Mémento sur le calcul matriciel

1 Définitions générales

Les définitions qui suivent sont relatives à un corps, généralement $\mathbb{R}$ ou $\mathbb{C}$, noté K. On désigne par $\mathcal{M}_K(m,n)$ l'ensemble des **matrices** à m lignes et n colonnes, à éléments dans K. C'est un espace vectoriel. En particulier, l'addition et la multiplication par un scalaire sont des lois internes :

$$(A,B) \in \mathcal{M}_K(m,n)^2,\ (\lambda,\mu) \in K^2 \quad \Longrightarrow \quad \lambda A + \mu B \in \mathcal{M}_K(m,n)$$

L'élément neutre pour l'addition est la **matrice nulle**, notée O. Dans le cas particulier où $n = 1$, les éléments de $\mathcal{M}_K(m,n)$ sont appelés des **vecteurs**. L'espace vectoriel $\mathcal{M}_K(m,1)$ est isomorphe à K^m.

De manière générale, le **produit matriciel** n'est pas une loi interne. Il n'est défini que si le nombre de colonnes de la première matrice coïncide avec le nombre de lignes de la deuxième matrice :

$$A \in \mathcal{M}_K(m,n),\ B \in \mathcal{M}_K(n,p) \quad \Longrightarrow \quad A\,B \in \mathcal{M}_K(m,p)$$

Plus précisément, le terme général de $C = A\,B$ vaut, pour $i \in [1,m]$ et $j \in [1,p]$:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Le produit matriciel n'est pas commutatif.

La **transposition** d'une matrice n'est pas une loi interne :

$$A \in \mathcal{M}_K(m,n) \quad \Longrightarrow \quad A^\top \in \mathcal{M}_K(n,m)$$

On peut donc toujours calculer les produits matriciels suivants, qui sont des matrices carrées :

$$\begin{array}{l} A^\top A \in \mathcal{M}_K(n,n) \\ A\,A^\top \in \mathcal{M}_K(m,m) \end{array}$$

On vérifie facilement les identités suivantes :

$$\begin{array}{l} (A+B)^\top = A^\top + B^\top \\ (\lambda\,A)^\top = \lambda\,A^\top \\ (A^\top)^\top = A \end{array}$$

Une identité moins évidente est :

$$(A\,B)^\top = B^\top A^\top$$

2 Matrices carrées

On s'intéresse dorénavant aux **matrices carrées**, pour lesquelles $m = n$. On note $\mathcal{M}_K(n,n) = \mathcal{M}_K(n)$. On introduit la notion intuitive de **diagonale**. Le produit matriciel devient une loi interne dans $\mathcal{M}_K(n)$. Relativement à cette loi interne, l'élément neutre est la **matrice identité**, notée I_n, qui comporte des 1 sur sa diagonale et des 0 ailleurs. On appelle **matrice diagonale** une matrice dont tous les éléments non diagonaux sont nuls.

La matrice identité est un exemple de matrice diagonale.

Le **déterminant** d'une matrice carrée A est un scalaire de K, noté $\det A$, dont la définition n'est pas intuitive. Il peut être calculé, par exemple, en « développant suivant la 1$^{\text{ère}}$ colonne » :

$$\det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det A_{i,1}$$

où $A_{i,1}$ désigne la matrice de $\mathcal{M}_K(n-1)$ obtenue en enlevant la $i^{\text{ème}}$ ligne et la 1$^{\text{ère}}$ colonne de A. Sachant que $|a_{1,1}| = a_{1,1}$, on en déduit les formules suivantes :

— Pour $n = 2$:

$$\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$$

— Pour $n = 3$:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{2,1}a_{1,3}a_{3,2} - (a_{3,1}a_{2,2}a_{1,3} + a_{1,2}a_{2,1}a_{3,3} + a_{1,1}a_{2,3}a_{3,2})$$

Le déterminant d'une matrice diagonale est égal au produit de ses termes diagonaux. En particulier, $\det I_n = 1$. On a également les deux propriétés suivantes :

$$\begin{array}{l} \det(A)^\top = \det A \\ (A,B) \in \mathcal{M}_K(n)^2 \quad \Longrightarrow \quad \det(A\,B) = \det A \, \det B \end{array}$$

Certaines matrices carrées A admettent une matrice inverse vis-à-vis du produit matriciel, notée A^{-1}, telle que :

$$A^{-1}\,A = A\,A^{-1} = I_n$$

Ces matrices sont dites **inversibles**. Par conséquent :

$$\det(A\,A^{-1}) = \det I_n$$

donc :

$$\det A \, \det(A^{-1}) = 1$$

Cela montre qu'une matrice inversible A a un déterminant non nul. La condition $\det A \neq 0$ est donc nécessaire pour que A soit inversible. C'est aussi une condition suffisante. Le calcul de l'inverse d'une matrice n'est pas trivial, sauf pour les matrices diagonales :

$$A = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{bmatrix} \quad \Longrightarrow \quad A^{-1} = \begin{bmatrix} {a_{1,1}}^{-1} & 0 & \cdots & 0 \\ 0 & {a_{2,2}}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & {a_{n,n}}^{-1} \end{bmatrix}$$

La **trace** d'une matrice carrée A est un scalaire de K, noté $\mathrm{tr}A$, égal à la somme de ses termes diagonaux. Quelques propriétés de la trace sont :

$$\begin{array}{l} \mathrm{tr}(A^\top) = \mathrm{tr}A \\ \mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B \\ \mathrm{tr}(A\,B) = \mathrm{tr}(B\,A) \end{array}$$

Attention :

$$\mathrm{tr}(A\,B) \neq \mathrm{tr}A \; \mathrm{tr}B$$

3 Matrices orthogonales et matrices unitaires

Si C_1, C_2, $\cdots$, C_n désignent les n colonnes d'une matrice A de $\mathcal{M}_K(n)$, qui sont des vecteurs de K^n, alors **on admet** que A est inversible si et seulement si $(C_1, C_2, \cdots, C_n)$ constitue une base de K^n. Si l'on munit K d'un produit scalaire, un cas particulier intéressant est celui où $(C_1, C_2, \cdots, C_n)$ constitue une base orthonormée, ou « bon », de K^n.

3.1 Matrices orthogonales de $\mathcal{M}_{\mathbb{R}}(n)$

Soit $(C_1, C_2, \cdots, C_n)$ une base de $\mathbb{R}^n$. Cette base est orthonormée, relativement au produit scalaire euclidien « canonique » dans $\mathbb{R}^n$, si pour $(i, j) \in [1, n]^2$:

$$C_i^\top C_j = \delta_{i,j}$$

où $\delta_{i,j}$ vaut 1 si $i = j$ et 0 sinon (**symbole de Kronecker**). Une matrice A dont les colonnes C_1, C_2, $\cdots$, C_n constituent une bon de $\mathbb{R}^n$ possède donc une propriété très remarquable :

$$A^\top A = \begin{bmatrix} C_1^\top \\ C_2^\top \\ \vdots \\ C_n^\top \end{bmatrix} \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix} = \begin{bmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,n} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n,1} & \delta_{n,2} & \cdots & \delta_{n,n} \end{bmatrix}$$

c'est-à-dire que :

$$A^\top A = I_n$$

Or, on sait que A est inversible, puisque $(C_1, C_2, \cdots, C_n)$ constitue une base de $\mathbb{R}^n$, donc nécessairement :

$$A^{-1} = A^\top \tag{1}$$

On appelle **matrice orthogonale** toute matrice de $\mathcal{M}_{\mathbb{R}}(n)$ telle que $A^\top A = I_n$. Ce sont des matrices inversibles, dont l'inverse est particulièrement simple à calculer par (1).

3.2 Matrices unitaires de $\mathcal{M}_{\mathbb{C}}(n)$

Soit $(C_1, C_2, \cdots, C_n)$ une base de $\mathbb{C}^n$. Cette base est orthonormée, relativement au produit scalaire euclidien « canonique » dans $\mathbb{C}^n$, si pour $(i, j) \in [1, n]^2$:

$$C_i^\top \overline{C_j} = \delta_{i,j}$$

où $\overline{C_j}$ désigne le vecteur complexe conjugué de C_j. Il s'ensuit donc que :

$$A^\top \overline{A} = I_n$$

c'est-à-dire :

$$A^{-1} = \overline{A}^\top \tag{2}$$

On appelle **adjointe** d'une matrice A de $\mathcal{M}_{\mathbb{C}}(n)$, et on note A^*, la matrice $\overline{A}^\top$. On appelle **matrice unitaire** toute matrice de $\mathcal{M}_{\mathbb{C}}(n)$ telle que $A^* A = I_n$. Pour ces matrices aussi, l'inverse est très facile à calculer par (2).

4 Matrices semblables

Deux matrices carrées A et A' sont **semblables** s'il existe une matrice inversible P telle que :

$$A' = P^{-1} A P$$

Si tel est le cas, alors il existe également une matrice inversible Q pour laquelle :

$$A = Q^{-1} A' Q$$

Cette matrice est $Q = P^{-1}$. Une matrice est **diagonalisable** si elle est semblable à une matrice diagonale.

On a les deux identités suivantes, pour deux matrices A et A' semblables :

$$\begin{aligned} \det A' &= \det A \\ \mathrm{tr} A' &= \mathrm{tr} A \end{aligned}$$

5 Valeurs propres et vecteurs propres

Une autre propriété des matrices carrées est que, si $A \in \mathcal{M}_K(n)$ et $X \in K^n$, alors $A\,X \in K^n$, c'est-à-dire que X et $A\,X$ appartiennent au même espace vectoriel. Ces deux vecteurs peuvent donc éventuellement être égaux ou, ce qui est moins rare, colinéaires. Un vecteur X non nul de K^n est un **vecteur propre** de A, associé à la **valeur propre** $\lambda \in K$, si :

$$A\,X = \lambda X \quad \Longleftrightarrow \quad (A - \lambda\,I_n)\,X = O$$

Si la matrice $(A - \lambda\,I_n)$ est inversible, alors cette égalité n'est vérifiée que pour $X = O$. Or on cherche $X \neq O$, donc pour que λ soit valeur propre de A, il faut que :

$$\det(A - \lambda\,I_n) = 0$$

qui est appelée l'**équation caractéristique** de A.

L'expression $\det(A - \lambda\,I_n)$, qui est appelée le **polynôme caractéristique** de A, est un polynôme de degré n en λ. L'équation caractéristique a donc toujours n solutions dans $\mathbb{C}$, c'est-à-dire n valeurs propres complexes λ_1, λ_2, $\cdots$, λ_n. Plusieurs de ces valeurs propres peuvent être égales. **On admet** que, si ces n valeurs propres sont distinctes, alors il existe une base de $\mathbb{C}^n$ composée de vecteurs propres $(X_1, X_2, \cdots, X_n)$ associés aux n valeurs propres $(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Dans le cas contraire, on ne peut rien affirmer de tel.

6 Diagonalisation dans $\mathbb{C}$ d'une matrice

Soit A une matrice de $\mathcal{M}_{\mathbb{R}}(n)$ ou de $\mathcal{M}_{\mathbb{C}}(n)$. Cette matrice a toujours n valeurs propres complexes λ_1, λ_2, $\cdots$, λ_n. Si ces n valeurs propres sont distinctes, alors il existe alors une base de $\mathbb{C}^n$ de vecteurs propres $(X_1, X_2, \cdots, X_n)$ associés aux n valeurs propres $(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Soit P la matrice ayant ces n vecteurs propres comme colonnes. On sait que P est inversible. Montrons que le produit matriciel $B = A\,P$ est remarquable. L'élément de B situé sur la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne s'écrit :

$$b_{i,j} = (A\,X_j)_i = (\lambda_j\,X_j)_i = \lambda_j\,p_{i,j}$$

où $p_{i,j}$ désigne l'élément de P situé sur la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne. On introduit la matrice diagonale D suivante :

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Si l'on note $d_{i,j}$ son terme courant, alors on peut réécrire l'égalité précédente sous la forme :

$$b_{i,j} = \sum_{k=1}^{n} p_{i,k}\,d_{k,j}$$

ce qui veut dire que :

$$B = P\,D \quad \Longleftrightarrow \quad P^{-1}\,A\,P = D$$

La matrice A est donc diagonalisable dans $\mathbb{C}$. Ce résultat est faux dans le cas général, c'est-à-dire si les valeurs propres de A ne sont pas toutes distinctes, comme par exemple pour la matrice suivante :

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

On déduit de ce résultat que :

$$\det A = \prod_{i=1}^{n} \lambda_i$$
$$\mathrm{tr} A = \sum_{i=1}^{n} \lambda_i$$

Ces deux identités, qui sont vraies même lorsque les valeurs propres ne sont pas toutes distinctes, permettent de faire des vérifications sur le calcul des valeurs propres.

7 Diagonalisation dans $\mathbb{R}$ d'une matrice symétrique réelle

On appelle **matrice symétrique réelle** toute matrice carrée $A \in \mathcal{M}_{\mathbb{R}}(n)$ telle que $A^\top = A$. Le résultat du paragraphe précédent se généralise aux matrices symétriques réelles. Nous allons le démontrer uniquement pour les matrices symétriques réelles admettant n valeurs propres distinctes. Soit A une matrice symétrique réelle, et soit λ_1, λ_2, $\cdots$, λ_n ses n valeurs propres, qui sont *a priori* complexes. Si ces valeurs propres sont distinctes, alors il existe une base de $\mathbb{C}^n$ composée de vecteurs propres $(X_1, X_2, \cdots, X_n)$ de A. Pour $i \in [1, n]$:

$$A\, X_i = \lambda_i\, X_i \quad \Longrightarrow \quad X_i^\top\, A^\top = \lambda_i\, X_i^\top$$

Puisque, par hypothèse, A est symétrique :

$$X_i^\top\, A = \lambda_i\, X_i^\top$$

d'où :

$$X_i^\top\, A\, \overline{X_i} = \lambda_i\, X_i^\top\, \overline{X_i} \tag{3}$$

Par ailleurs :

$$A\, X_i = \lambda_i\, X_i \Longrightarrow \overline{A}\, \overline{X_i} = \overline{\lambda_i}\, \overline{X_i}$$

Puisque, par hypothèse, A est réelle :

$$A\, \overline{X_i} = \overline{\lambda_i}\, \overline{X_i}$$

d'où :

$$X_i^\top\, A\, \overline{X_i} = \overline{\lambda_i}\, X_i^\top\, \overline{X_i} \tag{4}$$

De (3) et (4), on déduit que :

$$\overline{\lambda_i} = \lambda_i$$

ce qui signifie que toutes les valeurs propres de A sont réelles, et donc que A est diagonalisable dans $\mathbb{R}$.

Par ailleurs, **on admet** que, pour toute matrice symétrique réelle, il existe une bon de $\mathbb{R}^n$ composée de vecteurs propres.

8 Matrices carrées réelles (semi-)définies positives

Une matrice carrée réelle A est **définie positive** si :

$$\forall X \in \mathbb{R}^n \backslash \{O\}, X^\top\, A\, X > 0$$

Cette définition n'aurait pas de sens dans $\mathbb{C}$. Une matrice carrée réelle A est **semi-définie positive** si :

$$\forall X \in \mathbb{R}^n \backslash \{O\}, X^\top\, A\, X \geq 0$$

Montrons qu'une matrice symétrique réelle, définie positive, admet n valeurs propres strictement positives. En effet, il existe une bon de $\mathbb{R}^n$ constituée de vecteurs propres de A. La matrice P dont les colonnes sont les vecteurs de cette bon vérifie l'égalité suivante, où D est la matrice diagonale déjà définie plus haut :

$$A = P\, D\, P^{-1} \quad \Longrightarrow \quad A = P\, D\, P^\top$$

En notant $(X_1, \ldots, X_n)$ les vecteurs propres de A associés aux valeurs propres $(\lambda_1, \ldots, \lambda_n)$, il vient :

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

Or le produit matriciel $P^\top X_i$ s'écrit très simplement, puisque la base $(X_1, \ldots, X_n)$ est orthonormée :

$$P^\top X_i = \begin{bmatrix} X_1^\top \\ \vdots \\ X_i^\top \\ \vdots \\ X_n^\top \end{bmatrix} X_i = \begin{bmatrix} X_1^\top X_i \\ \vdots \\ X_i^\top X_i \\ \vdots \\ X_n^\top X_i \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Par conséquent :

$$X_i^\top A\, X_i = \lambda_i$$

Par ailleurs, on sait que $X_i^\top A\, X_i > 0$, puisque A est définie positive. On en déduit donc que $\lambda_i > 0$.

Si A est seulement semi-définie positive, on montre par un raisonnement similaire que ses n valeurs propres sont positives ou nulles.