

Matrices reloaded : inversion de matrices

ECE3 Lycée Carnot

23 mars 2011

Dans notre premier chapitre consacré aux matrices, nous avons notamment étudié assez longuement une opération à priori assez élémentaire : le produit. Nous avons ainsi constaté que celui-ci ne possédait pas vraiment toutes les propriétés auxquelles on serait en droit de s'attendre, en particulier il n'est pas commutatif. Un autre point qu'on soulève souvent quand on étudie une opération sur un ensemble (on reviendra sur ce sujet dans les derniers chapitres de l'année, consacrés aux espaces vectoriels), est de savoir si on peut effectuer l'opération inverse de l'opération en question. Pour la somme de matrices, par exemple, pas de problème, on peut faire des soustractions, car toute matrice admet une matrice opposée. Pour le produit, ça marche encore une fois nettement moins bien : il n'existe pas de division sur les matrices, et cela est du en partie à l'absence de commutativité, mais aussi au fait que la notion d'inverse fonctionne moins bien sur les matrices que sur les réels. Pour ces derniers, seul le nombre 0 n'est pas inversible, et c'est la cause de l'impossibilité de la division par 0. Pour les matrices, la matrice nulle est loin d'être la seule à poser problème.

1 Inversion de matrices

Définition 1. Une matrice carrée $M \in \mathcal{M}_n(\mathbb{R})$ est **inversible** s'il existe une matrice $N \in \mathcal{M}_n(\mathbb{R})$ telle que $MN = NM = I_n$. La matrice N est alors notée M^{-1} et on l'appelle **matrice inverse** de la matrice M .

Remarque 1. La notion n'a pas de sens dans le cas des matrices qui ne sont pas carrées.

Remarque 2. L'inverse d'une matrice, quand il existe, est unique. En effet, supposons qu'il existe deux matrices N et N' vérifiant $MN = NM = I$ et $MN' = N'M = I$. On a alors $NMN' = (NM)N' = N'$ d'une part et $NMN' = N(MN') = N$ d'autre part, donc $N = N'$.

Exemple : L'inverse de la matrice I_n est bien sûr I_n elle-même. La matrice nulle n'est pas inversible.

Exemple : La matrice $A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ est inversible et a pour inverse $B = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$.

Exemple : La matrice $M = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix}$ n'est pas nulle, mais elle n'est pas inversible pour

autant : on peut la multiplier à droite par ce qu'on veut, la première ligne du résultat sera toujours constituée de 3 zéros, et la matrice produit ne peut donc pas être égale à I_3 .

Proposition 1. Une matrice diagonale est inversible si et seulement si tous ses coefficients diagonaux

sont non nuls. On a alors, si $A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$, $A^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn}^{-1} \end{pmatrix}$.

Remarque 3. Nous verrons un peu plus loin que cette caractérisation s'étend en fait aux matrices triangulaires.

Proposition 2. Principales propriétés calculatoires de l'inversion de matrices.

- Si M est inversible, alors M^{-1} aussi et $(M^{-1})^{-1} = M$.
- Si $M, N \in \mathcal{M}_n(\mathbb{R})^2$ sont deux matrices inversibles, le produit MN est inversible et $(MN)^{-1} = N^{-1}M^{-1}$.
- Si M est une matrice inversible, M^k est inversible pour tout entier $k \in \mathbb{N}$, et $(M^k)^{-1} = (M^{-1})^k$.
- Si $M, N \in \mathcal{M}_n(\mathbb{R})$ vérifient $MN = I$, alors M et N sont inversibles et inverses l'une de l'autre.

Démonstration. La première proposition est évidente : par définition, $MM^{-1} = M^{-1}M = I$, donc M est l'inverse de M^{-1} . La deuxième ne pose pas vraiment de problème non plus : on a $(N^{-1}M^{-1})(MN) = N^{-1}(M^{-1}M)N = N^{-1}IN = N^{-1}N = I$, et de même $(MN)(N^{-1}M^{-1}) = I$, donc $N^{-1}M^{-1}$ est bien l'inverse de MN . La troisième propriété ne pose guère de problème non plus, il suffit de vérifier que $M^k(M^{-1})^k = (M^{-1})^kM^k = I$, ce qui est essentiellement immédiat (on peut utiliser que M et M^{-1} commutent pour éviter de recourir à une récurrence). La dernière proposition dit qu'on n'a en fait pas besoin de vérifier que le produit à gauche et à droite est égal à I pour trouver l'inverse d'une matrice, l'un des deux est suffisant. On ne démontrera pas ce résultat beaucoup moins facile qu'il n'en a l'air. \square

Remarque 4. Un des principaux intérêts de travailler avec des matrices inversibles est qu'on peut simplifier un peu plus naturellement certains calculs, notamment : si M est une matrice inversible et $MA = MB$, alors $A = B$ (il suffit de multiplier l'égalité à gauche par M^{-1} pour obtenir le résultat). Autre remarque utile : si A et B sont deux matrices non nulles telles que $AB = 0$, alors aucune des deux matrices n'est inversible (sinon, par l'absurde, en multipliant à gauche par l'inverse de A ou à droite par l'inverse de B , on constaterait que l'autre matrice est nulle).

Exemple : Le calcul d'inverse de matrices peut faire intervenir de petites astuces comme celle-ci : soit

$$M = \begin{pmatrix} -3 & 4 & 2 \\ -2 & 3 & 1 \\ 2 & -2 & 0 \end{pmatrix}. \text{ Un petit calcul permet d'obtenir } M^2 = \begin{pmatrix} 5 & -4 & -2 \\ 2 & -1 & -1 \\ -2 & 2 & 2 \end{pmatrix} \text{ et de constater}$$

que $M^2 = 2I - M$, ce qu'on peut écrire $M + M^2 = 2I$, ou encore $\frac{1}{2}M(M+I) = I$. Ceci suffit à montrer

$$\text{que } M \text{ est inversible et que son inverse est } \frac{1}{2}(M+I). \text{ Autrement dit, } M^{-1} = \begin{pmatrix} -1 & 2 & 1 \\ -1 & 2 & \frac{1}{2} \\ 1 & -1 & \frac{1}{2} \end{pmatrix}.$$

2 Lien entre matrices et systèmes linéaires

Venons-en à un résultat annoncé depuis un certain temps, l'équivalence entre systèmes linéaires et équations matricielles. C'est en fait tout simple :

Théorème 1. Soit (S) le système linéaire

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p = b_n \end{cases}$$

En posant $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$; $X = (x_i)_{1 \leq i \leq n}$ et $B = (b_i)_{1 \leq i \leq n}$, le système (S) est équivalent à l'équation matricielle $AX = B$.

Démonstration. Le produit des matrices A et X est une matrice-colonne à n termes, le i -ème terme valant $\sum_{j=1}^p a_{ij}x_j$. En identifiant coefficient par coefficient, on a donc $AX = B \Leftrightarrow \forall 1 \leq i \leq n$,

$$\sum_{j=1}^p a_{ij}x_j = b_j, \text{ ce qui est exactement le système } (S). \quad \square$$

Exemple : Soit (S) le système
$$\begin{cases} 2x - 3y + z = 5 \\ -5x + 4y - z = -1 \\ x + y - 3z = -6 \end{cases}$$

Il est équivalent à l'équation matricielle
$$\begin{pmatrix} 2 & -3 & 1 \\ -5 & 4 & -1 \\ 1 & 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ -6 \end{pmatrix}.$$

Théorème 2. Le système (S) est un système de Cramer si et seulement si la matrice associée est une matrice inversible.

Démonstration. Il y a un sens très facile : si A est inversible, l'équation $AX = B$ est équivalente à $X = A^{-1}B$, ce qui montre que le système (S) a une unique solution. Dans l'autre sens, c'est malheureusement plus difficile, nous nous contenterons de donner un algorithme permettant de calculer l'inverse de A dans ce cas au paragraphe suivant. \square

Ce théorème nous donne une grande motivation à l'introduction de la notion d'inverse de matrice, mais ne nous dit toujours pas comment le calculer. Remarquons tout de même le cas particulier suivant :

Proposition 3. Un système triangulaire a pour matrice associée une matrice triangulaire supérieure. Le système est de Cramer (et donc la matrice triangulaire inversible) si et seulement si tous ses coefficients diagonaux sont non nuls.

3 Pivot de Gauss sur les matrices

Nous allons enfin donner une méthode systématique d'inversion des matrices, qui est en fait l'équivalent matriciel de la résolution des systèmes linéaires par le pivot de Gauss. Commençons par préciser qu'on ne parlera plus dans ce paragraphe que de matrices carrées, les autres n'étant de toute façon pas inversibles. Comme dans le cas des systèmes, le but est tout d'abord de se ramener à un système triangulaire, donc ici à une matrice triangulaire supérieure. Dans le cas où on a un système de Cramer (c'est-à-dire une matrice triangulaire n'ayant pas de 0 sur la diagonale), la matrice sera inversible, et il ne restera qu'à compléter la résolution du système en « remontant » le triangle. Dans le cas contraire, la matrice n'est pas inversible.

Toute l'idée est en fait de faire une correspondance entre les opérations élémentaires sur les lignes et colonnes que l'on effectue pour résoudre les systèmes linéaires et les produits par certaines matrices particulières. Lorsqu'on résout un système linéaire, si on note $(L_i)_{1 \leq i \leq n}$ les lignes du système (on suppose dans tout ce paragraphe que le système, et donc la matrice associée, est carré), on effectue les opérations suivantes : échange de lignes $L_i \leftrightarrow L_j$; produit d'une ligne par un réel non nul $L_i \leftarrow \alpha L_i$; combinaison linéaire de deux lignes $L_i \leftarrow L_i + \alpha L_j$. La correspondante est sur la page suivante car elle ne tenait pas sur le reste de celle-ci. Le principe est ensuite simple :

- On effectue sur la matrice A les opérations du pivot de Gauss (comme si on résolvait le système linéaire associé) jusqu'à obtenir une matrice triangulaire supérieure. Ces opérations correspondent à des produits $B_k \times \cdots \times B_1$ par des matrices inversibles (k étant le nombre d'étapes nécessaires pour atteindre une matrice triangulaire), produits que l'on effectue en parallèle en partant de la matrice I .
- Si la matrice triangulaire obtenue a un coefficient diagonal nul, elle n'est pas inversible, et la matrice A non plus.
- Si tous les coefficients diagonaux sont non nuls, le système associé est résoluble en « remontant » le triangle. On fait de même avec notre matrice : on la rend diagonale en commençant par annuler les termes non diagonaux sur la dernière colonne. Ces nouvelles opérations correspondent à de nouveaux produits, et on finit par transformer A en I via un produit de matrices inversibles $B_{k'} \times \cdots \times B_1$. Ce produit n'est autre que l'inverse de la matrice A , qu'on a donc sous les yeux si on a pris soin de l'effectuer en parallèle à partir de la matrice I .

Deux exemples sont donnés après le gros tableau.

Opération sur les lignes du système	Produit de la matrice A par :
Échange $L_i \leftrightarrow L_j$	$ \begin{matrix} (L_i) \\ \\ (L_j) \end{matrix} \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & & & & & \vdots \\ \vdots & \ddots & 1 & & & & & & & & \vdots \\ \vdots & & & 0 & \dots & \dots & \dots & 1 & & & \vdots \\ \vdots & & & \vdots & 1 & & & \vdots & & & \vdots \\ \vdots & & & \vdots & & \ddots & & \vdots & & & \vdots \\ \vdots & & & \vdots & & & & 1 & \vdots & & \vdots \\ (L_j) \vdots & & & \vdots & & & & & & 1 & \vdots \\ \vdots & & & 1 & \dots & \dots & \dots & 0 & & & \vdots \\ \vdots & & & & & & & & & 1 & \ddots & \vdots \\ \vdots & & & & & & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} $
Produit par un réel $L_i \leftarrow \alpha L_i$	$ (L_i) \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & & \vdots \\ \vdots & & \ddots & \alpha & \ddots & & & \vdots \\ \vdots & & & & \ddots & 1 & \vdots & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} $
Combinaison linéaire $L_i \leftarrow L_i + \alpha L_j$	$ \begin{matrix} (L_i) \\ (L_j) \end{matrix} \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & 1 & \dots & \alpha & & & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & & \vdots \\ (L_j) \vdots & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} $

Nous allons calculer l'inverse de la matrice suivante en utilisant le pivot de Gauss : à gauche, les opérations sur la matrice A , à droite les mêmes opérations à partir de I pour obtenir l'inverse.

$$\begin{array}{ccc}
 A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -1 \\ -2 & -5 & 3 \end{pmatrix} & I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \\
 \\
 \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} & \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 + 2L_1 \end{array} \\
 \\
 \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} & L_2 \leftrightarrow L_3 \\
 \\
 \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} -1 & 1 & 0 \\ 4 & -1 & 1 \\ -2 & 1 & 0 \end{pmatrix} & \begin{array}{l} L_1 \leftarrow L_1 + L_3 \\ L_2 \leftarrow L_2 - L_3 \end{array} \\
 \\
 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 7 & -1 & 2 \\ 4 & -1 & 1 \\ -2 & 1 & 0 \end{pmatrix} & L_1 \leftarrow L_1 + 2L_2 \\
 \\
 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 7 & -1 & 2 \\ -4 & 1 & -1 \\ -2 & 1 & 0 \end{pmatrix} & L_2 \leftarrow -L_2
 \end{array}$$

Conclusion de ce long calcul : $A^{-1} = \begin{pmatrix} 7 & -1 & 2 \\ -4 & 1 & -1 \\ -2 & 1 & 0 \end{pmatrix}$.

4 Diagonalisation de matrices

Une constatation pas si bête qu'elle n'en a l'air quand on travaille avec des matrices, c'est qu'on s'en sort toujours plus facilement avec des matrices diagonales (pour les calculs de puissances notamment). Le principe de la diagonalisation est très simple, il s'agit de « transformer » une matrice A en matrice diagonale.

Définition 2. Une matrice A est dite diagonalisable s'il existe une matrice inversible P telle que le produit $P^{-1}AP$ soit une matrice diagonale.

Cette propriété permet effectivement de calculer très facilement les puissances de la matrice A , car si on a $D = P^{-1}AP$, alors pour tout entier n , $D^n = P^{-1}A^nP$ (ce qui se prouve à l'aide d'une petite récurrence). Malheureusement, l'étude des conditions assurant la diagonalisabilité d'une matrice est loin d'être simple, et c'est un sujet que nous n'aborderons pas cette année. Nous nous contenterons d'un exemple montrant l'utilité de la notion.

Exemple : Considérons les matrices $A = \begin{pmatrix} 10 & -6 & -3 \\ -4 & 12 & 2 \\ -4 & -4 & 6 \end{pmatrix}$ et $P = \begin{pmatrix} \frac{3}{4} & 0 & -\frac{3}{2} \\ 0 & -\frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & 1 & 0 \end{pmatrix}$. On montre en utilisant le pivot de Gauss que P est inversible et que $P^{-1} = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -1 & -1 & \frac{1}{2} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$, puis par le calcul

que $P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 16 \end{pmatrix}$, dont on déduit $A^n = P^{-1} \begin{pmatrix} 4^n & 0 & 0 \\ 0 & 8^n & 0 \\ 0 & 0 & 16^n \end{pmatrix} P$ (qu'on peut chercher à expliciter ou non selon son courage).

Exemple 2 : Le retour des chaînes de Markov.

Doudou le hamster partage sa passionnante existence entre trois activités : manger, dormir, et faire de l'exercice sur sa roue. En l'observant toutes les minutes pendant un moment, on constate les choses suivantes :

- si Doudou fait de l'exercice à une certaine minute, il y aura une chance sur deux qu'il dorme et une chance sur deux qu'il mange à la minute suivante.
- si Doudou mange à une certaine minute, il y a une chance sur deux qu'il retourne dormir et une chance sur deux qu'il tente de faire de l'exercice à la minute suivante.
- enfin, s'il dort, une chance sur quatre (seulement) qu'il parte faire de l'exercice, une chance sur quatre qu'il mange, et une chance sur deux qu'il continue sa sieste.

Au moment du début de l'observation, Doudou est en train de manger et, vous l'aurez sûrement deviné, le but est d'étudier son comportement après n minutes et notamment rechercher des limites éventuelles aux probabilités de chaque activité. Notons donc A_n : « Doudou dort après n minutes », B_n : « Doudou mange après n minutes » et C_n : « Doudou fait de l'exercice après n minutes », ainsi que a_n , b_n et c_n les probabilités correspondantes. Une application (ou plutôt trois) de la formule des probabilités totales au système complet d'évènements formé par A_n , B_n et C_n permet d'obtenir le système de relations suivant :

$$\begin{cases} a_{n+1} &= \frac{1}{2}a_n + \frac{1}{2}b_n + \frac{1}{2}c_n \\ b_{n+1} &= \frac{1}{4}a_n + \frac{1}{2}b_n + \frac{1}{2}c_n \\ c_{n+1} &= \frac{1}{4}a_n + \frac{1}{2}b_n \end{cases}$$

Nous allons plutôt exploiter la forme matricielle : en notant $X_n = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$ et $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix}$, le

système précédent s'écrit $X_{n+1} = AX_n$. Par une petite récurrence désormais classique, on en déduit que $X_n = A^n X_0$, ne reste plus que le délicat problème du calcul de A^n . On peut en fait constater

qu'en posant $P = \begin{pmatrix} 1 & 0 & -1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, $P^{-1}AP$ est une matrice diagonale D (de coefficients diagonaux

$1, -\frac{1}{2}$ et 0). Naturellement, à notre niveau, nous n'avons pas vraiment de moyen de deviner à quoi va ressembler la matrice P sans connaître D (ou vice-versa), l'une des deux sera donc toujours donnée dans l'énoncé. Une fois cette matrice diagonale obtenue, une petite récurrence permet de prouver que $A^n = PD^nP^{-1}$ (en effet, c'est vrai au rang 1 d'après les calculs précédents, et si on le suppose vrai pour A^n , alors $A^{n+1} = A \times A^n = (PDP^{-1})(PD^nP^{-1}) = PD^{n+1}P^{-1}$). On en déduit que

$$A^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} + \frac{(-1)^n}{2^{n+1}} & \frac{1}{4} + \frac{(-1)^{n+1}}{2^{n+1}} \\ \frac{1}{4} & \frac{1}{4} + \frac{(-1)^{n+1}}{2^{n+1}} & \frac{1}{4} + \frac{(-1)^n}{2^{n+1}} \end{pmatrix}, \text{ donc } X_n = A^n X_0 = A^n \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} + \frac{(-1)^n}{2^{n+1}} \\ \frac{1}{4} + \frac{(-1)^{n+1}}{2^{n+1}} \end{pmatrix}.$$

En particulier, les suites (a_n) , (b_n) et (c_n) ont pour limites respectives $\frac{1}{2}$, $\frac{1}{4}$ et $\frac{1}{4}$.