Méthode de Gauss pour les inversions de matrices

1 Rappel de la méthode :

On cherche à inverser une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$. La méthode de Gauss consiste à écrire dans un tableau la matrice A à gauche et la matrice identité de taille n I_n à droite :

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & \ddots & & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right)$$

et à effectuer des *opérations élémentaires* sur les lignes de A et I_n simultanément. Ceci va modifier progressivement les matrices de part et d'autre de la ligne de séparation, et on continue jusqu'à obtenir I_n du côté gauche :

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 & b_{21} & \ddots & & b_{2n} \\ \vdots & & & & \vdots & & & \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right)$$

La matrice $B = (b_{ij})$ qui se trouve à droite est alors l'inverse de A : $B = A^{-1}$.

Deux questions se posent alors :

1. Quelles sont les opérations élémentaires ? Comment les utiliser pour "transformer" la matrice A en matrice identité ?
2. Quelle est cette magie noire ? Autrement dit, pourquoi ça marche ?

2 La mise en oeuvre

Le principe est le même que celui du pivot de Gauss pour les systèmes (voir la fiche dédiée à ce sujet). Appelons $L_1, \dots, L_n$ les lignes de notre matrice. Les opérations élémentaires sur les lignes sont :

1. $L_i \leftarrow \alpha L_i$, pour $\alpha \neq 0$: on multiplie une ligne par un scalaire *non nul* α ;
2. $L_i \leftrightarrow L_j$: on échange deux lignes ;
3. $L_i \leftarrow L_i + \lambda L_j$, $\lambda \in \mathbb{K}$: on ajoute à la i-ème ligne un multiple de la j-ième.

Dans un premier temps, on va utiliser ces opérations sur A, d'une part et I_n, d'autre part, pour se ramener à une matrice triangulaire du côté gauche :

$$\left(\begin{array}{cccc|ccc} t_{11} & t_{12} & \dots & t_{1n} & * & \dots & * \\ 0 & t_{22} & \dots & t_{2n} & * & \dots & * \\ \vdots & & & & \vdots & & \\ 0 & 0 & \dots & t_{nn} & * & \dots & * \end{array}\right)$$

Pour arriver là, on commence par s'assurer que le coefficient a_{11} de A est non nul. S'il l'est, on échange avec une ligne dont le premier coefficient est non nul (on n'oublie pas de faire la même opération sur I_n du côté droit !) ; si toute la première colonne est nulle, alors A n'est pas inversible, et on peut aller dormir.

Une fois que le coefficient en haut à gauche est non nul, on l'utilise pour éliminer tous les coefficients de la première colonne qui sont en dessous, en faisant, pour chaque ligne $L_2, \dots, L_n$, l'opération $L_i \leftarrow L_i - \frac{a_{i1}}{a_{11}} L_1$ (et on fait les mêmes opérations sur le côté droit du tableau). On se retrouve, à gauche, avec une nouvelle matrice, dont le coefficient en haut à gauche est non nul, et le reste de la première colonne est nul.

On passe alors au deuxième coefficient de la deuxième ligne, et on applique le même procédé : quitte à échanger L_2 avec $L_3, \ldots L_n$ (et pareil à droite), on s'arrange pour que le coefficient de la 2ème ligne, 2ème colonne soit non nul, et on l'utilise pour éliminer les coefficients de la deuxième colonne sur $L_3, \ldots L_n$.

On se ramène ainsi à un système triangulaire, comme annoncé plus haut. La matrice n'est inversible que si on peut mener à bien ce procédé, et obtenir un système triangulaire dont les coefficients diagonaux sont non nuls : $t_{ii} \neq 0$. Toutes ces opérations ont bien sûr changé I_n en une autre matrice à droite.

On va maintenant utiliser les mêmes opérations élémentaires pour "remonter" les lignes de la matrice triangulaire et la transformer en matrice identité. On comment par diviser L_n par t_{nn} : la dernière ligne à gauche est maintenant $(0, \ldots, 0, 1)$. Puis on utilise cette ligne pour éliminer tous les coefficients de la dernière colonne sur $L_1, \ldots, L_{n-1}$ en faisant $L_i \leftarrow L_i - t_{i1} L_n$ (et pareil à droite).

L'avant dernière ligne est maintenant $L_{n-1} = (0, \ldots, t_{n-1,n-1}, 0)$: on la divise par $t_{n-1,n-1}$ et on s'en sert pour éliminer les coefficients de la $n-1$-ième colonne sur les lignes d'au dessus. Et ainsi de suite, en faisant les mêmes opérations à droite du tableau, jusqu'à obtenir

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ 0 & 1 & \ldots & 0 & b_{21} & \ddots & & b_{2n} \\ \vdots & & & & \vdots & & & \\ 0 & 0 & \ldots & 1 & b_{n1} & b_{n2} & \ldots & b_{nn} \end{array}\right)$$

Alors $B = A^{-1}$.

3 Pourquoi ?

En réalité, appliquer les opérations élémentaires revient à multiplier A par certaines matrices, appelées...matrices élémentaires.

Echange de lignes : Faire l'opération $L_i \leftrightarrow L_j$ revient à multiplier A à gauche par la matrice E_{ij} obtenue en échangeant les i-ème et j-ème lignes de la matrice identité I_n. Par exemple, pour $n = 4$,

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad E_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

La matrice E_{ij} est inversible, d'inverse elle-même : $E_{ij}^{-1} = E_{ij}$, et vérifie $E_{ij} = E_{ji}$. Et, comme on l'a dit, $E_{ij}A$ est la matrice obtenue en échangeant les lignes i et j de A

Preuve (optionnelle) : Le coefficient de E_{ij} en position (k, l) est donné par

$$(E_{ij})_{kl} = \begin{cases} 1 \text{ si } (k,l) = (i,j) \text{ ou } (k,l) = (j,i) \\ 1 \text{ si } k = l \text{ et } k \neq i, j \\ 0 \text{ sinon.} \end{cases}$$

Soit $C = E_{ij}A$, alors les coefficients de C sur la i-ème ligne sont

$$c_{il} = \sum_{p=1}^{n} (E_{ij})_{ip} a_{pl} = a_{jl}$$

car $(E_{ij})_{ip} = 1$ si $p = j$, 0 sinon. Les coefficients de la i-ème ligne de C sont donc ceux de la j-ème ligne de A. De même, les coefficients sur la j-ème ligne sont

$$c_{jl} = \sum_{p=1}^{n} (E_{ij})_{jp} a_{pl} = a_{il} :$$

ce sont ceux de la i-ème ligne de A. Pour les autres lignes, avec $k \neq i, j$,

$$c_{kl} = \sum_{p=1}^{n} (E_{ij})_{kp} a_{pl} = a_{kl}$$

autrement dit les autres lignes ne changent pas.

Exemple : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, alors $E_{12}A = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$.
Pour des matrices de taille 3, prenons une matrice quelconque

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \text{ alors } E_{23}A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

Multiplication d'une ligne par α : Faire l'opération $L_i \leftarrow \alpha L_i$ revient à multiplier A à gauche par la matrice $D_i(\alpha)$, obtenue en multipliant la i-ème de la matrice identité I_n par α. Par exemple, pour $n = 4$,

$$D_2(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La matrice $D_i(\alpha)$ est inversible, d'inverse $D_i(\alpha)^{-1} = D_i(\frac{1}{\alpha})$. Et, $D_i(\alpha)A$ est la matrice obtenue en multipliant la i-ème ligne de A par α.

Preuve (optionnelle) : Le coefficient de E_{ij} en position (k, l) est donné par

$$(D_i(\alpha))_{kl} = \begin{cases} 1 \text{ si } k = l \text{ et } k \neq i \\ \alpha \text{ si } (k, l) = (i, i) \\ 0 \text{ sinon.} \end{cases}$$

Soit $C = D_i(\alpha)A$, alors les coefficients de C sur la i-ème ligne sont

$$c_{il} = \sum_{p=1}^{n} (D_i(\alpha))_{ip} a_{pl} = \alpha a_{il}$$

ce sont donc ceux de la i-ème ligne de A, multipliés par α. Pour les autres lignes, avec $k \neq i$,

$$c_{kl} = \sum_{p=1}^{n} (D_i(\alpha))_{kp} a_{pl} = a_{kl}$$

autrement dit les autres lignes ne changent pas.

Ajout à une ligne de λ fois une autre : Faire l'opération $L_i \leftarrow L_i + \lambda L_j$ revient à multiplier A à gauche par la matrice $T_{ij}(\lambda)$, obtenue en ajoutant λ en position (i, j) à la matrice identité. Par exemple, pour $n = 4$,

$$T_{34}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \lambda \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad T_{13}(\lambda) = \begin{pmatrix} 1 & 0 & \lambda & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La matrice $T_{ij}(\lambda)$ est inversible, d'inverse $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$. Et $T_{ij}(\lambda)A$ est la matrice obtenue en ajoutant à la i-ème ligne de A λ fois la j-ème.

Preuve (optionnelle) : Le coefficient de $T_{ij}(\lambda)$ en position (k, l) est donné par

$$(T_{ij}(\lambda))_{kl} = \begin{cases} 1 \text{ si } k = l \\ \lambda \text{ si } (k, l) = (i, j) \\ 0 \text{ sinon.} \end{cases}$$

Soit $C = T_{ij}(\lambda)A$, alors les coefficients de C sur la i-ème ligne sont

$$c_{il} = \sum_{p=1}^{n} (T_{ij}(\lambda))_{ip} a_{pl} = \lambda a_{jl} + a_{ij}$$

ce sont donc ceux de la i-ème ligne de A, plus λ fois ceux de la j-ème. Pour les autres lignes, avec $k \neq i$,

$$c_{kl} = \sum_{p=1}^{n} (T_{ij}(\lambda))_{kp} a_{pl} = a_{kl}$$

autrement dit les autres lignes ne changent pas.

Exemple : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $T_{12}(\lambda) = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, alors $T_{12}(\lambda)A = \begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix}$.
Pour des matrices de taille 3 :

$$T_{12}(\lambda) = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \text{ alors } T_{12}(\lambda)A = \begin{pmatrix} a_1+\lambda a_2 & b_1+\lambda b_2 & c_1+\lambda c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Donc, pourquoi ça marche : Ainsi, en faisant successivement des opérations élémentaires sur A dans la partie gauche, on la multiplie par une série de matrices élémentaires jusqu'à obtenir I_n ; ainsi, on fait

$$T_{ij}(\lambda)E_{kl}\dots T_{i',j'}(\mu)A = I_n$$

Dans l'exemple donné en cours, avec $A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}$, vous pouvez vérifier qu'on a fait :

$$T_{12}(-2)D_2(-\frac{1}{8})T_{23}(5)T_{13}(-1)D_3(2)T_{32}(\frac{1}{2})T_{31}(1)T_{21}(-4)A$$

et ainsi obtenu I_3. Ce qui veut dire que A est inversible, et $T_{12}(-2)D_2(-\frac{1}{8})T_{23}(5)T_{13}(-1)D_3(2)T_{32}(\frac{1}{2})T_{31}(1)T_{21}(-4)$ est son inverse. Or, du côté droit, on a fait les mêmes opérations à I_n : on a donc fait

$$T_{12}(-2)D_2(-\frac{1}{8})T_{23}(5)T_{13}(-1)D_3(2)T_{32}(\frac{1}{2})T_{31}(1)T_{21}(-4)I_n = B.$$

Or, puisque multiplier par I_n ne change pas la matrice, on a donc bien que B est l'inverse de A.

Remarque : Les plus observateurs remarqueront qu'on a montré que $BA = I_n$, mais pas que $AB = I_n$. C'est vrai ; on verra plus tard dans le cours que, pour une matrice carrée, il suffit en fait d'avoir soit $BA = I_n$, soit $AB = I_n$, pour pouvoir affirmer que A est inversible et $A^{-1} = B$. A suivre...