Matrices normales

Nous allons nous intéresser à une classe assez générale de matrices, qui sont celles qui sont diagonalisables dans une base orthonormée et qui sont appelées matrices normales et dont nous verrons des sous ensembles particuliers telles que les matrices orthogonales, les matrices unitaires et les matrices symétriques.

I Problématique

Nous avons vu tout l'intérêt que présente la diagonalisation des matrices réelles ou complexes. Elle permet un traitement des systèmes différentiels comme en analyse modale ou encore la détermination des axes propres d'une conique (ellipse, hyperbole, parabole) ou encore en modélisation mécanique, la détermination des contraintes principales, et bien d'autres encore.

Les vecteurs propres représentent bien souvent dans de nombreuses applications, des vecteurs géométriques et il est fort intéressant de pouvoir exhiber une base orthonormée de vecteurs propres.

La question qui se pose alors est la suivante : quelle est la nature des matrices qui possèdent une base orthonormée de vecteurs propres au sens du produit scalaire réel ou bien complexe (hermitien) ?

Pour le savoir, considérons une telle matrice A, et P une matrice formée d'une base de vecteurs propres orthonormés. Commençons par le cas où A est réelle, tout comme P et les valeurs propres de A.

En notant D la matrice diagonale des vecteurs propres associés à P, nous avons :

$$\boxed{A\,P = P\,D}$$

Le caractère orthornormé des vecteurs colonnes de P s'écrit :

$$\boxed{{}^{t}P\,P = P\,{}^{t}P = I}$$

Nous en déduisons :

$$^{t}(A\,P) = {}^{t}(P\,D)$$

$$^{t}P\,{}^{t}A = {}^{t}D\,{}^{t}P$$

$$^{t}P\,{}^{t}A = D\,{}^{t}P$$

$$A\,P\,{}^{t}P\,{}^{t}A = P\,D\,D\,{}^{t}P$$

$$A\,{}^{t}A = P\,D^2\,{}^{t}P$$

En transposant cette dernière relation nous voyons une propriété apparaître :

$$\boxed{A\,{}^{t}A = {}^{t}A\,A}$$

Autrement dit, si une matrice réelle est diagonalisable dans une base orthonormée de vecteurs propres au sens du produit scalaire des vecteurs sur les nombres réels, alors elle commute dans le produit avec sa transposée.

Voyons alors ce qu'il en est pour une matrice A complexe, diagonalisable dans une base de vecteurs propres complexes, orthonormés pour le produit scalaire hermitien.

La démarche se calque sur la précédente. Il suffit de remplacer la relation sur P par :

$$\boxed{{}^{t}\overline{P}\,P = P\,{}^{t}\overline{P} = I}$$

Elle aboutit à :

$$\boxed{A\,{}^{t}\overline{A} = P\,|D|^2\,{}^{t}\overline{P}}$$

Où $|D|^2$ désigne la matrice diagonale dont les coefficients diagonaux sont les carrés des modules de ceux de D.

En prenant la transconjuguée de la relation précédente, nous avons donc la propriété remarquable :

$$\boxed{A\,{}^{t}\overline{A} = {}^{t}\overline{A}\,A}$$

Autrement dit, si une matrice complexe est diagonalisable dans une base orthonormée de vecteurs propres au sens du produit scalaire hermitien des vecteurs complexes, alors elle commute dans le produit avec sa transconjuguée.

Se pose alors la question des points de vue réciproque : Une matrice réelle, qui commute avec sa transposée est elle diagonalisable dans une base orthonormée réelle, et une matrice complexe qui commute avec sa transconjuguée est elle diagonalisable dans une base orthonormée complexe ?

La réponse a la première question est négative comme le montre l'exemple simple d'une matrice associée à une rotation vectorielle de 90° dans un plan :

$$\mathrm{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad {}^{t}\mathrm{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\mathrm{A}\ {}^{t}\mathrm{A} = {}^{t}\mathrm{A}\ \mathrm{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}$$

Pourtant A n'est pas diagonalisable dans $\mathbb{R}$, en effet les valeurs propres de A sont racines du polynôme caractéristique en λ

$$\boxed{\mathrm{Det}(\mathrm{A} - \lambda\,\mathrm{I}) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1}$$

Et ce polynôme n'admet pas de racines réelles.

En revanche A commute avec sa transconjuguée et est bien diagonalisable en base orthonormée complexe, ses valeurs propres étant i et –i.

Nous voyons donc que nous sommes amenés à étudier les propriétés des matrices complexes qui commutent avec leur transconjuguée et qui sont qualifiées de matrices normales, et nous allons voir que ces matrices ne sont rien d'autres que les matrices diagonalisables dans une base orthonormée complexe, ce qui sera déjà un résultat remarquable.

II Définitions et exemples

Etant donné une matrice A, cette matrice est dite **normale** si elle commute avec la transposée de sa conjuguée, autrement dit si :

$$A\,{}^t\overline{A} = {}^t\overline{A}\,A$$

Par commodité, il est d'usage de noter :

$$A^* = {}^t\overline{A}$$

Une matrice normale est donc une matrice qui vérifie :

$$A\,A^* = A^*\,A$$

Des sous familles de matrices normales très utiles peuvent être identifiées il s'agit :

Des matrices réelles symétriques, c'est-à-dire pour lesquelles :

$${}^tA = A$$

Nous verrons qu'elles ont des valeurs propres réelles qui peuvent être associées à une base de vecteurs propres réels.

Des matrices complexes hermitiennes, c'est-à-dire pour lesquelles :

$${}^t\overline{A} = A$$

Des matrices réelles orthogonales, c'est-à-dire pour lesquelles :

$${}^tA\,A = A\,{}^tA = I \qquad \text{soit encore} \qquad {}^tA = A^{-1}$$

Des matrices complexes unitaires, c'est-à-dire pour lesquelles :

$${}^t\overline{A}\,A = A\,{}^t\overline{A} = I \qquad \text{soit encore} \qquad {}^t\overline{A} = A^{-1}$$

De manière générale, toute matrice diagonalisable dans une base orthonormée complexe est une matrice normale, comme nous l'avions vu au chapitre 1. Nous allons voir que la réciproque est vraie.

III Propriétés de diagonalisation

C'est la propriété caractéristique des matrices normales :

Toute matrice normale complexe est diagonalisable dans une base orthonormée complexe.

Une autre propriété voisine est :

Toute matrice symétrique réelle est diagonalisable dans une base orthonormée réelle et ses valeurs propres sont réelles.

Nous allons prouver le premier théorème en faisant un raisonnement par récurrence sur l'ordre n de la matrice.

La première étape de la démonstration consiste à noter qu'une matrice complexe quelconque possède toujours au moins une valeur propre λ comme une racine de son polynôme caractéristique.

La seconde étape consiste à considérer le sous espace propre E_λ associé à cette valeur propre et à former une base orthonormée complexe $P_1, P_2, \ldots, P_m$ de ce sous espace de dimension m, puis à la compléter en une base orthonormée $P_1, P_2, \ldots, P_m, P_{m+1}, \ldots, P_n$ de l'espace vectoriel des vecteurs colonnes complexes de taille n, que nous noterons pour simplifier $\mathbb{C}_{n,1}$. Le procédé pratique pour faire ce genre de choses est le procédé d'orthonormalisation de Schmidt que nous exposerons dans un autre fichier.

La troisième étape consiste à montrer que l'espace engendré par les vecteurs $P_{m+1}, \ldots, P_n$ est stable par la matrice A. Soit donc une des colonnes P_j précédentes, montrons que $A\,P_j$ est orthogonale à chaque colonne P_i de $P_1, P_2, \ldots, P_m$. Pour cela notons que :

$$A\,P_i = \lambda\,P_i$$

$$A^*A\,P_i = A^*(\lambda\,P_i)$$

$$A\,(A^*\,P_i) = \lambda(A^*\,P_i)$$

Le vecteur colonne $A^* P_i$ est donc un vecteur propre de A associé à la valeur propre λ. Il se décompose donc sur les vecteurs de base de E_λ.

$$\boxed{A^* P_i = x_1 P_1 + x_2 P_2 + \cdots + x_m P_m}$$

Il est maintenant aisé d'en déduire le produit scalaire complexe (hermitien) des colonnes $A P_j$ et P_i.

$${}^t\left(\overline{A P_j}\right) P_i = {}^t\overline{P_j}\, {}^t\overline{A}\, P_i = {}^t\overline{P_j}\, A^* P_i$$

$$= {}^t\overline{P_j}\, (x_1 P_1 + x_2 P_2 + \cdots + x_m P_m)$$

$$= x_1\, {}^t\overline{P_j}\, P_1 + x_2\, {}^t\overline{P_j}\, P_2 + \cdots + x_m\, {}^t\overline{P_j}\, P_m = 0$$

Les vecteurs colonnes $A P_j$ et P_i sont donc bien orthogonaux pour le produit scalaire hermitien. On en déduit que $A P_j$ a des coordonnées nulles sur les vecteurs $P_1, P_2, .., P_m$ de la base $P_1, P_2, .., P_n$ ce qui traduit le fait que le sous espace orthogonal à E_λ est stable par A.

Formons alors une matrice P carrée d'ordre n avec les colonnes $P_1, P_2, .., P_n$. Cette matrice est alors unitaire, c'est-à-dire vérifie, rappelons le :

$$P^* = P^{-1}$$

Nous avons alors, en faisant apparaître les colonnes de la matrice produit :

$$A P = \left(\begin{bmatrix} \vdots \\ A P_1 \\ \vdots \\ \vdots \end{bmatrix} \cdots \begin{bmatrix} \vdots \\ A P_m \\ \vdots \\ \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ A P_{m+1} \\ \vdots \\ \vdots \end{bmatrix} \cdots \begin{bmatrix} \vdots \\ A P_{m+k} \\ \vdots \\ \vdots \end{bmatrix} \right)$$

Or le caractère orthogonal des colonnes $A P_j$ et P_i vu précédemment se traduit par le fait que :

$${}^t\overline{P_i}\, \left(A P_j\right) = 0$$

Soit en écriture matricielle détaillée :

$$\left(\ldots \ldots \overline{P_i} \ldots \ldots \ldots \right) \begin{pmatrix} \vdots \\ \vdots \\ A\,P_j \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = 0$$

On en déduit l'allure de la matrice :

$$P^*\ A\ P = \begin{pmatrix} \begin{bmatrix} \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & B & \end{bmatrix} \end{pmatrix}$$

Cette matrice est formée de 4 blocs :

- un bloc carré formé par une matrice diagonale d'ordre n
- un bloc carré formé d'une matrice B d'ordre k = n – m
- deux blocs rectangulaires ne contenant que des zéros

En prenant la transconjuguée de cette matrice, nous avons alors :

$$P^*\ A^*\ \ P = \begin{pmatrix} \begin{bmatrix} \overline{\lambda} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{\lambda} \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & B^* & \end{bmatrix} \end{pmatrix}$$

Notons alors que la multiplication des deux précédentes matrices commute.

En effet, compte tenu de $P^*\ P = I$, nous avons, d'une part :

$$(P^*\ A\ P)\ (P^*\ A^*\ \ P) = P^*\ A\ A^*\ \ P$$

$$(P^*\ A^*\ \ P)\ (P^*\ A\ P)\ = P^*\ A^*\ A\ \ \ P$$

Ces deux matrices étant égale compte tenu de : $A\ A^* = A^*\ A$

et d'autre part , en effectuant le produit :

$$(P^* \, A \, P) \, (P^* \, A^* \; P) = \begin{pmatrix} \begin{bmatrix} |\lambda|^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\lambda|^2 \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & B \, B^* & \end{bmatrix} \end{pmatrix}$$

$$(P^* \, A^* \; P)(P^* \, A \, P) \; = \begin{pmatrix} \begin{bmatrix} |\lambda|^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\lambda|^2 \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & B^* B & \end{bmatrix} \end{pmatrix}$$

L'égalité de ces deux matrices conduit alors à la relation :

$$B \, B^* = B^* B$$

B est donc également une matrice normale d'ordre k strictement inférieur à n. En appliquant une hypothèse de récurrence forte, à savoir que toute matrice normale d'ordre strictement inférieur à n (pour $n > 1$) est diagonalisable dans une base orthonormée complexe, l'initialisation étant triviale (toute matrice d'ordre 1 est évidemment diagonalisable dans une base orthonormée qui est le nombre réel 1, l'unique valeur propre étant l'unique terme de cette matrice) nous pouvons donc dire qu'il existe une matrice Q unitaire d'ordre k et une matrice diagonale D' d'ordre k telles que :

$$Q^* \, B \, Q = D'$$

Ou encore :

$$B \; = Q \, D' \, Q^*$$

En notant $[I]_m$ la matrice identité d'ordre m, nous avons alors :

$$A = P\begin{pmatrix} [I]_m \begin{bmatrix} \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{bmatrix} [I]_m{}^{*} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & Q\,D\,Q^{*} & \end{bmatrix} \end{pmatrix} P^{*}$$

Posons alors :

$$P' = P\begin{pmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & Q & \end{bmatrix} \end{pmatrix}$$

Soit :

$$P'^{*} = \begin{pmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & Q^{*} & \end{bmatrix} \end{pmatrix} P^{*}$$

Le calcul par bloc montre alors que P' est unitaire et que :

$$A = P'\begin{pmatrix} \begin{bmatrix} \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{bmatrix} & \begin{bmatrix} & 0 & \end{bmatrix} \\ \begin{bmatrix} & 0 & \end{bmatrix} & \begin{bmatrix} & D' & \end{bmatrix} \end{pmatrix} P'^{*}$$

La matrice A est donc bien diagonalisée dans la base orthonormée complexe des vecteurs propres formant les vecteurs colonnes de P' et ses valeurs propres sont λ et les éléments diagonaux de D', ce qui achève la démonstration