in \mathbb{R} \times \mathbb{C}^n$, $1 \le j \le n$ (les valeurs propres sont toutes réelles), tels que :*

$$A.\vec{v}_j = \lambda_j \vec{v}_j, \quad \forall j = 1, ..., n \qquad \text{et} \qquad (\vec{v}_i, \vec{v}_j)_{\mathbb{C}^n} = \delta_{ij}, \quad \forall i, j = 1, ..., n. \tag{7.1}$$

Notant P la matrice dont la colonne j donne les composantes du vecteur $\vec{v}_j$ dans la base canonique, et $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ la matrice diagonale réelle contenant les valeurs propres, (7.1) se réécrit sous la forme :

$$A.P = P.D, \qquad P^*.P = I, \tag{7.2}$$

soit :

$$D = P^*.A.P, \qquad P^{-1} = P^*, \tag{7.3}$$

soit :

$$A = P.D.P^*, \qquad P^{-1} = P^*. \tag{7.4}$$

De plus les espaces propres sont orthogonaux 2 à 2 (relation $P^*.P = I$).

Preuve. Similaire au cas réel puisqu'on dispose de la proposition 5.1 (les vp sont réelles), avec ici $(P^*.P)_{ij} = (\overline{P}^T.P)_{ij} = [\overline{\vec{v}}_i]^T.[\vec{v}_j] = (\vec{v}_j, \vec{v}_i)_{\mathbb{C}^n} = \delta_{ji} = \delta_{ij}$ pour tout i, j, soit $P^*.P = I$.

Puis $A = P.D.P^*$ donne $A^T = (P^*)^T.D.P^T = \overline{P}.D.P^T$, d'où $A^* = P.D.P^*$ car D est réelle. ▪

Remarque 7.2 On retrouve le cas A symétrique réelle (cas particulier des matrices hermitiennes) : dans ce cas $P^* = P^T$ car P est réelle. ▪

Exercice 7.3 Diagonaliser $A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$.

Réponse. $A^T = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$ donne $A^* = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = A$: ici A est hermitienne.

On a $\det(A - \lambda I) = (1-\lambda)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda - 2)$, d'où $D = \mathrm{diag}(0, 2)$, et $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$. On vérifie que $P^*.P = I$, et en particulier les vecteurs colonnes de P vérifient $(\vec{v}_1, \vec{v}_2)_{\mathbb{C}^2} = [\vec{v}_1]^T.[\overline{\vec{v}_2}] = (1 \quad i).\begin{pmatrix} -i \\ 1 \end{pmatrix} = -i + i = 0$. ▪

Remarque 7.4 La proposition 5.8 est fausse si on se place dans $\mathbb{C}$: si $A \in \mathbb{C}^{n^2}$ est diagonalisable et si on peut choisir des vecteurs propres qui forment une b.o.n. de $(\mathbb{C}^n, (\cdot,\cdot)_{\mathbb{C}^n})$, alors on n'a pas nécessairement A hermitienne.

Voir exemple suivant 7.5.

Si on essaie la démonstration de la proposition 5.8 : $(\vec{v}_i)$ b.o.n. dans $(\mathbb{C}^n, (\cdot,\cdot)_{\mathbb{C}^n})$, P comme précédemment, avec ici $P.P^* = I$ et $A = P.D.P^{-1} = P.D.P^*$, donc $A^* = P.D^*.P^* = P.\overline{D}.P^* \neq A$ en général quand $D \neq \overline{D}$ (voir remarque 1.53). ▪

Exercice 7.5 Diagonaliser $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Réponse. A a deux valeurs propres distinctes $\lambda_\pm = \pm i$: $D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Et des vecteurs propres sont donnés à l'exercice 4.20. Prenons $P = \begin{pmatrix} \frac{1}{\sqrt{2}}[\vec{v}_1] & \frac{1}{\sqrt{2}}[\vec{v}_2] \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, qui vérifie $P^*.P = I$. Ici A n'est pas hermitienne : $A^* = A^T = -A \neq A$ (ici D n'est pas réelle). ▪

8 Diagonalisation des matrices antisymétriques et des matrices normales

8.1 Matrices antisymétriques

Définition 8.1 Une matrice carrée réelle $A = [A_{ij}] \in \mathbb{R}^{n^2}$ est anti-symétrique (skew-symmetric) ssi $A^T = -A$, i.e. ssi $A_{ji} = -A_{ij}$ pour tout i, j.

En particulier une telle matrice a ses termes diagonaux tous nuls : $a_{ii} = -a_{ii}$ pour tout i.

Exemple 8.2 $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est antisymétrique. Sa diagonalisation a été faite exercice 4.20. ▪

Exercice 8.3 Montrer : si A est anti-symétrique réelle alors $(A.\vec{v}, \vec{v})_{\mathbb{R}^n} = 0$ pour tout $\vec{v} \in \mathbb{R}^n$. Autrement dit A transforme tout vecteur réel en un vecteur réel orthogonal. En déduire que la seule valeur propre réelle éventuelle de A est $\lambda = 0$.

Réponse. On a $(A.\vec{v}, \vec{v})_{\mathbb{R}^n} = -(\vec{v}, A.\vec{v})_{\mathbb{R}^n} = -(A.\vec{v}, \vec{v})_{\mathbb{R}^n}$ pour tout $\vec{v} \in \mathbb{R}^n$, d'où $(A.\vec{v}, \vec{v})_{\mathbb{R}^n} = 0$ pour tout $\vec{v} \in \mathbb{R}^n$. Donc si $A.\vec{v} = \lambda\vec{v}$ avec $\lambda \in \mathbb{R}$ et $\vec{v} \neq 0$ (vecteur propre), on a $\lambda||\vec{v}||^2_{\mathbb{R}^n} = 0$, donc $\lambda = 0$. ▪

Proposition 8.4 *Si A est une matrice antisymétrique réelle, alors la matrice iA est hermitienne, i.e. $(iA)^* = iA$, cf. définition 1.55. Et donc iA est diagonalisable dans une b.o.n. de $\mathbb{C}^n$, et toutes ses valeurs propres sont réelles. Donc A est diagonalisable dans une b.o.n. de $\mathbb{C}^n$, et toutes ses valeurs propres sont imaginaires pures ou nulles.*

Preuve. $\overline{iA} = -i\overline{A} = -iA$ car A est réelle. Donc $\overline{iA}^T = -iA^T = +iA$ car A est antisymétrique (déjà fait à l'exercice 1.60). Donc iA est hermitienne, donc diagonalisable dans $\mathbb{C}$, cf. théorème 7.1 : il existe une b.o.n $(\vec{e}_k)$ dans $\mathbb{C}^n$ telle que $iA.\vec{e}_k = \lambda_k\vec{e}_k$, les λ_k étant tous réels. Donc $A.\vec{e}_k = -i\lambda_k\vec{e}_k$, et les $\mu_k = -i\lambda_k$ sont les valeurs propres de A, sont imaginaires purs ou nulles, et $(\vec{e}_k)_{k=1,\dots,n}$ est une b.o.n de vecteurs propres associés. ▪

Exercice 8.5 Montrer : si n impair et $A \in \mathbb{R}^{n^2}$ est anti-symétrique, alors 0 est valeur propre de A et A donc n'est pas inversible.

Réponse. On a $\det(A) = \det(A^T) = \det(-A) = (-1)^n \det(A) = -\det(A)$ si n est impair, et donc $\det(A) = 0$ (donc A n'est pas inversible), donc 0 est valeur propre. ▪

Exercice 8.6 Montrer : si A est anti-symétrique réelle, alors A^2 est symétrique négative. En déduire le résultat déjà vu : si λ est valeur propre non nulle de A, alors $\lambda \in i\mathbb{R}$.

Réponse. On a (utilisation implicite de la base canonique) $(A^2.\vec{x}, \vec{y})_{\mathbb{R}^n} = (A.\vec{x}, A^T.\vec{y})_{\mathbb{R}^n} = (A.\vec{x}, -A.\vec{y})_{\mathbb{R}^n} = (\vec{x}, -A^T.A.\vec{y})_{\mathbb{R}^n} = (\vec{x}, A^2.\vec{y})_{\mathbb{R}^n}$ pour tout $\vec{x}, \vec{y} \in \mathbb{R}^n$, et A^2 est symétrique réelle. Et $(A^2.\vec{x}, \vec{x})_{\mathbb{R}^n} = (A.\vec{x}, -A.\vec{x})_{\mathbb{R}^n} = -||A.\vec{x}||^2_{\mathbb{R}^n} \leq 0$. Donc A^2 est diagonalisable et ses valeurs propres μ sont réelles négatives. Or si λ est valeur propre de A alors λ^2 est valeur propre de A^2, donc $\lambda^2 = \mu$ est réelle négative, donc $\lambda = \pm i\sqrt{-\mu}$ est imaginaire pure ou nulle. ▪

Exercice 8.7 Montrer : si A est anti-symétrique réelle, si $\lambda \in i\mathbb{R}$ est une valeur propre non nulle, alors $\overline{\lambda} = -\lambda \in i\mathbb{R}$ est aussi valeur propre : si on connaît une valeur propre non nulle, on en connaît deux.

Réponse. Si $A\vec{v} = \lambda\vec{v}$ avec $\vec{v} \neq 0$ et $\lambda \neq 0$, alors de $\overline{A\vec{v}} = \overline{\lambda\vec{v}}$, donc $A\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}$ car A est réelle, donc $\overline{\lambda}$ est valeur propres de A associé au vecteur propre conjuguée. ▪

Exercice 8.8 Montrer : si $\alpha \in \mathbb{R}^*$ et A est antisymétrique réelle, alors $A + \alpha I$ est inversible.

Réponse. Posons $M = (A^T + \alpha I)(A + \alpha I)$. On a $(Mx, x) = ((A + \alpha I)x, (A + \alpha I)x) = ||(A + \alpha I)x||^2$, avec $\alpha \notin i\mathbb{R}$ donc α n'est pas valeur propre (pour $\alpha \neq 0$), donc M est définie positive, donc inversible, donc $A + \alpha I$ est inversible. ▪

Exercice 8.9 Montrer : si A est anti-symétrique réelle, si $\vec{x}$ et $\vec{y}$ sont vecteurs propres de A de valeurs propres respectives λ et μ telles que $\lambda \neq \mu$, alors $\vec{x} \perp \vec{y}$ dans $\mathbb{C}^n$, i.e. $(\vec{x}, \vec{y})_{\mathbb{C}^n} = 0$.

Réponse. On a $(A\vec{x}, \vec{y})_{\mathbb{C}^n} = \lambda(\vec{x}, \vec{y})_{\mathbb{C}^n}$ et, A étant réelle, $(A\vec{x}, \vec{y})_{\mathbb{C}^n} = (\vec{x}, A^*\vec{y})_{\mathbb{C}^n} = (\vec{x}, A^T\vec{y})_{\mathbb{C}^n} = -(\vec{x}, A\vec{y})_{\mathbb{C}^n} = -(\vec{x}, \mu\vec{y})_{\mathbb{C}^n} = -\overline{\mu}(\vec{x}, \vec{y})_{\mathbb{C}^n}$. D'où $\lambda(\vec{x}, \vec{y})_{\mathbb{C}^n} = -\overline{\mu}(\vec{x}, \vec{y})_{\mathbb{C}^n} = +\mu(\vec{x}, \vec{y})_{\mathbb{C}^n}$, car $\mu \in i\mathbb{R}$ donc $\overline{\mu} = -\mu$. D'où $\lambda(\vec{x}, \vec{y})_{\mathbb{C}^n} = \mu(\vec{x}, \vec{y})_{\mathbb{C}^n}$. D'où $(\vec{x}, \vec{y})_{\mathbb{C}^n} = 0$ si $\lambda \neq \mu$. ▪

Exercice 8.10 En déduire : si A est anti-symétrique réelle, alors A est décomposable par blocs diagonaux réels, blocs 2*2 correspondants aux valeurs propres non nulles, et bloc diagonal correspondant à la valeur propre nulle éventuelle. Le faire pour la matrice $\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$.

Réponse. On a vu qui si λ étant valeur propre, alors $\overline{\lambda}$ l'était aussi, avec $\lambda = ia \in i\mathbb{R}$. D'où le polynôme caractéristique (à une constante multiplicative près) de la forme $P(\lambda) = \lambda^{k_0} \prod_k (\lambda - ia_k)(\lambda + ia_k) = \lambda^{k_0} \prod_k (\lambda^2 + a_k^2)$, si $\lambda = 0$ est valeur propre de multiplicité k_0.

Et la matrice $D = \lambda \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ est similaire à la matrice $B = \lambda \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, i.e. il existe un matrice inversible P telle que $D = P^{-1}BP$, voir exemple 8.2. D'où les blocs diagonaux. ▪

8.2 Matrices normales

Définition 8.11 Une matrice $A \in \mathbb{C}^{n^2}$ est normale ssi elle commute avec son adjoint, i.e. ssi $A.A^* = A^*.A$.

(Et on rappelle qu'elle est orthonormale (ou orthogonale) ssi $A.A^* = A.A^* = I$, ce qu'on ne suppose pas ici.)

Exercice 8.12 Montrer qu'une matrice symétrique réelle est normale.

Réponse. Ici $A^* = A^T$ (matrice réelle) et $A = A^T$ (matrice symétrique), donc $A.A^T = A^2 = A^T.A$. ▪

Exercice 8.13 Montrer qu'une matrice antisymétrique réelle est normale.

Réponse. $A^* = A^T = -A$, donc $AA^* = A(-A) = -A^2 = (-A)A = A^*A$. ▪

Exemple 8.14 $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ n'est pas normale : $A.A^T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $A^T.A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. ▪

Exercice 8.15 Montrer que : A est normale alors pour tout $\vec{x} \in \mathbb{C}^n$ on a $||A\vec{x}||_{\mathbb{C}^n} = ||A^*\vec{x}||_{\mathbb{C}^n}$.

Réponse. Pour $\vec{x}, \vec{y} \in \mathbb{C}^n$ on a ici $(AA^*\vec{x}, \vec{y})_{\mathbb{C}^n} = (A^*A\vec{x}, \vec{y})_{\mathbb{C}^n}$, donc $(A^*\vec{x}, A^*\vec{y})_{\mathbb{C}^n} = (A\vec{x}, A\vec{y})_{\mathbb{C}^n}$. ▪

Exercice 8.16 Montrer que : A est une matrice normale, si $(\lambda, \vec{v})$ est un couple vp-vp pour A alors $(\overline{\lambda}, \vec{v})$ est un couple vp-vp pour A^*.

Réponse. Si $\lambda \in \mathbb{C}$ et $A^*.A = A.A^*$ on a $(A - \lambda I)$ matrice normale : en effet $(A - \lambda I)^* = A^* - \overline{\lambda} I$ donne $(A - \lambda I)(A - \lambda I)^* = A.A^* - (\lambda + \overline{\lambda})A + \lambda\overline{\lambda} = A^*.A - (\overline{\lambda} + \lambda)A + \overline{\lambda}\lambda = (A - \lambda I)^*(A - \lambda I)$.

D'où, avec l'exercice précédent, pour $\lambda \in \mathbb{C}$ et $\vec{v} \in \mathbb{C}^n$ on a $||(A - \lambda I)\vec{v}|| = ||(A - \lambda I)^*\vec{v}|| = ||(A^* - \overline{\lambda} I)\vec{v}||$. En particulier vrai quand $(\lambda, \vec{v})$ est un couple vp–vp de A, qui donne $(\overline{\lambda}, \vec{v})$ est un couple vp–vp de A^*.

Et on vérifie sur la matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et son vecteur propre $\vec{v}_+ = \begin{pmatrix} 1 \\ -i \end{pmatrix} \in \mathbb{C}^2$ associé à la valeur propre $+i$ que $A^*.\vec{v}_+ = (-A).\vec{v}_+ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} -i \\ -1 \end{pmatrix} = -i \begin{pmatrix} 1 \\ -i \end{pmatrix} = -i\vec{v}_+$. ▪

Exercice 8.17 Montrer : si A est normale, alors A est diagonalisable dans $\mathbb{C}$. En déduire qu'une matrice anti-symétrique réelle est diagonalisable dans $\mathbb{C}$.

Réponse. Démontrons-le par récurrence sur la dimension de l'espace $\mathbb{C}^n$. C'est trivial pour $n = 1$. Soit $n \geq 2$. Soit λ une valeur propre (une racine du polynôme caractéristique existe toujours dans $\mathbb{C}$). Soit $\vec{v}$ un vecteur propre associé. Soit $V = \mathrm{Vect}\{\vec{v}\}^\perp$: on a V est stable par A. En effet, pour $\vec{w} \in V$, avec l'exercice précédent on a $(A\vec{w}, \vec{v})_{\mathbb{C}^n} = (\vec{w}, A^*\vec{v})_{\mathbb{C}^n} = (\vec{w}, \overline{\lambda}\vec{v})_{\mathbb{C}^n} = \lambda(\vec{w}, \vec{v})_{\mathbb{C}^n} = 0$. De même V est stable par A^* : $(A^*\vec{w}, \vec{v})_{\mathbb{C}^n} = (\vec{w}, A\vec{v})_{\mathbb{C}^n} = (\vec{w}, \lambda\vec{v})_{\mathbb{C}^n} = \overline{\lambda}(\vec{w}, \vec{v})_{\mathbb{C}^n} = 0$. Soit $B = A_{|V}$ la restriction de A à V : on vient de voir grâce à la stabilité que $AA^* = A^*A$ s'écrit $BB^* = B^*B$ dans V. Avec $\dim V = n-1$, donc B est diagonalisable dans V. Donc A est diagonalisable dans $\mathbb{C}^n$. ▪

9 Trigonalisation de matrices

9.1 Définitions

Rappel : une matrice $T \in K^{n^2}$ est triangulaire supérieure ssi $T_{ij} = 0$ pour tout $i > j$ (nulle sous la diagonale). Elle est triangulaire inférieure si sa transposée est triangulaire supérieure.

Soit $(\vec{E}_i)$ la base canonique de K^n. Donc $T \in K^{n^2}$ est triangulaire supérieure ssi, pour tout j :

$$T.\vec{E}_j = \sum_{i=1}^{j} T_{ij}\vec{E}_j, \qquad [T.\vec{E}_j] = \begin{pmatrix} T_{1j} \\ \vdots \\ T_{jj} \\ 0 \\ \vdots \end{pmatrix}. \tag{9.1}$$

Définition 9.1 Une matrice $A \in K^{n^2}$ est trigonalisable (ou triangulisable) dans K ssi il existe une matrice inversible $P \in K^{n^2}$ et une matrice triangulaire supérieure $T \in K^{n^2}$ telles que :

$$T = P^{-1}.A.P, \tag{9.2}$$

i.e. ssi A est semblable à une matrice triangulaire supérieure.

Remarque 9.2 Ainsi P est une matrice de changement de base qui, lorsque A est la matrice d'un endomorphisme dans une base donnée exprime cet endomorphisme sous forme matrice triangulaire dans la nouvelle base. ▟

Donc A est trigonalisable ssi il existe P inversible et T triangulaire supérieure telles que :

$$A.P = P.T, \tag{9.3}$$

soit, notant $P = (\,[\vec{p}_1] \quad ... \quad [\vec{p}_n]\,)$, pour tout $j = 1, ..., n$:

$$A.\vec{p}_j = P.T.\vec{E}_j = \sum_{i=1}^{j} T_{ij}\vec{p}_i. \tag{9.4}$$

9.2 Exemples et contre-exemples

Exemple 9.3 Toute matrice triangulaire supérieure (en particulier toute matrice diagonale) est trigonalisable (c'est la moindre des choses !) : prendre $P = I$ dans (9.2). ▟

Exercice 9.4 Montrer que si T et U sont triangulaires supérieures, alors $T.U$ est triangulaire supérieure, et $(T.U)_{ii} = T_{ii}U_{ii}$ pour tout i (les éléments diagonaux de $T.U$ sont ceux les produits des éléments diagonaux).

Réponse. Soit $\vec{u}_j = \sum_{i=1}^{j} u_{ij}\vec{E}_i$ le j-ème "vecteur colonne" de U (U est triangulaire supérieure). Alors la j-ème colonne de $T.U$ est $T.\vec{u}_j = \sum_{i=1}^{j} u_{ij}T.\vec{E}_i$ (linéarité de T), soit $T.\vec{u}_j = \sum_{i=1}^{j} u_{ij}\sum_{k=1}^{i} T_{ki}\vec{E}_k$ (T est triangulaire supérieure), soit $T.\vec{u}_j = \sum_{i=1}^{j}\sum_{k=1}^{i} T_{ki}u_{ij}\vec{E}_k \in \mathrm{Vect}\{\vec{E}_1, ..., \vec{E}_j\}$. D'où $T.U$ est triangulaire supérieure, et $(T.U)_{jj}$ est donné par le coefficient devant $\vec{E}_k$ pour $k = j$, donc pour $i = j$ et $k = j$, donc par $T_{jj}u_{jj}$.

(Ou encore, $(T.U)_{jj} = [\vec{\ell}_j].[\vec{u}_j]$ où $\vec{\ell}_j$ est la ligne j de T, avec $\vec{\ell}_j = \sum_{k=j}^{n} T_{jk}\vec{E}_k$, car T est triangulaire supérieure, d'où $(T.U)_{jj} = \sum_{k=j}^{n} T_{jk}\sum_{i=1}^{j} u_{ij}[\vec{E}_k].[\vec{E}_i] = \sum_{k\geq j}\sum_{i\leq j} T_{jk}u_{ij}\delta_{ik} = T_{jj}u_{jj}$.) ▟

Exercice 9.5 Montrer que si T est triangulaire supérieure réelle, si $T^2 = I$ et si T est positive alors nécessairement $T = I$. (C'est faux si T n'est pas positive, cf. exemple 5.28.)

Réponse. $T^2 = I$ donne $T_{ii}^2 = 1$ (exercice précédent), donc $T_{ii} = \pm 1$. Comme T est positive, nécessairement $T_{ii} = 1$: en effet, T triangulaire donne $\lambda_i = T_{ii}$ est valeur propre, et si $\vec{v}_i$ est vecteur propre associé à T_{ii} alors $0 \leq (T.\vec{v}_i, \vec{v}_i)_{K^n} = (\lambda_i\vec{v}_i, \vec{v}_i)_{K^n} = T_{ii}||\vec{v}_i||^2_{K^n}$ donne $T_{ii} \geq 0$. Et le j-ème "vecteur colonne" de T^2 est $T^2.\vec{E}_j = T.\vec{t}_j$ où $\vec{t}_j$ est le j-ème "vecteur colonne" de T. Donc $T^2 = I$, donne $T.\vec{t}_j = \vec{E}_j$. D'où $\vec{t}_j = \vec{E}_j$ puisque $\vec{t}_j$ est donné par (9.1). D'où $T = I$. ▟

Exercice 9.6 Montrer que $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ n'est pas trigonalisable dans $\mathbb{R}$.

Réponse. Si $A.P = P.T$, avec $T = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix} \in \mathbb{R}^{n^2}$, alors pour l'égalité de la première colonne :

$$\begin{cases} P_{21} = P_{11}T_{11}, \\ -P_{11} = P_{21}T_{11}. \end{cases} \tag{9.5}$$

Donc $P_{21} = -P_{21}T_{11}^2$. Mais $T_{11} \neq 0$ car $\det T = T_{11}T_{22} \neq 0$ puisque T est triangulaire inversible et $1 = \det A = \det P \det T \neq 0$.

Donc dans $\mathbb{R}$, nécessairement $T_{11}^2 > 0$, donc $P_{21} = 0$ (sinon $T_{11}^2 = -1$ impossible dans $\mathbb{R}$). Donc $P_{11} = 0$, donc la première colonne de P est nulle. Donc P n'est pas inversible. Absurde si A est trigonalisable. Donc A n'est pas trigonalisable dans $\mathbb{R}^n$. Ou encore utiliser la proposition suivante. ▟

Exercice 9.7 Une matrice $A \in \mathbb{C}^{n^2}$ est nilpotente ssi il existe $k \in \mathbb{N}$ telle que $A^k = 0$. Une matrice $T = [T_{ij}]$ est triangulaire supérieure stricte ssi $T_{ij} = 0$ pour tout $j \leq i$. Montrer qu'une matrice $T \in \mathbb{C}^{n^2}$ triangulaire supérieure stricte est nilpotente.

Réponse. $T.\vec{E}_j = \sum_{i=1}^{j-1} T_{ij}\vec{E}_i \in \mathrm{Vect}\{\vec{E}_1, ..., \vec{E}_{j-1}\}$, donc T^2 a sa diagonale nulle (par définition de triangulaire stricte) et la première diagonale au dessus également nulle. Puis par récurrence $T^n = 0$. ▟

9.3 Théorème de trigonalisation des matrices complexes

Proposition 9.8 *Si A est trigonalisable, alors ses valeurs propres sont données par la diagonale de T avec la multiplicité correspondante.*

Preuve. $\det(A-\lambda I) = \det(P.T.P^{-1}-\lambda I) = \det(P(T-\lambda I).P^{-1}) = \det P \det(T-\lambda I)\det(P^{-1}) = \det(T-\lambda I)$: donc A et T ont le même polynôme caractéristique, donc les mêmes valeurs propres. ▪

Exemple 9.9 Retour à l'exercice 9.6 : les valeurs propres de A sont $\pm i$ (non réelles), et donc la matrice triangulaire T a sa diagonale qui n'est pas réelle, cf. proposition précédente. Donc A n'est pas trigonalisable dans $\mathbb{R}$. ▪

Théorème 9.10 *(Décomposition de Schur.) Si $A \in \mathbb{C}^{n^2}$ (matrice complexe $n*n$), alors A est semblable à une matrice triangulaire : il existe $P \in \mathbb{C}^{n^2}$ inversible et il existe $T \in \mathbb{C}^{n^2}$ triangulaire supérieure telles que :*

$$A = P.T.P^{-1}, \tag{9.6}$$

la diagonale de T étant constituée des valeurs propres de A.

De plus, parmi toutes les décompositions (9.6) possibles, on peut en choisir une telle que P est une matrice normale, i.e. t.q. ;

$$P^*.P = I, \tag{9.7}$$

les colonnes de P formant donc une b.o.n. dans $\mathbb{C}^n$.

On encore, on peut en choisir une telle que T soit formée des blocs de Jordan.

Preuve. Démonstration similaire à la démonstration du théorème 5.2, la différence étant que dans la formule (5.4) la première ligne est pleine (pas de zéros a priori), et on replace B diagonalisable par trigonalisable. ▪

Exercice 9.11 Trigonaliser $A = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$ avec a et b réels (matrice générique d'une équation différentielle linéaire d'ordre 2 mise sous forme système différentiel d'ordre 1).

Réponse. On a $p(\lambda) = \det(A-\lambda I) = -\lambda(b-\lambda) - a = \lambda^2 - b\lambda - a$. Discriminant : $\Delta = b^2+4a$.

Cas $b^2+4a \neq 0$: 2 racines distinctes et A est diagonalisable dans $\mathbb{C}$.

Cas $b^2+4a = 0$, une racine double $\lambda = \frac{b}{2}$, et $\mathrm{Ker}(A-\lambda I) = \mathrm{Vect}\{\begin{pmatrix} 1 \\ \lambda \end{pmatrix}\}$ de dimension 1 alors que λ est valeur de multiplicité 2. Donc $A = \begin{pmatrix} 0 & 1 \\ -\lambda^2 & 2\lambda \end{pmatrix}$ n'est pas diagonalisable.

Et A est trigonalisable avec $T = \begin{pmatrix} \lambda & t \\ 0 & \lambda \end{pmatrix}$ (car λ est vp double) et par exemple $P = \begin{pmatrix} 1 & p \\ \lambda & q \end{pmatrix}$ (la première colonne de P est imposée par $A.\vec{e}_1 = \lambda\vec{e}_1$). Et $A.P = P.T$ donne pour la deuxième colonne $\begin{pmatrix} q \\ -\lambda^2 p + 2\lambda q \end{pmatrix} = \begin{pmatrix} t+\lambda p \\ \lambda t + \lambda q \end{pmatrix}$. Donc $q = t+\lambda p$ (les deux équations sont les mêmes). Seul $t=0$ est interdit car dans ce cas P n'est pas inversible, sinon toutes les combinaisons vérifiant $q = t+\lambda p$ conviennent.

Si on veut par exemple que $T = \lambda\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (soit λ fois le premier bloc de Jordan), on prend $t=\lambda$, puis par exemple $p=0$ et $q=\lambda$, soit $P = \begin{pmatrix} 1 & 0 \\ \lambda & \lambda \end{pmatrix}$ (matrice mal conditionnée pour λ grand).

Si on préfère P orthonormale (pour travailler dans une nouvelle base orthonormée), on prend par exemple $P = \begin{pmatrix} \frac{1}{\sqrt{1+\lambda^2}} & \frac{p}{\sqrt{1+\lambda^2}} \\ \frac{\lambda}{\sqrt{1+\lambda^2}} & \frac{q}{\sqrt{1+\lambda^2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \\ \frac{\lambda}{\sqrt{1+\lambda^2}} & -\frac{1}{\sqrt{1+\lambda^2}} \end{pmatrix}$ (on a bien $P^T.P = I$), et donc $t = q - \lambda p = -(1+\lambda^2)$. ▪

Exercice 9.12 Trigonaliser $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Réponse. Voir exercice 7.5 : ici A est diagonalisable dans $\mathbb{C}$. ▪

Exercice 9.13 Montrer que si A est définie positive et trigonalisable dans $\mathbb{R}^n$, alors T est définie positive et ses valeurs propres λ_i sont strictement positives.

Réponse. Pour $\vec{x} \neq \vec{0}$ on a $0 < (A.\vec{x},\vec{x})_{\mathbb{R}^n} = (P.T.P^T.\vec{x},\vec{x})_{\mathbb{R}^n} = (T.\vec{y},\vec{y})_{\mathbb{R}^n}$ où on a posé $\vec{y} = P^T.\vec{x}$, et donc $\vec{y} \neq \vec{0}$ car P^T est inversible et $\vec{x} \neq \vec{0}$. Donc T est définie positive.

Et pour λ valeur propre, soit $\vec{v} \in \mathrm{Ker}(A-\lambda I)$ un vecteur propre associé (donc $\vec{v} \neq \vec{0}$). Donc $0 < (A.\vec{v},\vec{v})_{\mathbb{R}^n} = (\lambda\vec{v},\vec{v})_{\mathbb{R}^n} = \lambda||\vec{v}||_{\mathbb{R}^n}$, donc $0 < \lambda$. ▪

Références

[1] Cairoli B. : *Algèbre linéaire*. Presses Polytechniques et Universitaires Romandes, 1991.

[2] Ciarlet P.G. : *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, 1992.

[3] Golub G., Van Loan C. : *Matrix Computations*. John Hopkins University Press, 1996.

[4] Lascaux P., Théodor R. : *Analyse numérique matricielle appliquée à l'art de l'ingénieur*. Masson, 1998.

[5] Strang G. : *Linear Algebra and its Applications*. Harcourt Brace (1988).