CALCUL MATRICIEL

PLAN

I : Résolution de systèmes

1– Méthode du pivot de Gauss-Jordan

On appelle système linéaire à n équations et p inconnues un système de la forme :

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

En ce qui concerne les coefficients a_{ij}, la convention universelle adoptée est telle que l'indice de gauche i est le numéro de ligne, l'indice j de droite est le numéro de colonne correspondant à l'inconnue x_j.

On procède à la résolution d'un tel système par un algorithme appelé méthode du pivot de Gauss-Jordan. Cette méthode consiste à effectuer des opérations sur les lignes de façon à obtenir un système échelonné facile à résoudre, tout en garantissant que le système final est équivalent au système initial, de sorte que toute solution du premier système est solution du second et inversement.

Si tous les a_{i1} sont nuls, alors x_1 n'intervient pas et on passe à la deuxième colonne. Si l'un des a_{i1} est non nul, on permute la i-ème ligne et la première de façon à se ramener au cas $a_{11} \neq 0$. On remplace ensuite, pour $2 \leq i \leq n$, la ligne L_i par $L_i' = a_{11}L_i - a_{i1}L_1$, ce qui a pour effet d'annuler le coefficient de x_1 dans les lignes 2 à n. On obtient un nouveau système :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ \qquad a_{22}'\,x_2 + \ldots + a_{2p}'\,x_p = b_2' \\ \quad \ldots \\ \qquad a_{n2}'\,x_2 + \ldots + a_{np}'\,x_p = b_n' \end{cases}$$

et on itère le procédé sur les $n - 1$ dernières lignes. Ce système est équivalent au précédent. En effet, connaissant L_i' et L_1, il est possible de reconstituer $L_i = \dfrac{L_i' + a_{i1}L_1}{a_{11}}$. On remarquera pour cela qu'il est essentiel que a_{11} soit non nul.

Nous symboliserons le calcul $L_i' = a_{11}L_i - a_{i1}L_1$ par $L_i \leftarrow a_{11}L_i - a_{i1}L_1$, analogue à une opération d'affectation de variables en informatique, ceci afin de pouvoir indiquer les calculs effectués sans avoir besoin d'introduire de nouvelles notations pour les nouvelles lignes.

Un système est échelonné si sa matrice des coefficients vérifie les propriétés suivantes :

- Si une ligne est nulle, toutes les lignes qui suivent sont nulles
- Si la ligne i est non nulle, son premier coefficient a_{ij} non nul s'appelle son pivot et se trouve à une colonne j de rang strictement supérieur à celui du pivot de la ligne précédente.

Ainsi, la matrice $\begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 0 & 0 & 1 & 6 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée. Les inconnues en facteur d'un pivot sont appelées inconnues principales, les autres sont appelées inconnues secondaires ou paramètres. Avec la matrice précédente, les inconnues principales sont x_1, x_3, x_4 ; les inconnues secondaires sont x_2 et x_5.

EXEMPLE :

$$\begin{cases} 2x + 4y - z + t = 2 \\ x - 2y + 3z - t = 1 \\ x + y + \dfrac{z}{2} + \dfrac{t}{2} = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 2x + 4y - z + t = 2 \\ \quad 8y - 7z + 3t = 0 & \qquad L_2 \leftarrow L_1 - 2L_2 \\ \quad 2y - 2z = 2 & \qquad L_3 \leftarrow L_1 - 2L_3 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 2x + 4y - z + t = 2 \\ \quad 8y - 7z + 3t = 0 \\ \quad y - z = 1 & \qquad L_3 \leftarrow L_3/2 \end{cases}$$

Dans la dernière équation, on peut exprimer y en fonction de z et reporter dans l'équation précédente, mais pour des raisons didactiques, nous préférons appliquer la méthode de Gauss-Jordan de façon systématique jusqu'au bout.

$$\Leftrightarrow \quad \begin{cases} 2x + 4y - z + t = 2 \\ \quad 8y - 7z + 3t = 0 \\ \qquad z + 3t = -8 \end{cases} \qquad \mathrm{L}_3 \leftarrow \mathrm{L}_2 - 8\mathrm{L}_3$$

Le système final obtenu est échelonné. Les coefficients non nuls 2, 8, 1 en début de chaque équation sont les pivots. Les inconnues x, y, z devant ces pivots sont les inconnues principales. L'inconnue t est inconnue secondaire ou paramètre. Sous cette forme finale, on termine la résolution en exprimant de bas en haut les inconnues principales en fonction des paramètres, et en reportant dans les équations qui précèdent :

$$\Leftrightarrow \quad \begin{cases} 2x + 4y - z + t = 2 \\ \quad 8y - 7z + 3t = 0 \\ \qquad z = -8 - 3t \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 2x + 4y - z + t = 2 \\ \quad y = -7 - 3t \\ \quad z = -8 - 3t \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x = 11 + 4t \\ y = -7 - 3t \\ z = -8 - 3t \end{cases}$$

Ce système admet une infinité de solution.

Les opérations élémentaires effectuées sur les lignes sont de trois types :

(i)	$\mathrm{L}_i \leftrightarrow \mathrm{L}_j$	permuter deux lignes
(ii)	$\mathrm{L}_i \leftarrow \lambda \mathrm{L}_i$	multiplier par une constante non nulle λ
(iii)	$\mathrm{L}_i \leftarrow \mathrm{L}_i + \alpha \mathrm{L}_j$	ajouter à une ligne un multiple d'une autre

Ces opérations garantissent de conserver un système équivalent. C'est évident pour (i). Pour (ii), on retrouve la valeur initiale de la ligne en effectuant $\mathrm{L}_i \leftarrow \frac{1}{\lambda} \mathrm{L}_i$. Il est donc essentiel que λ soit non nul. Pour (iii), on retrouve la valeur initiale de L_i en effectuant $\mathrm{L}_i \leftarrow \mathrm{L}_i - \alpha \mathrm{L}_j$. α peut être nul, mais dans ce cas, aucune transformation n'est effectuée. Il est possible de combiner en même temps plusieurs de ces opérations afin de gagner du temps, ce qu'on a fait plus haut en effectuant $\mathrm{L}_i \leftarrow a_{11}\mathrm{L}_i - a_{i1}\mathrm{L}_1$. Plus généralement, on peut affecter à L_i une combinaison linéaire de toutes les lignes $\mathrm{L}_i \leftarrow \sum_{k=1}^{p} \lambda_k \mathrm{L}_k$ à condition que λ_i soit non nul, afin de garantir l'équivalence des systèmes.

2– Notation matricielle

On écrit le système initial sous la forme AX = B où A est le tableau constitué des coefficients a_{ij}, X la colonne des inconnues, élément de $\mathbb{K}^p$, et B la colonne constituée des termes du second membre, élément de $\mathbb{K}^n$:

$$\mathrm{A} = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \dots & & \dots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$$ A s'appelle matrice à n lignes et p colonnes.

$$\mathrm{X} = \begin{pmatrix} x_1 \\ \dots \\ x_p \end{pmatrix}$$

$$\mathrm{B} = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}$$

La notation AX désigne un véritable produit, celui d'une matrice par une colonne. Pour pouvoir effectuer ce produit, il est nécessaire que le nombre de termes dans la colonne X soit égal au nombre de colonnes de A. Si Y = AX, Y est une colonne ayant un nombre de termes égal au nombre de lignes de A, et, pour tout i :

$$y_i = \sum_{j=1}^{p} a_{ij}\, x_j$$

Ce produit caractérise A dans le sens où si deux matrices A et B vérifient AX = BX pour toute colonne X de $\mathbb{K}^p$, alors A = B. Il suffit en effet d'appliquer A et B sur le j-ème vecteur $\begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$ de la base canonique de $\mathbb{K}^p$ (où le 1 occupe la j-ème place) pour voir que A et B ont même colonne j. Ainsi :

$$\forall\, X \in \mathbb{K}^p,\ AX = BX \Rightarrow A = B$$

Deux matrices A et A' sont dites équivalentes en lignes si on peut passer de l'une à l'autre par des opérations élémentaires sur les lignes, ce qu'on note : $A \underset{L}{\sim} A'$. Il s'agit d'une relation d'équivalence. En effet :

$A \underset{L}{\sim} A$	en n'effectuant aucune opération sur les lignes
$A \underset{L}{\sim} A' \Rightarrow A' \underset{L}{\sim} A$	car les opérations étant inversibles, on peut passer de A' à A par les opérations inverses.
$A \underset{L}{\sim} A'$ et $A' \underset{L}{\sim} A'' \Rightarrow A \underset{L}{\sim} A''$	car on passe de A à A'' en effectuant naturellement les opérations faisant passer de A à A', puis de A' à A''

La méthode de résolution de Gauss-Jordan montre qu'une matrice quelconque est équivalente en lignes à une matrice échelonnée en ligne.

3– Rang d'un système

On appelle rang r du système ou rang de la matrice des coefficients A = (a_{ij}) le nombre d'inconnues principales (ou de pivots non nuls) une fois obtenu un système échelonné ou une matrice échelonnée. On note r = rg(A).

Le rang ne dépend pas de la façon dont les opérations ont été menées pour obtenir une matrice échelonnée. En effet, si la matrice A conduit à deux matrices échelonnées A' et A'', alors on peut passer de A' à A'' par des opérations sur les lignes. Mais la première ligne est la seule à avoir un pivot non nul sur la colonne d'indice le plus petit et aucune combinaison linéaire $\sum_{k=1}^{p} \lambda_k\, L_k$ avec $\lambda_1 \neq 1$ ne peut modifier ou diminuer cet indice puisque les coefficients des autres lignes sur la même colonne sont nuls. Ainsi, A' et A'' ont sur la première ligne un pivot occupant la même colonne. En ce qui concerne la deuxième ligne, on ne peut passer de la matrice A' à la matrice A" que par une combinaison linéaire de la forme $\sum_{k=2}^{p} \lambda_k \, L_k$, $\lambda_2 \neq 0$, la ligne L_1 ne pouvant intervenir faute de quoi un pivot occuperait la même colonne pour les deux premières lignes de A". Mais une combinaison linéaire $\sum_{k=2}^{p} \lambda_k \, L_k$, $\lambda_2 \neq 0$, donne un pivot sur la deuxième ligne occupant la même colonne pour les deux matrices. On poursuit le raisonnement successivement sur chacune des lignes, montrant que les pivots occupent les mêmes colonnes dans les deux matrices. Enfin, une ligne nulle dans A' ne peut par combinaison linéaire être transformée en une ligne non nulle de A" car cette dernière ligne aurait un pivot occupant une colonne identique à un pivot d'une ligne précédente. Ainsi, A' et A" ont même nombre de pivots.

L'unicité du rang de la matrice échelonnée permet également de conclure que les opérations élémentaires sur les lignes conservent le rang d'une matrice. Soit en effet une matrice A, modifiée en une matrice échelonnée A' ayant *r* pivots non nul, et soit B une matrice obtenue à partir de A par des opérations élémentaires sur les lignes. Alors $A \underset{L}{\sim} A'$ et $A \underset{L}{\sim} B$ donc $B \underset{L}{\sim} A'$ et A et B sont toutes deux de rang *r*.

En ce qui concerne la résolution d'un système, plusieurs cas se présentent :

❑ Si $r = n = p$: autant d'inconnues que d'équations, le nombre d'inconnues étant égal au rang. On dit qu'on a un système de Cramer. Après échelonnement, on obtient un système triangulaire sans paramètres. Il y a alors une solution et une seule :
EXEMPLE : $r = n = p = 3$

$$\begin{cases} x + 2y - 3z = 5 \\ \quad\; y - 2z = 2 \\ \qquad\quad z = 3 \end{cases}$$

Cela signifie que l'application $X \to AX$ est bijective, où A est la matrice des coefficients du système et X la colonne des inconnues.

❑ Si $r = n < p$: plus d'inconnues que d'équations, le nombre d'équations étant égal au rang. On obtient après échelonnement un système échelonné ayant $p - r$ paramètres. Il y a une infinité de solutions, quel que soit le second membre :
EXEMPLE : $r = n = 3 < p = 4$

$$\begin{cases} x + 2y - 3z + 2t = 5 \\ \quad\; y - 2z - t = 2 \\ \qquad\quad z + 3t = 3 \end{cases}$$

On résout le système en prenant *t* quelconque, et on en déduit *z*, puis *y*, puis *x* en fonction de *t*. Concrètement, le fait que le rang soit égal au nombre d'équations signifie que les équations sont indépendantes entre elles.
Cela signifie que l'application $X \to AX$ est surjective.

❑ Si $r = p < n$: plus d'équations que d'inconnues, le nombre d'inconnues étant égal au rang. Après échelonnement, on obtient un système échelonné sans paramètre, mais dont les $n - r$ dernières équations ont leur membre de gauche nul. Si les membres de droite sont également nuls, alors ces

équations sont inutiles et on se ramène au système de Cramer. Mais si l'un des membres de droite est non nul, cela signifie une incompatibilité entre les équations initiales et il n'y a pas de solution.
EXEMPLE : $r = p = 2 < n = 3$

$$\begin{cases} x + 2y = 5 \\ \quad y = 2 \\ \quad 0 = 0 \end{cases} \quad \text{provenant par exemple de} \begin{cases} x + 2y = 5 \\ 2x + 5y = 12 \\ x + 3y = 7 \end{cases} \quad \text{solution unique}$$

$$\begin{cases} x + 2y = 5 \\ \quad y = 2 \\ \quad 0 = 1 \end{cases} \quad \text{provenant par exemple de} \begin{cases} x + 2y = 5 \\ 2x + 5y = 12 \\ x + 3y = 8 \end{cases} \quad \text{pas de solution}$$

Cela signifie que l'application $X \to AX$ est injective. Dans le cas d'un second membre B nul, la seule solution est la solution nulle.

❑ Si $r < p$ et $r < n$, il y a $p - r$ paramètres, mais comme ci-dessus, $n - r$ équations ont leur membre de gauche nul. Ou bien il n'y a pas de solution si les équations sont incompatibles, ou bien il y en a une infinité.
EXEMPLE : $r = 2 < p = n = 3$

$$\begin{cases} 2x + y - z = 1 \\ x + 3y + z = 3 \\ 4x + 7y + z = 7 \end{cases}$$

$$\Leftrightarrow \begin{cases} x + 3y + z = 3 \\ \quad 5y + 3z = 5 \\ \quad 5y + 3z = 5 \end{cases} \quad \begin{matrix} L_1 \leftarrow L_2 \\ L_2 \leftarrow 2L_2 - L_1 \\ L_3 \leftarrow 4L_2 - L_3 \end{matrix}$$

$$\Leftrightarrow \begin{cases} x + 3y + z = 3 \\ \quad 5y + 3z = 5 \\ \qquad 0 = 0 \end{cases} \quad \begin{matrix} \\ \\ L_3 \leftarrow L_2 - L_3 \end{matrix}$$

$$\Leftrightarrow \begin{cases} x = \dfrac{4z}{5} \\ y = -\dfrac{3z}{5} + 1 \end{cases} \quad \text{L'ensemble des solutions est une droite affine}$$

Mais :

$$\begin{cases} 2x + y - z = 1 \\ x + 3y + z = 3 \\ 4x + 7y + z = -1 \end{cases}$$

$$\Leftrightarrow \begin{cases} x + 3y + z = 3 \\ \quad 5y + 3z = 5 \\ \quad 5y + 3z = 13 \end{cases} \quad \begin{matrix} L_1 \leftarrow L_2 \\ L_2 \leftarrow 2L_2 - L_1 \\ L_3 \leftarrow 4L_2 - L_3 \end{matrix}$$

L_2 et L_3 sont incompatibles (5 = 13). Il n'y a pas de solution.

II : Opérations sur les matrices

1– Somme de matrices

On note $\mathcal{M}_{np}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes, à coefficients dans $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On souhaite définir une addition sur cet ensemble de façon que cette addition satisfasse la relation :

$$(A + B)X = AX + BX$$

où X est un élément quelconque de $\mathbb{K}^p$. Si A = (a_{ij}) et B = (b_{ij}), le i-ème coefficient de AX est $\sum_{j=1}^{p} a_{ij}\, x_j$, celui de BX est $\sum_{j=1}^{p} b_{ij}\, x_j$, donc celui de (A + B)X est $\sum_{j=1}^{p} a_{ij}\, x_j + \sum_{j=1}^{p} b_{ij}\, x_j = \sum_{j=1}^{p} (a_{ij} + b_{ij})\, x_j$. A + B est donc la matrice dont le terme (i, j) est $a_{ij} + b_{ij}$:

DEFINITION

Soient A *et* B *deux matrices de* $\mathcal{M}_{np}(\mathbb{K})$. *Alors* A + B *est une matrice de* $\mathcal{M}_{np}(\mathbb{K})$ *dont le terme général est* :

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

On vérifie facilement que :

A + B = B + A

(A + B) + C = A + (B + C)

A + O = O + A = A où O désigne la matrice dont tous les coefficients sont nuls (matrice dite nulle)

A + (–A) = (–A) + A = O où A désigne la matrice de terme général – a_{ij}

2– Produit par un scalaire

On souhaite définir un produit par un scalaire λ élément de $\mathbb{K}$ de façon que :

$$\lambda(AX) = (\lambda A)X$$

où X est un élément quelconque de $\mathbb{K}^p$. Si A = (a_{ij}), le i-ème coefficient de AX est $\sum_{j=1}^{p} a_{ij}\, x_j$, celui de λ(AX) est $\lambda\sum_{j=1}^{p} a_{ij}\, x_j = \sum_{j=1}^{p} \lambda a_{ij}\, x_j$, donc celui de (λA)X est λa_{ij} :

DEFINITION

Soit A *une deux matrice de* $\mathcal{M}_{np}(\mathbb{K})$ *et* λ *un élément de* $\mathbb{K}$. *Alors* λA *est la matrice de* $\mathcal{M}_{np}(\mathbb{K})$ *dont le terme général est* :

$$(\lambda A)_{ij} = \lambda a_{ij}$$

On vérifie facilement que :

λ(μA) = (λμ)A

(λ + μ)A = λA + μA

λ(A + B) = λA + λB

1A = A

$\mathcal{M}_{np}(\mathbb{K})$ muni de l'addition et du produit par un scalaire est donc un espace vectoriel. Sa dimension est np. Une base est en effet constituée des matrices E_{ij} ayant tous leurs termes nuls sauf un seul qui vaut 1 et qui occupe la position (i, j). Si A est une matrice de terme général a_{ij}, on a simplement :

$$A = \sum_{i,j} a_{ij}\, E_{ij}$$

3– Produit de matrices

On souhaite définir un produit de matrice de façon que :

$$(AB)X = A(BX)$$

où X est un élément quelconque de $\mathbb{K}^q$. B doit donc disposer de q colonnes, donc est élément de $\mathcal{M}_{pq}(\mathbb{K})$ pour un certain p. Dans ce cas BX possède p lignes. Pour effectuer le produit A(BX), A doit donc posséder p colonnes, donc est élément de $\mathcal{M}_{np}(\mathbb{K})$ pour un certain n. A(BX) sera alors élément de $\mathbb{K}^n$. Cela signifie que AB doit appartenir à $\mathcal{M}_{nq}(\mathbb{K})$.

Si A = (a_{ij}) et B = (b_{ij}), le k-ème coefficient de Y = BX est $y_k = \sum\limits_{j=1}^{q} b_{kj}\, x_j$, et le i-ème élément de A(BX) = AY est $\sum\limits_{k=1}^{p} a_{ik}\, y_k$. Or :

$$\sum_{k=1}^{p} a_{ik}\, y_k = \sum_{k=1}^{p} a_{ik} \sum_{j=1}^{q} b_{kj}\, x_j = \sum_{k=1}^{p} \sum_{j=1}^{q} a_{ik}\, b_{kj}\, x_j = \sum_{j=1}^{q} \sum_{k=1}^{p} a_{ik}\, b_{kj}\, x_j$$

$$= \sum_{j=1}^{q} \Big(\sum_{k=1}^{p} a_{ik}\, b_{kj}\Big)\, x_j$$

donc le terme général AB est $\sum\limits_{k=1}^{p} a_{ik}\, b_{kj}$

DEFINITION

Soit A *élément de* $\mathcal{M}_{np}(\mathbb{K})$ *et* B *élément de* $\mathcal{M}_{pq}(\mathbb{K})$. *Alors* AB *est un élément de* $\mathcal{M}_{nq}(\mathbb{K})$ *dont le terme général est* :

$$(AB)_{ij} = \sum_{k=1}^{p} a_{ik}\, b_{kj}$$

Pour calculer le produit C = AB, on effectue le produit de la $i^{\text{ème}}$ ligne de A par la $j^{\text{ème}}$ colonne de B. La disposition usuelle de calcul est la suivante :

$$\begin{array}{cc} & \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1q} \\ \dots & & \dots & & \dots \\ b_{i1} & \dots & b_{ij} & \dots & b_{iq} \\ \dots & & \dots & & \dots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pq} \end{pmatrix} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ \dots & & \dots & \\ a_{i1} & a_{i2} & \dots & a_{ip} \\ \dots & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} & \begin{pmatrix} & & | & & \\ & & | & & \\ — & — & c_{ij} & & \\ & & & & \\ & & & & \end{pmatrix} \end{array}$$

On peut aussi dire qu'on multiplie successivement chaque colonne de B par la matrice A.

EXEMPLE :

Calculons $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & -1 & 4 \end{pmatrix} \times \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}$. On a :

$$\begin{array}{cc} & \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & -1 & 4 \end{pmatrix} & \begin{pmatrix} 2 & 0 \\ 2 & 3 \\ 1 & 0 \\ -1 & -4 \end{pmatrix} \end{array}$$

Les propriétés du produit sont (en supposant les nombres de lignes et de colonnes entre matrices compatibles entre elles) :

$A(B + B') = AB + AB'$
$(A + A')B = AB + A'B$
$(AB)C = A(BC)$ que l'on note ABC
$A(\lambda B) = \lambda(AB) = (\lambda A)B$ que l'on note λAB

La seule propriété non évidente est l'associativité $(AB)C = A(BC)$. Or, pour toute colonne X :

$((AB)C)X = (AB)(CX)$ par définition du produit de AB par C sur la colonne X
$= A(B(CX))$ par définition du produit A par B sur la colonne CX
$= A((BC)X)$ par définition du produit de B par C sur la colonne X
$= (A(BC)X)$ par définition du produit de A par BC sur la colonne X

Cette relation étant vérifiée pour toute colonne X (de dimension adéquate), on peut conclure que $(AB)C = A(BC)$.

On notera que si A appartient à $\mathcal{M}_{np}(\mathbb{K})$ et n appartient à $\mathcal{M}_{pq}(\mathbb{K})$, alors MN existe et appartient à $\mathcal{M}_{nq}(\mathbb{K})$, mais que, si q est différent de n, BA n'existe pas. Le produit n'est donc pas commutatif. Il ne l'est pas même si les matrices sont carrées de même taille.

EXEMPLE :

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix}$$

mais $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix}$

On prendra donc garde qu'en général $AB \neq BA$

En particulier, $(A + B)^2$ n'est pas égal à $A^2 + 2AB + B^2$, mais à $A^2 + AB + BA + B^2$. Ce défaut de commutativité du produit est source importante d'erreur chez les étudiants qui débutent l'étude des matrices.

Une autre propriété source d'erreur provient du fait que **AB peut être nul alors que ni A ni B n'est nul**.

EXEMPLE :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

En particulier, **on ne peut rien conclure du fait que AB = 0 et certainement pas que l'une des deux matrices est nulle.**

Il résulte de la définition du produit de matrice qu'on peut assimiler une colonne X élément de $\mathbb{K}^p$ à une matrice à une colonne $\mathcal{M}_{n1}(\mathbb{K})$. Le résultat AX est le même dans les deux cas.

III : Anneau des matrices carrées

1– Définition

On note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées à n lignes et n colonnes. Dans ce cas, le produit des matrices est un produit interne à cet ensemble :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \forall B \in \mathcal{M}_n(\mathbb{K}), AB \in \mathcal{M}_n(\mathbb{K})$$

Pour les MPSI : $(\mathcal{M}_n(\mathbb{K}),+,\times)$ est un anneau.

2– Matrice identité

Cette matrice, notée I_n est associée à l'application identique : $X \in \mathbb{K}^n \to I_nX = X$. Elle vaut :

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & \dots & & \dots & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

comme on peut le vérifier facilement en multipliant une colonne X par cette matrice. Son terme général est noté δ_{ij} (symbole de Kronecker).

Si A est élément de $\mathcal{M}_{np}(\mathbb{K})$, alors :

$$I_nA = AI_p = A$$

En effet, pour tout $(i, j) \in \{1, ..., n\} \times \{1, ..., p\}$:

$$(I_nA)_{ij} = \sum_{k=1}^{n} \delta_{ik}\, a_{kj}$$

mais tous les δ_{ik} sont nuls sauf pour $k = i$, donc $(I_nA)_{ij} = a_{ij}$. Donc $I_nA = A$. On procède de même pour AI_p.

On aurait pu dire aussi que, pour toute colonne X élément de $\mathbb{K}^p$:

$(I_nA)X = I_n(AX)$ par définition du produit de I_n par A, appliqué sur la colonne X

$= AX$ par définition du I_n

donc $I_nA = A$.

3– Matrices particulières

<u>a) Matrices scalaires :</u>

Ce sont les matrices de la forme λI. Tous les termes sont nuls sauf sur la diagonale, où ils sont tous égaux entre eux. Ce sont les matrices associées aux homothéties $X \to \lambda X$.

b) Matrices diagonales :

Ce sont les matrices dont tous les termes sont nuls en dehors de la diagonale. Ceux de la diagonale sont quelconques : $i \neq j \Rightarrow a_{ij} = 0$.

$$A = \begin{pmatrix} a_1 & 0 & 0 & \ldots & 0 \\ 0 & a_2 & 0 & \ldots & 0 \\ 0 & 0 & a_3 & \ldots & 0 \\ & \ldots & & \ldots & \\ 0 & 0 & 0 & \ldots & a_n \end{pmatrix}$$

On pourra vérifier que, si A et B sont toutes deux diagonales, le produit AB est diagonale et que les termes de la diagonale sont les $a_i b_i$.

c) Matrices triangulaires :

On traitera des matrices triangulaires supérieures, le cas des matrices triangulaires inférieures se traitant de manière analogue. Elles sont de la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & a_{33} & \ldots & a_{3n} \\ & \ldots & & \ldots & \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

et caractérisées par : $i > j \Rightarrow a_{ij} = 0$

Elles forment un sous-espace vectoriel de dimension $\dfrac{n(n+1)}{2}$ (une base de ce sous-espace vectoriel est formée des matrices dont tous les coefficients sont nuls, sauf un seul qui vaut 1 et qui occupe l'une des $\dfrac{n(n+1)}{2}$ positions du triangle supérieur de la matrice). La seule chose non évidente est de vérifier la stabilité pour le produit de matrices. Soient donc deux matrices A et B triangulaires supérieures et soit $i > j$. Le terme (i,j) du produit est donné par $\sum_{k=1}^{n} a_{ik}\, b_{kj}$ qui est nul car $a_{ik} = 0$ pour $i > k$ et $b_{kj} = 0$ pour $k > j$. Comme $i > j$, pour chaque k, il existe l'un des deux termes a_{ik} ou b_{kj} qui est nul.

d) Matrices nilpotentes :

Une matrice M est dite nilpotente s'il existe un entier n tel que $M^n = 0$.

4– Matrices inversibles

On dit que A élément de $\mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si il existe M élément de $\mathcal{M}_n(\mathbb{K})$ tel que $AM = MA = I_n$. On note alors $M = A^{-1}$. Le produit de matrices étant non commutatif, il est a priori indispensable de préciser qu'on a les deux relations $AM = I_n$ et $MA = I_n$. Cependant, nous allons montrer ci-dessous qu'il suffit d'une égalité pour avoir l'autre.

Remarque préliminaire : effectuer une opération élémentaire sur les lignes de A revient à multiplier A à gauche par une matrice, et effectuer une opération élémentaire sur les colonnes revient à la multiplier à droite. On vérifiera en effet que :

- Multiplier la $j^{\text{ème}}$ ligne de A par λ est équivalent à multiplier A à gauche par la matrice :

$$\begin{array}{c} j \\ \downarrow \end{array}$$

$$\begin{pmatrix} 1 & 0 & & \dots & & 0 \\ 0 & 1 & & \dots & & 0 \\ & & \dots & & \dots & \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ & & \dots & & \dots & \\ 0 & 0 & & \dots & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j \\ \\ \\ \end{matrix} \text{(matrice de dilatation)}$$

- Retrancher la ligne j à la ligne i est équivalent à multiplier A à gauche par la matrice :

$$\begin{array}{c} \begin{matrix} & & & j & & \\ & & & \downarrow & & \end{matrix} \\ \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ & & \dots & -1 & \dots & \\ & \dots & & \dots & & 0 \\ 0 & 0 & \dots & 1 & \dots & 0 \\ & & \dots & & \dots & \\ 0 & 0 & & \dots & & 1 \end{pmatrix} \end{array} \begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix} \text{(matrice de transvection } t_{ij}(-1)\text{)}$$

- Permuter les colonnes i et j est équivalent à multiplier A à gauche par la matrice :

$$\begin{array}{c} \begin{matrix} & & i & & j & \\ & & \downarrow & & \downarrow & \end{matrix} \\ \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ & \dots & 0 & \dots & 1 & \\ \dots & & \dots & & \dots & \\ & \dots & 1 & \dots & 0 & \dots \\ & & \dots & & \dots & \\ 0 & 0 & & \dots & & 1 \end{pmatrix} \end{array} \begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix} \text{(matrice de transposition)}$$

Ainsi, toutes les opérations élémentaires sur les lignes de A sont équivalentes à des produits de matrices à gauche de A. Effectuer plusieurs opérations élémentaires sur les lignes revient donc à multiplier A à gauche par des matrices élémentaires M_1, puis M_2, ..., puis M_k et donc à multiplier A à gauche par la matrice $M = M_k M_{k-1} \dots M_2 M_1$.

On vérifiera de même que les produits à droite donne les opérations élémentaires sur les colonnes de A. On notera $A \underset{C}{\sim} B$ le fait de passer de A à B par des opérations élémentaires sur les colonnes.

Remarquons qu'une matrice M produit de matrices élémentaires est de rang n puisque les opérations élémentaires appliquées sur M en sens inverse vont mener de la matrice M à la matrice échelonnée I_n. Il en est de même si on raisonne sur les colonnes : le rang des colonnes d'une matrice N produit de matrices correspondant à des opérations élémentaires sur les colonnes est n.

PROPOSITION :

Dans $\mathcal{M}_n(\mathbb{K})$, *il y a équivalence entre :*

i) A *est inversible*

ii) $A \underset{L}{\sim} I_n$

iii) L'équation $AX = 0$, $X \in \mathbb{K}^n$, admet 0 comme seule solution.

iv) $\forall B \in \mathbb{K}^n$, l'équation $AX = B$ admet au plus une solution

v) $\forall B \in \mathbb{K}^n$, l'équation $AX = B$ admet au moins une solution

Démonstration :

❑ i) $\Rightarrow$ iii)
car si AX = 0 et si M est une matrice telle que MA = I_n, alors X = MAX = M0 = 0.

❑ ii) $\Rightarrow$ iii)
Si $A \underset{L}{\sim} I_n$, alors des opérations élémentaires sur les lignes de A transforment A en I_n. D'après la remarque préliminaire, cela revient à dire qu'il existe une matrice M, égale au produit des matrices élémentaires correspondant à ces opérations élémentaires, telle que MA = I_n, et on conclut comme ci-dessus.

❑ iii) $\Rightarrow$ ii)
car le fait qu'il y a seulement la solution nulle à AX = 0 montre que la résolution de Gauss va conduire à un système échelonné avec n lignes. Comme il y a n équations, cela signifie que A est équivalente en lignes à une matrice triangulaire supérieure $\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ & \dots & & \dots & \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$, les éléments diagonaux a_{ii} étant non nuls. Mais il n'est pas difficile, à partir de cette matrice triangulaire, d'effectuer des opérations élémentaires sur les lignes de façon à annuler le triangle supérieur, arrivant ainsi à la matrice $\begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & 0 \\ & \dots & & \dots & \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$ (commencer par utiliser la dernière ligne pour annuler les coefficients de la dernière colonne en utilisant le fait que $a_{nn} \neq 0$). Il suffit enfin d'effectuer les opérations $L_i \leftarrow \dfrac{1}{a_{ii}} L_i$ pour obtenir I_n. On a alors montré que $A \underset{L}{\sim} I_n$.

❑ iii) $\Rightarrow$ iv)
En effet, si X_1 et X_2 sont deux solutions de AX = B, alors $A(X_1 - X_2) = B - B = 0$, donc, d'après iii), $X_1 - X_2 = 0$.

❑ iv) $\Rightarrow$ iii)
Appliquer iv) en prenant B = 0

❑ ii) $\Rightarrow$ v)
ii) signifie que des opérations élémentaires sur les lignes va transformer le système AX = B en un système équivalent MAX = X = MB où MB est obtenu à partir de B par les mêmes opérations qui ont transformé A en I_n, correspondant à un produit à gauche par une matrice M. Donc il y a une solution au système, à savoir X = MB.

❑ v) $\Rightarrow$ ii)
De même, si v) est vraie, cela signifie que la résolution du système AX = B conduit à un système échelonné avec n lignes non nulles. En effet, si l'une des lignes du système est nulle et si M est la matrice dont le produit à gauche de A conduit à ce système, on aurait MAX = MB avec l'une des

lignes de MA nulle. S'il s'agit de la ligne k, on aurait donc $0 = \sum_{j=1}^{n} m_{kj}\, b_j$. Mais si on prend $b_j = \overline{m_{kj}}$, on obtient $0 = \sum_{j=1}^{n} |m_{kj}|^2$ et donc la k-ème ligne de M serait nulle. Mais cela est impossible, car cela voudrait dire que le rang de M est strictement inférieur à n, or M est obtenu à partir d'opérations élémentaires sur les lignes donc est de rang n. A est donc équivalente en lignes à une matrice de la forme $\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ & \dots & & \dots & \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$ avec les $a_{ii} \neq 0$ et on conclut de la même façon que iii) $\Rightarrow$ ii).

On a donc montré que ii) $\Leftrightarrow$ iii) $\Leftrightarrow$ iv) $\Leftrightarrow$ v) et que i) $\Rightarrow$ iii). Pour conclure, il suffit donc de montrer que ii) $\Rightarrow$ i) :

❑ Nous avons déjà vu que, si $A \underset{L}{\sim} I_n$, alors des opérations élémentaires sur les lignes de A transforment A en I_n et donc, d'après la remarque préliminaire, il existe une matrice M égale au produit des matrices élémentaires correspondant à ces opérations élémentaires, telle que $MA = I_n$. Procédons maintenant à des opérations élémentaires sur les colonnes, de façon à obtenir une matrice échelonnée selon les colonnes. Cela revient à multiplier à droite par une matrice, produit des matrices élémentaires correspondant chacune à une opération élémentaire sur les colonnes.
Si on parvient à une matrice échelonnée en colonne, triangulaire avec ses termes diagonaux non nuls, on pourra poursuivre les opérations élémentaires sur les colonnes pour parvenir à la matrice I_n (comme on l'a fait dans la démonstration de iii) $\Rightarrow$ ii) pour les lignes), ce qui signifie qu'il existe une matrice N telle que $AN = I_n$. Mais on a alors $MAN = (MA)N = I_nN = N = M(AN) = MI_n = M$, donc $M = N$ et A est bien inversible.
Dans le cas contraire, on obtiendrait une matrice échelonnée en colonne avec une colonne nulle. Autrement dit, il existerait une matrice N produit de matrices d'opérations élémentaires en colonnes, telle que AN possède une colonne nulle. Mais ceci est impossible. Si AN a une colonne nulle, il en est de même de M(AN) or $MAN = (MA)N = I_nN = N$, donc N possède une colonne nulle. Donc le rang de ses colonnes est strictement inférieur à n. Ceci est contradictoire avec le fait que N est produit de matrices d'opérations élémentaires en colonnes, puisque, pour une telle matrice le rang de ses colonnes vaut n.

L'ensemble des matrices inversibles, muni du produit des matrices s'appelle groupe linéaire et est noté $GL_n(\mathbb{K})$. Si A et B sont inversibles, il en est de même de leur produit et

$$(AB)^{-1} = B^{-1}A^{-1}$$

En effet, $(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$.

Le calcul de l'inverse d'une matrice se fait de la façon suivante :

❑ Pour la matrice 2×2 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, elle est inversible si et seulement si $ad - bc \neq 0$ et son inverse vaut $\dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ comme on le vérifie en multipliant cet inverse par la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ initiale.

❑ Si $A \in \mathcal{M}_n(\mathbb{K})$ et s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ tel que $AB = I_n$ (ou $BA = I_n$) alors $B = A^{-1}$. Par exemple, une matrice A vérifiant la relation $A^3 + 3A^2 - A + I_n = 0$ est inversible puisque :

$$I = -A^3 - 3A^2 + A = A \times (-A^2 - 3A + I)$$

et $\quad A^{-1} = -A^2 - 3A + I$.

❑ Méthode générale : on résout le système $AX = Y$ où X et Y sont des vecteurs colonnes à n lignes. Si A est inversible, alors on aura $X = A^{-1}AX = A^{-1}Y$ et la solution du système est unique. L'expression de X en fonction de Y possède des coefficients qui sont précisément ceux de la matrice inverse.
Réciproquement, s'il existe une solution unique au système pour tout Y, alors A est inversible. Prenons en effet pour Y les colonnes successives E_k de base canonique de $\mathbb{K}^n$ et notons B la matrice dont la k-ème colonne est la solution X du système $AX = E_k$. On aura alors $AB = I_n$.

EXEMPLE : Calculer l'inverse de $\begin{pmatrix} 10 & 9 & 1 \\ 9 & 10 & 5 \\ 1 & 5 & 9 \end{pmatrix}$

$$\begin{cases} 10x + 9y + z = x' \\ 9x + 10y + 5z = y' \\ x + 5y + 9z = z' \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 41y + 89z = 10z' - x' & 10L_3 - L_1 \to L_1 \\ 35y + 76z = 9z' - y' & 9L_3 - L_2 \to L_2 \\ x + 5y + 9z = z' \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x + 5y + 9z = z' & L_3 \to L_1 \\ 35y + 76z = 9z' - y' \\ z = 19z' - 41y' + 35x' & 41L_2 - 35L_1 \to L_3 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x = 65x' + 76y' + 35z' \\ y = -76x' + 89y' - 41z' \\ z = 35x' - 41y' + 19z' \end{cases}$$

Donc la matrice inverse vaut $\begin{pmatrix} 65 & 76 & 35 \\ -76 & 89 & -41 \\ 35 & -41 & 19 \end{pmatrix}$

❑ Nous donnons enfin une méthode utilisant les matrices élémentaires, correspondant aux opérations élémentaires sur les lignes ou les colonnes de A. On peut alors inverser une matrice de la façon suivante : on choisit une fois pour toutes d'effectuer les mêmes opérations sur les lignes, ou (strictement) sur les colonnes, et ceci à la fois sur A et sur la matrice Identité. Cela revient à multiplier A et I_n par la même matrice M (à droite si on agit sur les colonnes, à gauche si on agit sur les lignes). Dans le cas d'opérations sur les colonnes, en cours de calcul, on dispose donc des matrices AM et $I_nM = M$. Si l'on parvient finalement à AM et I_nM avec $AM = I_n$, alors nécessairement $I_nM = M = A^{-1}$.

EXEMPLE 1 : Calculer l'inverse de $\begin{pmatrix} 2 & 4 & -1 \\ 1 & -2 & 3 \\ 1 & 1 & 1/2 \end{pmatrix}$

$$\begin{pmatrix} 2 & 4 & -1 \\ 1 & -2 & 3 \\ 1 & 1 & 1/2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$C_2 \leftarrow 2C_1 - C_2$
$C_3 \leftarrow C_1 + 2C_3$

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 7 \\ 1 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$C_3 \leftarrow 4C_3 - 7C_2$

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & -10 \\ 0 & -1 & 7 \\ 0 & 0 & 8 \end{pmatrix}$$

$C_2 \leftarrow C_2 - C_3$

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 12 & -10 \\ 0 & -8 & 7 \\ 0 & -8 & 8 \end{pmatrix}$$

$C_2 \leftarrow C_2/4$

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 & -10 \\ 0 & -2 & 7 \\ 0 & -2 & 8 \end{pmatrix}$$

$C_1 \leftarrow (C_1 - C_2 - C_3)/2$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 4 & 3 & -10 \\ -5/2 & -2 & 7 \\ -3 & -2 & 8 \end{pmatrix}$$

Donc la matrice inverse vaut $\begin{pmatrix} 4 & 3 & -10 \\ -5/2 & -2 & 7 \\ -3 & -2 & 8 \end{pmatrix}$

EXEMPLE 2 : Calculer l'inverse de $\begin{pmatrix} 10 & 9 & 1 \\ 9 & 10 & 5 \\ 1 & 5 & 9 \end{pmatrix}$

$$\begin{pmatrix} 10 & 9 & 1 \\ 9 & 10 & 5 \\ 1 & 5 & 9 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$C_2 \leftarrow 10C_2 - 9C_1$
$C_3 \leftarrow 10C_3 - C_1$

$$\begin{pmatrix} 10 & 0 & 0 \\ 9 & 19 & 41 \\ 1 & 41 & 89 \end{pmatrix} \qquad \begin{pmatrix} 1 & -9 & -1 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{pmatrix}$$

$C_3 \leftarrow -41C_2 + 19C_3$

$$\begin{pmatrix} 10 & 0 & 0 \\ 9 & 19 & 0 \\ 1 & 41 & 10 \end{pmatrix} \qquad \begin{pmatrix} 1 & -9 & -350 \\ 0 & 10 & -410 \\ 0 & 0 & 190 \end{pmatrix}$$

$C_3 \leftarrow C_3/10$

$$\begin{pmatrix} 10 & 0 & 0 \\ 9 & 19 & 0 \\ 1 & 41 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & -9 & -35 \\ 0 & 10 & -41 \\ 0 & 0 & 19 \end{pmatrix}$$

$C_2 \leftarrow C_2 - 41C_3$

$$\begin{pmatrix} 10 & 0 & 0 \\ 9 & 19 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & -1444 & -35 \\ 0 & 1691 & -41 \\ 0 & -779 & 19 \end{pmatrix}$$

$C_2 \leftarrow C_2/19$

$$\begin{pmatrix} 10 & 0 & 0 \\ 9 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & -76 & -35 \\ 0 & 89 & -41 \\ 0 & -41 & 19 \end{pmatrix}$$

$C_1 \leftarrow (C_1 - 9C_2 - C_3)/10$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 65 & -76 & -35 \\ -76 & 89 & -41 \\ 35 & -41 & 19 \end{pmatrix}$$

Donc la matrice inverse vaut $\begin{pmatrix} 65 & -76 & -35 \\ -76 & 89 & -41 \\ 35 & -41 & 19 \end{pmatrix}$

EXEMPLE 3 : Soit A la matrice $\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ \dots & & \dots & & \dots & \\ 1 & 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix}$

On obtient directement la matrice I_n au moyen des opérations $C_i \leftarrow C_i - C_{i+1}$. En appliquant les mêmes règles sur I_n, on obtient A^{-1}, à savoir $\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 \\ \dots & & \dots & & \dots & \\ 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}$

L'algorithme général est le suivant :

i) Obtenir une matrice triangulaire inférieure : Pour i variant de 1 à n–1, placer en colonne i un vecteur parmi ceux occupant les colonnes i à n dont le $i^{\text{ème}}$ terme est non nul (ce qui est possible sinon, la matrice ne serait pas de rang n et ne serait pas inversible) et faire des combinaisons entre ce $i^{\text{ème}}$ vecteur et les vecteurs placés à sa droite de façon à annuler les termes situés en la ligne i et en colonne k, $k > i$. Ceci s'obtient par utilisation répétée des trois manipulations élémentaires. On remplit donc de 0 le triangle supérieur.

ii) D'une façon analogue, obtenir une matrice diagonale en remplissant de 0 le triangle inférieur, mais en procédant de gauche à droite, en remplissant de 0 les lignes les plus basses.

iii) Obtenir la matrice identité, en divisant chaque colonne par le terme de la diagonale occupant cette colonne.

On peut également effectuer le calcul sur les lignes. Il est comparable à celui qui est mené dans la méthode de Gauss pour résoudre un système. En particulier, si le rang de la matrice est n, on arrivera de façon certaine à une matrice échelonnée triangulaire à diagonale non nulle, ce qui assure qu'on pourra terminer le calcul.

IV : Transposition

1– Définition

On considère l'application :

$$\mathcal{M}_{np}(\mathbb{K}) \longrightarrow \mathcal{M}_{pn}(\mathbb{K})$$

$$A \longrightarrow {}^{t}A \text{ appelée transposée de A.}$$

définie par : $({}^{t}A)_{ij} = A_{ji}$, $1 \leq i \leq p$, $1 \leq j \leq n$. On note aussi la transposée A^T.

EXEMPLE :

$${}^{t}\begin{pmatrix} 2 & 3 \\ 1 & 0 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 6 \\ 3 & 0 & 5 \end{pmatrix}$$

2– Propriétés

Il est aisé de vérifier :

i) ${}^{t}({}^{t}A) = A$

ii) ${}^{t}(A + B) = {}^{t}A + {}^{t}B$

iii) ${}^{t}(\lambda A) = \lambda {}^{t}A$ la transposée est donc linéaire

iv) ${}^{t}(AB) = {}^{t}B{}^{t}A$ pour A élément de $\mathcal{M}_{np}(\mathbb{K})$ et B élément de $\mathcal{M}_{pq}(\mathbb{R})$

v) ${}^{t}(A^{-1}) = ({}^{t}A)^{-1}$ pour A élément de $GL_n(\mathbb{K})$

Pour la relation iv), on a en effet :

$$({}^{t}(AB))_{ij} = (AB)_{ji} = \sum_{k=1}^{p} A_{jk}\, B_{ki} = \sum_{k=1}^{p} ({}^{t}A)_{kj}({}^{t}B)_{ik} = \sum_{k=1}^{p} ({}^{t}B)_{ik}({}^{t}A)_{kj} = ({}^{t}B{}^{t}A)_{ij}$$

En ce qui concerne v), on a :

A inversible

$\Leftrightarrow$ $AA^{-1} = I$

$\Leftrightarrow$ ${}^{t}(AA^{-1}) = I$

$\Leftrightarrow$ ${}^{t}(A^{-1}){}^{t}A = I$

ce qui prouve que ${}^{t}A$ est inversible d'inverse ${}^{t}(A^{-1})$.

Dans le chapitre L(E-F).PDF portant sur les applications linéaires, on montre que A et sa transposée ont même rang.

3– Matrices symétriques et antisymétriques

Soit A une matrice carrée. On dit que A est symétrique si ${}^{t}A = A$. On dit que A est antisymétrique si ${}^{t}A = -A$. L'ensemble des matrices symétriques et l'ensemble des matrices antisymétriques sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$. Montrons qu'ils sont supplémentaires :

❑ Si A est à la fois symétrique et antisymétrique, ${}^{t}A = A = -A$ donc A est nulle. La somme est donc directe.

❑ Toute matrice carrée s'écrit comme somme d'une matrice symétrique et antisymétrique :

$$M = \frac{M + {}^tM}{2} + \frac{M - {}^tM}{2}$$

Ainsi $\begin{pmatrix} 7 & 5 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 2 & 5 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$

La dimension du sous-espace des matrices symétriques est $\dfrac{n(n+1)}{2}$, une base étant $(E_{ij} + E_{ji})$, $i \le j$; la dimension du sous-espace des matrices antisymétriques est $\dfrac{n(n-1)}{2}$, une base étant $(E_{ij} - E_{ji})$, $i < j$.

Annexe : utilisation des matrices en physique

Les utilisations des matrices en physique sont innombrables. On en donne quelques exemples ci-dessous.

1– Matrice d'inertie

Soit M un point de masse m se déplaçant à la vitesse $\boldsymbol{V}$ par rapport à un référentiel et O un point de ce référentiel. On appelle moment cinétique de M par rapport à O le vecteur :

$$\boldsymbol{L}_O = \boldsymbol{OM} \wedge m\boldsymbol{V}$$

L'intérêt du moment cinétique tient dans le fait que sa dérivée est égale aux moments des forces appliquées en M par rapport à O (y compris les forces d'inertie si le référentiel n'est pas galiléen), lorsque O est fixe dans le référentiel considéré.

Supposons que M effectue un mouvement de rotation autour d'un axe passant par O. La vitesse $\boldsymbol{V}$ est alors de la forme $\boldsymbol{V} = \omega \wedge \boldsymbol{OM}$, où ω est un vecteur appelé vecteur instantané de rotation et $\wedge$ désigne le produit vectoriel (voir l'annexe du chapitre sur les déterminants). L'axe (O,ω) est l'axe de rotation. On a alors :

$$\boldsymbol{L}_O = m\,\boldsymbol{OM} \wedge (\omega \wedge \boldsymbol{OM})$$

On introduit alors un opérateur, appelé opérateur d'inertie, qui, à tout vecteur $\boldsymbol{u}$, associe le vecteur $\boldsymbol{J}(\boldsymbol{u}) = m\,\boldsymbol{OM} \wedge (\boldsymbol{u} \wedge \boldsymbol{OM})$, de sorte que $\boldsymbol{L}_O = \boldsymbol{J}(\omega)$. En utilisant la formule du double produit vectoriel, on obtient également :

$$\boldsymbol{J}(\boldsymbol{u}) = m\,OM^2\boldsymbol{u} - m\,\langle \boldsymbol{u},\boldsymbol{OM}\rangle\,\boldsymbol{OM} = m\,(\boldsymbol{OM} \wedge \boldsymbol{u}) \wedge \boldsymbol{OM}$$

En ce qui concerne un solide, on opère de même en sommant au moyen d'une intégrale triple sur tous les points du solide. L'opérateur d'inertie prend alors la forme :

$$\boldsymbol{J}(\boldsymbol{u}) = \iiint_V \boldsymbol{OM} \wedge (\boldsymbol{u} \wedge \boldsymbol{OM})\,\mathrm{d}m$$

$$= \iiint_V OM^2\boldsymbol{u} - \langle \boldsymbol{u},\boldsymbol{OM}\rangle\,\boldsymbol{OM}\,\mathrm{d}m$$

$\boldsymbol{J}$ est clairement linéaire et est donc représenté par une matrice 3×3 dont on montre qu'elle est symétrique. La matrice de $\boldsymbol{J}$ s'écrit donc $\begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix}$ avec :

$A = \langle \boldsymbol{J}(\boldsymbol{i}),\boldsymbol{i}\rangle = \iiint_V \mathrm{OM}^2 - (\boldsymbol{i}.\boldsymbol{OM})^2 \,\mathrm{d}m = \iiint_V (y^2 + z^2)\,\mathrm{d}m$,

moment d'inertie par rapport à l'axe $(O,\boldsymbol{i})$

$B = \iiint_V (x^2+z^2)\,\mathrm{d}m$, moment d'inertie par rapport à l'axe $(O,\boldsymbol{j})$

$C = \iiint_V (x^2+y^2)\,\mathrm{d}m$, moment d'inertie par rapport à l'axe $(O,\boldsymbol{k})$

$D = -\langle \boldsymbol{J}(\boldsymbol{k}),\boldsymbol{j}\rangle = \iiint_V yz\,\mathrm{d}m$, quantité appelée produit d'inertie.

$E = \iiint_V xz\,\mathrm{d}m$

$F = \iiint_V xy\,\mathrm{d}m$

Etant symétrique, on montre qu'elle est diagonalisable dans un repère orthonormé dont les axes sont appelés axes principaux d'inertie. (cf le fichier PREHILB.PDF traitant du chapitre *Espaces préhilbertiens* du cours de deuxième année pour plus de détails).

2– Réseaux de conducteurs électriques

On considère un réseau électrique formé de n mailles. On attribue à chaque maille un courant électrique fictif orienté dit courant de maille (ou courant de Maxwell). Il y a donc n courants de maille I_1, ..., I_n. Le courant réel observé dans un conducteur commun à plusieurs mailles est la somme algébrique des courants des mailles auxquelles il appartient. Si chaque maille dispose d'une force électromotrice E_1, ..., E_n, alors on dispose d'une relation matricielle :

$$(E) = (R)(I)$$

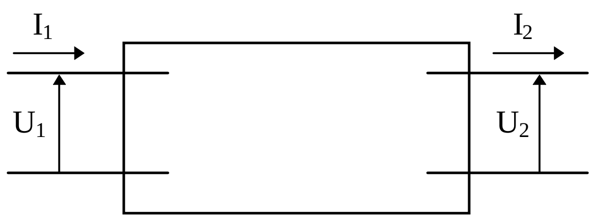
où R est une matrice $n \times n$. Dans le cas de mailles extérieures les unes aux autres, toutes orientées dans le même sens (par exemple le sens trigonométrique), on montre que :

- ❑ R_{ii} est la somme des résistances totales de la maille i
- ❑ R_{ij} est l'opposé de la somme des résistances communes aux mailles i et j.

En particulier, cette matrice est symétrique. Pour connaître les intensités connaissant les forces électromotrices, il suffit d'inverser la matrice R.

3– Quadripôles

On considère le système électrique suivant, appelé quadripôle :



Le cadre contient un réseau électrique purement passif. Dans ce cas, les courants I_1 et I_2 dépendent linéairement de U_1 et U_2. En effet, on peut les considérer comme les courants de mailles extérieures de sorte que l'on a :

$$\begin{pmatrix} I_1 \\ I_2 \\ \dots \\ \dots \end{pmatrix} = (R)^{-1} \begin{pmatrix} U_1 \\ -U_2 \\ \dots \\ \dots \end{pmatrix}$$

avec $(R)^{-1}$ inverse de la matrice (R) définie au paragraphe précédent. Le signe de U_2 provient de l'orientation positive choisie pour U_2 ici, en sens contraire du courant. $(R)^{-1}$ est symétrique. On obtient donc, en introduisant le signe de $-U_2$ dans la matrice :

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & c \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

On en déduit une expression :

$$\begin{pmatrix} U_2 \\ I_2 \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} U_1 \\ I_1 \end{pmatrix}$$

dont on peut vérifier, en l'exprimant à partir de a, b, c que le déterminant est égal à 1. La dernière matrice utilisée s'appelle matrice de transfert. Cette disposition permet le calcul aisé d'une suite de quadripôles en séries. Il apparaît alors un produit de matrices.

4– Electrostatique

On considère n conducteurs électriques de charges $Q_1, Q_2, \dots, Q_n$. Ces conducteurs sont portés aux potentiels $V_1, V_2, \dots, V_n$. L'expérience prouve que la dépendance des charges en fonction des potentiels est linéaire, et donc qu'il existe une matrice (C) telle que :

$$(Q) = (C)(V)$$

où (Q) est la matrice colonne de composantes Q_i et (V) la matrice colonne de composantes V_i.

Les éléments diagonaux valent C_{ii}, sont appelés capacité du conducteur i en présence des autres conducteurs, et sont positifs. Les éléments C_{ij}, pour i différent de j, s'appellent coefficient d'influence du conducteur i sur le conducteur j. Ils sont négatifs. Cette matrice est par ailleurs symétrique. Cette propriété repose sur le principe de conservation de l'énergie. En effet, on montre que, si l'on charge d'abord le conducteur j d'une charge Q_j, puis le conducteur i d'une charge Q_i, l'énergie nécessaire vaut $(C^{-1})_{ij}Q_{ii}^2 + (C^{-1})_{ij}Q_jQ_i + \frac{1}{2}(C^{-1})_{jj}Q_j^2$. Si l'on commence d'abord par le conducteur i puis par le conducteur j, l'énergie vaut $(C^{-1})_{ij}Q_{ii}^2 + (C^{-1})_{ji}Q_jQ_i + \frac{1}{2}(C^{-1})_{jj}Q_j^2$. On a donc $(C^{-1})_{ij} = (C^{-1})_{ji}$. La matrice C^{-1} et donc la matrice C est symétrique. Indiquons enfin que l'énergie électrostatique nécessaire pour charger les conducteurs i à la charge Q_i vaut :

$$E = \frac{1}{2}\,{}^tQ\,C^{-1}\,Q \quad \text{où Q est la matrice colonne de composantes } Q_i$$

$$= \frac{1}{2}\,{}^tV\,C\,V \quad \text{où V est la matrice colonne de composantes } V_i.$$

$$= \frac{1}{2}\,{}^tQ\,V$$

5– Inductance mutuelle

De même, considérons n circuits électriques parcourus par les courants I_i. Ils créent un champ magnétique $\boldsymbol{B}$ dont le flux à travers le circuit i est Φ_i. On a une relation du type :

$$(\Phi) = (L)(I)$$

où (Φ) est le vecteur colonne dont les composantes sont Φ_i et (I) le vecteur colonne dont les composantes sont I_i. (L) est une matrice carrée $n \times n$ de terme général L_{ij} avec :

❑ L_{ii} inductance propre du circuit *i*
❑ L_{ij} inductance mutuelle du circuit *i* sur le circuit *j*.
Cette matrice est elle aussi symétrique. L'énergie magnétostatique du système s'exprime sous la forme :

$$\begin{aligned} E &= \frac{1}{2}\,{}^{t}\Phi\ I \\ &= \frac{1}{2}\,{}^{t}I\ L\ I \\ &= \frac{1}{2}\,{}^{t}\Phi\ L^{-1}\ \Phi \end{aligned}$$

6– Polarisation

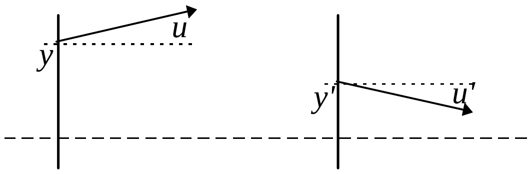
Considérons un corps, par exemple un cristal, plongé dans un champ électrique ***E***. Les atomes de ce corps subissent en général une polarisation. Il en résulte une polarisation globale ***P***, qui n'est pas en général colinéaire à ***E***. Il existe une transformation linéaire M telle que :

$$\boldsymbol{P} = \mathrm{M}\boldsymbol{E}$$

En considérant l'énergie nécessaire à une polarisation d'abord suivant l'axe des *x*, puis suivant l'axe des *y*, ou d'abord suivant l'axe des *y* puis suivant l'axe des *x*, on peut montrer là aussi que M est symétrique. Le fait qu'une matrice symétrique soit diagonalisable dans une base orthonormée signifie qu'il existe trois directions de polarisation privilégiée, pour lesquelles ***P*** est colinéaire à ***E***. L'énergie de polarisation est $\frac{1}{2}\,{}^{t}\boldsymbol{E}.\mathrm{M}.\boldsymbol{E}$

7– Optique matricielle

On définit en chaque point d'un axe optique le couple (*y*,*u*) où *y* est la distance du rayon à l'axe et *u* l'angle du rayon lumineux et de l'axe.



Entre deux points, on a en première approximation une transformation linéaire de (*y*,*u*) en (*y*',*u*') due à la présence d'appareils optiques tels qu'un système de lentielle. Cette approximation est obtenue en confondant *u*, sin(*u*) et tan(*u*) (autrement dit, on considère que la direction du rayon lumineux est très proche de celle de l'axe [approximation de Gauss]). La matrice de cette transformation est dite matrice de transfert. Par exemple, les matrices sont :
❑ milieu transparent de longueur L : *u* est invariant, mais *y* augmente de L*u*

$$\begin{pmatrix} 1 & \mathrm{L} \\ 0 & 1 \end{pmatrix}$$

❑ dioptre sphérique de rayon R, du milieu d'indice *n* au milieu d'indice *n*' : *y* est invariant, mais le rayon change de direction :

$$\begin{pmatrix} 1 & 0 \\ \dfrac{(n-n')}{n'\mathrm{R}} & \dfrac{n}{n'} \end{pmatrix}$$

❑ dioptre plan, du milieu d'indice n au milieu d'indice n' : idem

$$\begin{pmatrix} 1 & 0 \\ 0 & \dfrac{n}{n'} \end{pmatrix}$$

obtenu comme cas particulier du dioptre shérique lorsque R tend vers l'infini.

❑ miroir sphérique de rayon R : idem

$$\begin{pmatrix} 1 & 0 \\ -\dfrac{2}{\mathrm{R}} & -1 \end{pmatrix}$$

obtenu comme cas particulier du dioptre shérique lorsque $n' = -n$.

❑ miroir plan : idem

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

obtenu comme cas particulier du dioptre plan, lorsque $n' = -n$, ou du miroir sphérique lorsque R tend vers l'infini.

Disposer successivement plusieurs dioptres ou miroirs séparés les uns des autres revient à construire un système optique dont la matrice est le produit des matrices de ses éléments.

Deux plans sont conjugués lorsque $y = 0$ correspond à $y' = 0$ et ceci, quel que soit u. Cela signifie qu'un objet disposé sur l'axe dans le premier plan aura son image sur l'axe dans le second plan. On a alors nécessairement $a_{12} = 0$. Si la matrice est $\begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$, on a $y' = ay$ de sorte que a est l'agrandissement transversal. Si $y = 0$, on a $u' = cu$ et c est l'agrandissement angulaire. Les plans conjugués sont dits principaux lorsque l'agrandissement vaut $a = 1$.

Le foyer objet est atteint lorsque $a_{22} = 0$, le foyer image, lorsque $a_{11} = 0$. Dans le premier cas, $y = 0 \Rightarrow u' = 0$, autrement dit les rayons issus d'un objet placé sur l'axe au foyer objet ressorte parallèle à l'axe. Dans le second cas, $u = 0 \Rightarrow y' = 0$, autrement dit des rayons incidents parallèles à l'axe se concentre sur l'axe au foyer image. La matrice entre les deux foyers vaut $\begin{pmatrix} 0 & -f \\ -\dfrac{1}{f'} & 0 \end{pmatrix}$.

8– Transformation de Lorentz

En relativité restreinte, on suppose qu'un repère O' se déplace à la vitesse V par rapport à O. Alors le quadrivecteur (x', y', z', ct') constitué de la position (x', y', z') de O' et du temps t' multiplié par la vitesse de la lumière c se déduit du quadrivecteur (x, y, z, ct) par une tranformation linéaire. En particulier, si les axes sont alignés, et si le déplacement se fait suivant Ox, on a :

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{1-\gamma^2}} & -\dfrac{\gamma}{\sqrt{1-\gamma^2}} \\ -\dfrac{\gamma}{\sqrt{1-\gamma^2}} & \dfrac{1}{\sqrt{1-\gamma^2}} \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}$$

où $\gamma = \frac{V}{c}$, c vitesse de la lumière. De même, le quadrivecteur impulsion-energie ($\boldsymbol{P}$, $\frac{E}{c}$) d'une masse ponctuelle m_0 se déplaçant à la vitesse $\boldsymbol{V}$ et d'énergie cinétique E se transforme avec le même opérateur, avec $\boldsymbol{P} = \frac{1}{\sqrt{1-\gamma^2}} m_0 \boldsymbol{V}$ et $\frac{E}{c} = \frac{1}{\sqrt{1-\gamma^2}} m_0 c$

On transforme de même les quadrivecteurs densité de courant ($\boldsymbol{J}$, $c\rho$) ou le quadrivecteur potentiel ($\boldsymbol{A}$, $\frac{\Phi}{c}$).

◆