Les matrices - Multiplication

Notes rédigées par Laurent ZIMMERMANN

Résumé L'objectif de cette séquence est de généraliser la règle du produit matriciel au cas de matrices d'ordre $(m \times n)$ quelconque. La règle est formulée sur base de la notation indicielle compacte des matrices. La condition à respecter sur les dimensions des matrices est établie.

Rappel

Dans l'introduction aux matrices nous avons écrit un système de deux équations à deux inconnues en utilisant un produit de matrices. L'impératif de retrouver les équations de départ nous a guidé pour définir cette opération : *le produit de deux matrices est une nouvelle matrice dont chaque élément est calculé comme le produit scalaire entre une ligne de la première matrice et une colonne de la seconde matrice, la ligne et la colonne à considérer étant celles de l'élément de la nouvelle matrice que l'on calcule*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$$

Jusqu'ici nous avons envisagé le produit d'une matrice 2×2 avec une matrice 2×1. Nous nous proposons à présent de généraliser la règle à des matrices de dimensions quelconques.

Produit de matrices 2 x 2

Dans la vidéo d'introduction aux matrices nous avons écrit un système de deux équations linéaires à deux inconnues sous la forme symbolique

$$\mathbf{AX} = \mathbf{P}$$

et nous avons montré, en établissant l'expression de $\mathbf{A}^{-1}$, que sa solution pouvait s'écrire sous la forme

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{P}$$

Pour parvenir à cette dernière expression, nous avons utilisé une division par une matrice. Nous aurions aussi pu y parvenir en multipliant les deux membres de l'équation de départ par $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{P}$$

et en nous disant que $\mathbf{A}^{-1}\mathbf{A}$ devrait se simplifier. Mais ce produit est celui de deux matrices 2×2, une opération que nous n'avons pas encore examinée.

Envisageons donc le produit entre deux matrices 2×2 et une matrice colonne. Commençons par effectuer le second puisque nous connaissons déjà la règle à appliquer [1]. Nous obtenons successivement.

$$\begin{aligned}
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &\overset{\overset{1}{\downarrow}}{=} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} ex+fy \\ gx+hy \end{pmatrix} \\
&\overset{\overset{2}{\downarrow}}{=} \begin{pmatrix} a(ex+fy)+b(gx+hy) \\ c(ex+fy)+d(gx+hy) \end{pmatrix} \\
&\overset{\overset{3}{\downarrow}}{=} \begin{pmatrix} (ae+bg)x+(af+bh)y \\ (ce+dg)x+(cf+dh)y \end{pmatrix} \\
&\overset{\overset{4}{\downarrow}}{=} \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

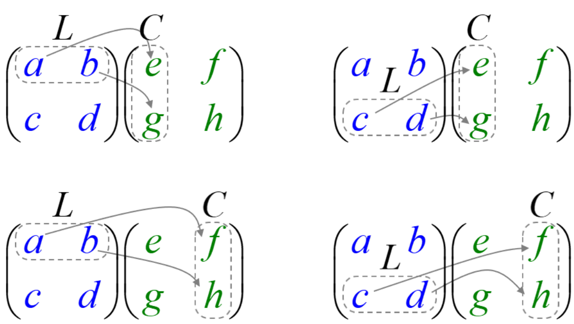
Les justifications des étapes du calcul sont les suivantes :

1. application de la loi de multiplication entre une matrice 2×2 et un vecteur colonne 2×1 (effectuer le produit)
2. idem
3. application de la loi de distributivité de la multiplication sur l'addition pour les scalaires (distribuer d'abord, mettre en évidence *in fine*) et de la loi de commutativité de l'addition des scalaires (réarranger les termes)
4. application de la loi de multiplication entre une matrice 2×2 et un vecteur colonne 2×1 (décomposer en un produit)

En comparant la première et la dernière étape, et compte tenu que le même vecteur colonne figure à droite dans les deux membres, nous déduisons que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

Nous constatons que le produit d'une matrice 2×2 par une autre matrice 2×2 est une nouvelle matrice 2×2. Chaque élément de la nouvelle matrice est le produit scalaire entre une ligne de la première matrice et une colonne de la seconde, la ligne et la colonne à choisir étant celles de l'élément de la nouvelle matrice que l'on calcule. La règle du produit scalaire est la même que celle vue dans l'introduction, sauf qu'il faut la répéter pour un plus grand nombre d'éléments à calculer (une colonne en plus).



1. Nous verrons dans une vidéo ultérieure que la multiplication des matrices est associative et nous aurions donc tout aussi bien pu commencer par le premier produit ; cela n'aurait en rien changé le résultat final.

Utilisation du symbole sommatoire

Par conséquent, si une matrice **C** est le produit des matrices **A** et **B**, nous pouvons écrire, en utilisant la notation standard :

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Écrivons les unes en dessous des autres les égalités entre éléments correspondants :

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} \\ c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} \\ c_{22} &= a_{21}b_{12} + a_{22}b_{22} \end{aligned}$$

Nous pouvons noter deux points communs :

— les premiers indices des c sont 1, 1, 2, 2 et se retrouvent toujours et seulement comme premiers indices des a ;

— les seconds indices des c sont 1, 2, 1, 2 et ne se retrouvent toujours et seulement comme seconds indices des b.

Cela signifie que le calcul de n'importe quel élément c se fait en utilisant toujours

— les éléments a pris à la même ligne que cet élément c ;

— les éléments b pris à la même colonne que cet élément c.

ce qui fait dire que le produit matriciel s'effectue « ligne par colonne ». Ainsi, quels que soient les indices i et j d'un élément c, nous avons toujours

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j}$$

Il s'agit ici d'une somme simple, puisqu'elle ne compte que deux termes. Nous pouvons en profiter pour la noter avec le symbole sommatoire Σ (la lettre grecque sigma majuscule qui est l'équivalent de notre lettre S comme dans le mot somme). Pour comprendre cette notation, nous devons d'abord avoir remarqué dans la formule précédente que les deux termes sont identiques à ceci près qu'un des indices (qui apparaît d'ailleurs deux fois), prend d'abord la valeur 1 dans le premier terme $(a_{i\mathbf{1}}b_{\mathbf{1}j})$, et ensuite la valeur 2 dans le second terme $(a_{i\mathbf{2}}b_{\mathbf{2}j})$.

La notation qui suit

— désigne cet indice changeant par le symbole k,

— fait donc apparaître le produit $a_{ik}b_{kj}$,

— indique que l'indice k prend toutes les valeurs comprises entre 1 et 2,

— rappelle enfin que tous ces produits construits avec les valeurs successives de k doivent être additionnées (somme Σ) :

$$c_{ij} = \sum_{k=1}^{2} a_{ik}b_{kj}$$

ce qui s'exprime de manière raccourcie par l'expression

« c_{ij} est la somme pour k allant de 1 à 2 des a_{ik} b_{kj}. »

Généralisation aux matrices carrées plus grandes

Ce résultat se généralise à des matrices carrées de plus grandes dimension, $n \times n$. Au lieu que chaque produit scalaire d'une ligne par une colonne compte 2 termes comme c'est le cas dans le produit de deux matrices 2×2, il comptera n termes dans le cas du produit de deux matrices $n \times n$.

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \vdots + a_{in}b_{nj}$$

$$\underset{n\times n}{A}\,\underset{n\times n}{B} = C = (c_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & \dots & a_{in} \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ b_{21} & \dots & b_{2j} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{nn} \end{pmatrix} = (c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj})$$

Généralisation aux matrices de dimensions quelconques

Ce résultat se généralise même à des matrices quelconques (rectangulaires) pour autant que le nombre de colonnes de la première (= nombre d'éléments d'un vecteur ligne) soit identique au nombre de lignes de la seconde (= nombre d'éléments d'un vecteur colonne), car si les nombres d'éléments des deux vecteurs sont différents il est impossible d'en effectuer le produit scalaire.

Attention ! Repère bien les mots « ligne » et « colonne » dans le paragraphe qui précède et sois bien conscient les deux points suivants.

- Vu que les éléments successifs d'une *ligne* dans une matrice sont écrits dans des *colonnes* adjacentes, le nombre d'éléments dans une *ligne* est forcément égal au nombre de *colonnes* de cette matrice.
- Vu que les éléments successifs d'une *colonne* dans une matrice sont écrits dans des *lignes* adjacentes, le nombre d'éléments dans une *colonne* est forcément égal au nombre de *lignes* de cette matrice.

$$\underset{l\times n}{A}\,\underset{n\times m}{B} = \underset{l\times m}{C} = (c_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & \dots & a_{in} \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{l1} & a_{l2} & \dots & \dots & a_{ln} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1m} \\ b_{21} & \dots & b_{2j} & \dots & b_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{nm} \end{pmatrix} = (c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj})$$

Ainsi le produit d'une matrice $l \times n$ et d'une matrice $n \times m$ sera une matrice $l \times m$ ayant le même nombre de lignes que la première et le même nombre de colonnes que la seconde

$$\underbrace{\mathbf{A}}_{l\times n}\underbrace{\mathbf{B}}_{n\times m} = \underbrace{\mathbf{C}}_{l\times m}$$

et chaque élément de **C** sera le produit scalaire de la i^{e} ligne de **A** avec la j^{e} colonne de **B** :

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

```
// Pour chaque ligne de la matrice produit...
for (i=0;i<l;i++)
    // et pour chaque colonne de la matrice produit...
    for (j=0;j<m;j++)
        // le produit scalaire ligne * colonne s'effectue...
        for (k=0;k<n;k++)
            // en cumulant les produits des éléments homologues
            c[i][j]+=a[i][k]*b[k][j];
```

L'essentiel

- Le produit de deux matrices n'est possible que si le nombre de colonnes de la première est égal au nombre de lignes de la seconde.
- Le produit de deux matrices est une nouvelle matrice qui possède le même nombre de lignes que la première et le même nombre de colonnes que la seconde.

$$(l \times n) \cdot (n \times m) \quad \rightarrow \quad (l \times m)$$

- Chaque élément de la matrice produit est égal au produit scalaire entre une ligne de la première matrice et une colonne de la seconde matrice, la ligne et la colonne à prendre en compté étant celles de l'élément en question.
- De cette manière de calculer les éléments de la matrice produit découle l'expression de *produit ligne par colonne* pour décrire le produit de deux matrices.
- Le produit scalaire d'une ligne de la première matrice avec une colonne de la seconde matrice n'est possible que si cette ligne et cette colonne possèdent le même nombre d'éléments ; il en résulte la condition énoncée en premier lieu ci-dessus.