UNIVERSITÉ DE RENNES 1

Pivot de Gauss et applications

Principales leçons concernées :

— 101 : groupes opérant sur un ensemble

— 106 : Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $GL(E)$, applications.

— 108 : Exemples de parties génératrices d'un groupe. Applications

— 122 : Anneaux principaux, exemples et applications

— 148 : Dimension d'un espace vectoriel, rang, exemples et applications.

— 149 : Déterminant, exemples et applications

— 150 : Polynôme d'endomorphisme en dimension finie, réduction d'un endomorphisme en dimension finie. Applications

— 162 : Systèmes d'équations linéaires, opérations élémentaires, aspects algorithmiques et conséquences théoriques

mais aussi :

— 103 : Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications

— 191 : Exemples d'utilisation de techniques d'algèbre en géométrie

I. Le théorème du pivot

On fixe un corps K. Rappelons l'énoncé du théorème du pivot, les définitions associées sont plus bas.

Théorème I.1 (Th du pivot, version opérations élémentaires)**.** *Soit $A \in \mathcal{M}_{n,p}(K)$.*

On peut effectuer une suite finie d'opérations élémentaires sur les lignes de A (sans utiliser les opérations de type dilatation) pour la rendre échelonnée.

En utilisant en plus les opérations de dilatation, on peut rendre A échelonnée et réduite.

De plus, il y a unicité de la matrice échelonnée réduite qu'on peut obtenir à partir de A par des opérations élémentaires sur les lignes (voir exercice 13).

Définition I.2 (Matrice échelonnée)**.** *On définit le pivot d'une ligne non nulle comme son premier coefficient non nul.*

Une matrice est échelonnée[a] *si elle vérifie les conditions suivantes :*

— *Si une ligne est nulle, toutes les lignes suivantes sont nulles*

— *Le pivot de chaque ligne non nulle est strictement à droite des pivots des lignes précédentes.*

Elle est échelonnée réduite si de plus, chaque pivot est un 1, et la colonne correspondante n'a que des zéros à part ce 1.

a. sous-entendu : en lignes

Remarque I.3. Si M est une matrice carrée échelonnée, elle est triangulaire supérieure (mais la réciproque est fausse). Si M est une matrice carrée échelonnée en colonnes, elle est triangulaire inférieure.

Une matrice carrée échelonnée est inversible ssi elle est triangulaire supérieure avec coefficients diagonaux non nuls ssi elle n'a pas de ligne nulle.

Pour une matrice carrée échelonnée et réduite, elle est inversible ssi c'est la matrice identité.

Définition I.4 (Opérations élémentaires)**.** *il y en a 3 types :*

- *(transvection)* $L_i \leftarrow L_i + \lambda L_j$ *(avec* $i \neq j$*,* $\lambda \in K$*) :*
- *(dilatation)* $L_i \leftarrow \alpha L_i$ *(avec* $\alpha \in K^*$*)*
- *(échange)* $L_i \leftrightarrow L_j$ *(avec* $i \neq j$*)*

Voici l'algorithme de calcul classique.

Algorithme 1 : Élimination par la méthode du pivot de Gauss

Entrée : une matrice $A = (a_{i,j})_{i \leq n, j \leq p}$

Sortie : une matrice échelonnée obtenue à partir de A par opérations sur les lignes

Initialisation : $i_0 = 1, j_0 = 1$

Tant que $i_0 < n$, $j_0 \leq p$, **faire:**

  si a_{i_0,j_0} *et tous les coefficients en dessous sont nuls* **:**

    incrémenter i_0, recommencer la boucle

  sinon:

    Si nécessaire, échanger L_{i_0} avec une ligne en dessous pour avoir $a_{i_0,j_0} \neq 0$

    Effectuer des opérations sur les lignes du type $L_i \leftarrow L_i - \lambda L_{i_0}$ pour $i > i_0$ pour annuler les coefficients sous a_{i_0,j_0}

    incrémenter i_0 et j_0 et recommencer la boucle.

Renvoyer A

Remarque I.5. On peut adapter l'algorithme pour renvoyer une matrice échelonnée *réduite* : lorsqu'on fait les éliminations $L_i \leftarrow L_i - \lambda L_{i_0}$ pour $i > i_0$, il faut aussi les faire pour $i > i_0$ afin d'annuler aussi les coefficients au-dessus de a_{i_0,j_0} ; et il faut mettre a_{i_0,j_0} à 1 par une dilatation $L_{i_0} \leftarrow \frac{1}{a_{i_0,j_0}} L_{i_0}$; et en fait, faire la dilatation avant les éliminations permet d'économiser quelques divisions...

Exercice 1.

Pour prouver le théorème du pivot ci-dessus, montrer que l'algorithme est correct en introduisant un bon invariant de boucle : quelle est la forme de la matrice à chaque itération ?

Version matricielle des opérations élémentaires. Effectuer une opération élémentaire *sur les lignes* de A revient à multiplier A *à gauche* par des matrices élémentaires, qui sont les matrices inversibles suivantes :

transvection $E_{ij}(\lambda)$, $i \neq j$

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \lambda & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

dilatation $Sc_i(\alpha)$, $\alpha \in K^*$

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

échange Sw_{ij}, $i \neq j$

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

Par exemple, effectuer sur A l'opération $L_i \leftarrow L_i + \lambda L_j$ revient à transformer A en $E_{ij}(\lambda) \times A$.

De même (et ça se voit en passant à la transposée), faire une opération élémentaire sur les colonnes de A revient à la multiplier à *à droite* par une matrice élémentaire.

Remarque I.6 (mnémo-technique). Pour se souvenir que la multiplication à gauche correspond aux opérations sur les lignes, on peut se souvenir qu'on utilise les opérations sur les lignes quand on résout un système (ce qui revient à déterminer le noyau d'une matrice), et les matrices A et EA ont le même noyau (lorsque E est inversible).

Remarque I.7 (Terminologie (transvections)). J'appelle les $E_{ij}(\lambda)$ les *matrices élémentaires de transvection* pour éviter de confondre avec la notion de *transvection* tout court (voir [Per95, Chap IV] proposition-definition 2.2) qui est plus générale [1]. Toute matrice élémentaire de transvection est une transvection mais pas réciproquement. Par contre, toutes les transvections (et donc aussi toutes les matrices élémentaires de transvections) sont conjuguées dans $GL_n(K)$ (et même dans $SL_n(K)$ pour $n \geq 3$, voir [Per95, Chap IV, Th 2.17]). Dans la littérature, selon les auteurs, le mot transvection peut signifier l'une ou l'autre de ces deux notions.

Théorème I.8 (Th du pivot, version matricielle)**.** *Soit $A \in \mathcal{M}_{n,p}(K)$.*

Il existe une suite finie de matrices élémentaires $E_1, \ldots, E_k$ et une matrice échelonnée réduite Ech telles que $E_k \ldots E_1 A = Ech$.

De plus, Ech est unique (pas les E_i, ni leur produit $E_k \ldots E_1$).

II. Dimension

Le théorème de la dimension dit que 2 bases d'un même espace vectoriel ont le même cardinal, et permet donc de définir correctement la dimension. On peut le démontrer facilement à partir du thm du pivot. Noter que la question (a) de l'exercice est une des rares énoncés où on peut dire quelque chose sur un système à partir du nombre d'équations et d'inconnues (Cf l'exercice vrai-faux ci-dessous).

Exercice 2. Application au théorème de la dimension

(a) En utilisant le pivot de Gauss, montrer qu'un système linéaire homogène de n équations à p inconnues avec $p > n$ admet au moins une solution non nulle.

(b) En déduire que deux bases (finies) d'un même espace vectoriel ont le même cardinal, (et donc que la dimension d'un espace vectoriel est bien définie).

III. Résoudre un système, représentation cartésienne et paramétrique d'un sous-espace

Référence : Grifone [Gri15], [CG17, §IV.3]

Préliminaire : que signifie résoudre ? Quand un énoncé demande de *résoudre* un système, qu'est-ce que ça signifie ? Quand le système a une unique solution, c'est clair : on veut déterminer cette solution. Mais quand le système a une infinité de solutions, c'est pas évident : il faut un moyen de décrire cette infinité de solutions.

Pour fixer les idées, disons qu'on a un système linéaire qu'on écrit matriciellement sous la forme $AX = b$ où X est le vecteur colonne des inconnues et b le second membre. Voici

1. par définition, une transvection est un endomorphisme de la forme $x \mapsto x + f(x)a$, avec f une forme linéaire non nulle, et $a \in \ker f \setminus \{0\}$

une telle description (absolument stupide ! !) de l'ensemble des solutions du système : c'est l'ensemble des X tels que $AX = b$! C'est parfaitement correct, mais ça n'apporte rien.

Alors que veut-on dire quand on demande de résoudre ce système ? Par exemple, si on cherche l'ensemble des $(x_1, \dots, x_5) \in K^5$ tels que

$$(*) \quad \begin{cases} x_1 & +2x_2 & & +3x_4 & +4x_5 = 0 \\ & & x_3 & +5x_4 & +6x_5 = 0 \end{cases}$$

(qui est déja sous forme échelonnée réduite), une description satisfaisante de l'ensemble F des solutions est

$$(**) \quad F = \left\{ \begin{bmatrix} -2t_1 - 3t_2 - 4t_3 \\ t_1 \\ -5t_2 - 6t_3 \\ t_2 \\ t_3 \end{bmatrix} \mid (t_1, t_2, t_3) \in \mathbb{K}^3 \right\} = \text{Vect} \left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ -6 \\ 0 \\ 1 \end{bmatrix} \right).$$

C'est un *paramétrage* de l'ensemble des solutions : chaque valeur de $t \in K$ produit une solution, et ce paramétrage est injectif : deux valeurs de (t_1, t_2, t_3) différentes produisent deux solutions différentes. De manière équivalente, on a trouvé une base de l'ensemble des solutions donnée par les 3 vecteurs ci-dessus.

En fait, on a deux façons de représenter un sous-espace $F \subset K^n$

— par un système d'équations cartésiennes, comme dans (*)

— ou un paramétrage comme dans (**).

Le système d'équation cartésiennes est pratique pour savoir si un point donné est dans F. Le paramétrage est plutôt une machine à produire des points de F.

Donc implicitement, *résoudre* un système, c'est passer d'une description de F par un système d'équations cartésiennes à une représentation paramétrique.

Passage d'une représentation cartésienne à une représentation paramétrique Si on définit notre sous-espace par $F = \ker A$, c'est une représentation de F par un système d'équation cartésiennes. Si on définit $F = \text{Im}B$, c'est une représentation paramétrique de F (injective si $\ker B = 0$). Si on définit $F = \text{Vect}(v_1, \dots, v_p)$, c'est une représentation paramétrique de F (injective si $(v_1, \dots, v_p)$ est libre).

Le système (*) ci-dessus définit F sous la forme $F = \ker A$, avec A la matrice 2×5 du système, et lorsqu'on le résoud, on écrit $F = \text{Im}B$ avec

$$B = \begin{bmatrix} -2 & -3 & -4 \\ 1 & 0 & 0 \\ 0 & -5 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

et les colonnes de B forment une base de F (le fait que la famille est libre est une conséquence immédiate du fait que la sous-matrice de B formée par les lignes 2,4,5 [2] est égale à I_3). [3]

Le pivot de Gauss permet donc de passer d'une représentation cartésienne à une représentation paramétrique injective de F, ou de manière équivalente, partant de la description $F = \ker A$ de trouver une base de F.

2. correspondant aux variables libres du système : x_2, x_4, x_5.

3. Notons que B a une forme particulière : elle n'est pas échelonnée en colonnes, mais si on définit les pivots des colonnes comme étant les premiers coefficients non nuls en partant du bas (ici ce sont des 1), le pivot de chaque colonne est au-dessous de la suivante : c'est ce que les auteurs appellent une matrice *co-échelonnée* dans [CG17, Chap IV] (mais leur rédaction de la définition 3.2.1 est assez approximative...)

Exercice 3.

Voir par exemple [CG17, Chap IV,§3].

Soit $A\mathcal{M}_{np}(K)$ une matrice échelonnée, ayant r lignes non nulles, et $n-r$ lignes nulles.

(a) Montrer que $\ker A$ est de dimension $n-r$.

(b) Si $A = \begin{bmatrix} 1 & 0 & a & b & c & d \\ 0 & 1 & a' & b' & c' & d' \end{bmatrix}$, donner une base de $\ker A$.

(c) Généraliser au cas où $A \in \mathcal{M}_{np}(K)$ est une matrice échelonnée réduite dont les pivots sont dans les r premières colonnes (c'est à dire $A = \left[\begin{smallmatrix} I_r & C \\ 0 & 0 \end{smallmatrix}\right]$) : on explicitera une matrice M dont les vecteurs forment une base de $\ker A$. Sans supposer que les pivots sont dans les r premières colonnes, montrer qu'il existe une matrice de permutation $Q \in \mathcal{M}_p(K)$ telle que AQ devienne comme ci-dessus, et déduire une base de $\ker A$.

Passage d'une représentation paramétrique à une représentation cartésienne. L'algorithme du pivot permet aussi de passer d'une représentation paramétrique à une représentation cartésienne d'un sous-espace. C'est basé sur la remarque suivante :

Exercice 4.

Soit $A\mathcal{M}_{np}(K)$ une matrice échelonnée, ayant r lignes non nulles, et $n-r$ lignes nulles.

Montrer que le système $AX = b$ a des solutions ssi les $n-r$ dernières coordonnées de b sont nulles.

Faisons le sur un cas particulier.

Exercice 5. Passage de paramétrique à cartésien

Soient $v_1 = (1, 3, -2, 2, 3)$, $v_2 = (1, 4, -3, 4, 2)$ et $v_3 = (2, 3, -1, -2, -9)$. Déterminer un système d'équations cartésiennes pour le sous-espace de $\mathbb{R}^5$ défini par $F = \text{Vect}(v_1, v_2, v_3)$.

Remarque III.1. Résoudre un système avec un second membre générique comme ci-dessus peut s'écrire matriciellement avec un second membre matriciel, initialement égal à I_n. Par exemple, la matrice du système ci-dessus peut s'écrire

$$\left[\begin{array}{ccc|ccccc} 1 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 3 & 4 & 3 & 0 & 1 & 0 & 0 & 0 \\ -2 & -3 & -1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 4 & -2 & 0 & 0 & 0 & 1 & 0 \\ 3 & 2 & -9 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Et les matrices de passage ? Gauss-Jordan. Lorsqu'on échelonne une matrice A, on a parfois besoin de calculer une matrice P telle que $PA = Ech$. Une méthode pratique consiste à appliquer sur la matrice I_n les mêmes opérations que sur A : à la fin, la matrice I_n a été transformée en $P.I_n = P$. Il y a une manière commode, c'est d'écrire I_n à droite de A [4] et de faire les opérations sur les lignes de la matrice augmentée obtenue.

Si on part d'une matrice inversible $A \in GL_n(K)$, la matrice échelonnée réduite qu'on va obtenir est la matrice identité, donc puisque $PA = I_n$, on obtient $P = A^{-1}$. C'est la fameuse *méthode de Gauss-Jordan* pour calculer l'inverse de A.

4. ce qui revient à ajouter un second membre générique comme dans la remarque ci-dessus

Exercice 6.

Voir [Gri15, §2.3 et §2.4].

Expliquer comment on peut, avec l'algorithme de Gauss, résoudre les problèmes suivants [a] :

(a) Etant donnée une famille de vecteurs $v_1, \dots v_p \in K^n$, déterminer si elle est libre ou non ; déterminer son rang

(b) Etant donnée une famille de vecteurs $v_1, \dots v_p \in K^n$, déterminer un système d'équations cartésiennes de $\mathrm{Vect}(v_1, \dots, v_p)$

(c) Etant donnée une famille de vecteurs $v_1, \dots v_p \in K^n$, en extraire une base de $\mathrm{Vect}(v_1, \dots, v_p)$

(d) Etant donnée une famille de vecteurs $v_1, \dots v_p \in K^n$, déterminer si elle est génératrice

(e) Etant donnée une famille libre de vecteurs $v_1, \dots v_p \in K^n$, la compléter en une base de K^n ; déterminer un supplémentaire

(f) Etant donnés $F = \mathrm{Vect}(v_1, \dots, v_p)$ et $G = \mathrm{Vect}(w_1, \dots, w_q)$, déterminer une base de $F + G$

(g) Etant donnée une matrice A, déterminer une base de $\mathrm{Im}A$

(h) Etant donnée une matrice A, déterminer une base de $\ker A$

(i) Etant données deux matrices A, B, déterminer une base de $\ker A \cap \ker B$

(j) Etant données deux matrices A, B, déterminer une base de $\mathrm{Im}A \cap \mathrm{Im}B$

(k) Etant données deux matrices A, B, déterminer une base de $\ker A \cap \mathrm{Im}B$

a. On suppose qu'on peut faire des calculs exacts ; par exemple $K = \mathbb{Q}$ ou un corps fini, (et pas $\mathbb{R}$ pour lequel on travaillerait avec des flottants, sujets aux erreurs d'arrondi).

IV. Générateurs, connexité

Exercice 7. Application : générateurs de $GL_n(K)$.

Soit K un corps. Le but de l'exercice est de montrer le théorème suivant (voir exercice B.3 [CG17, Chap IV]) :

Théorème.

1. *$GL_n(K)$ est engendré par les matrices élémentaires de transvection et de dilatation*
2. *$SL_n(K)$ est engendré par les matrices élémentaires de transvection*

(a) Ecrire une variante l'algorithme d'élimination de Gauss qui suppose la matrice d'entrée A inversible, et qui produit en sortie une matrice diagonale, sans utiliser d'opération de dilatation, et en remplaçant les opérations d'échange de lignes $L_i \leftrightarrow L_j$ (nécessaire quand un pivot a_{ii} est nul) par une opération $L_i \leftarrow L_i + L_j$ avec $j > i$.

(b) Déduire que pour toute matrice A inversible, il existe une matrice diagonale D et des matrices élémentaires de transvection E_i telles que $E_1 \dots E_k A = D$, et déduire le point 1 du théorème.

(c) Montrer que pour tout $\lambda \in K^*$, la matrice $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix}$ est un produit de matrices élémentaires de transvection.

(d) En déduire que dans la question (b), on peut ajouter que $D = diag(1, 1, \ldots, 1, a)$ pour un certain $a \in K^*$.

(e) En déduire le point 2 du théorème : $SL_n(K)$ est engendré par les matrices élémentaires de transvection.

Autre preuve du point 1, en utilisant des relations entre les matrices élémentaires.

(f) En utilisant la version matricielle du pivot de Gauss, montrer que toute matrice $A \in GL_n(K)$ est un produit de matrices élémentaires

(g) Remplacement des opérations de permutation : vérifier la relation

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

et en déduire que $GL_n(K)$ est engendré par les matrices élémentaires de transvection et de dilatation.

(h) En utilisant la relation de la question (g) et sa généralisation à n'importe quelle matrice $Sw_{i,j}$, déduire le point 1 du théorème.

Exercice 8.

Montrer que pour tout $n \geq 2$, il existe une constante $C \in \mathbb{N}$ telle que toute matrice de $SL_n(K)$ est un produit d'au plus C matrices élémentaires de transvection.

Exercice 9. Connexités des groupes linéaires sur $\mathbb{R}$ et $\mathbb{C}$

On note $GL_n^+(\mathbb{R})$ le groupe des matrices de déterminant positif (c'est à dire préservant l'orientation). Le but de l'exercice est de montrer le théorème suivant :

Théorème.

1. *$GL_n(\mathbb{C})$ et $SL_n(\mathbb{C})$ sont connexes par arcs.*
2. *$GL_n(\mathbb{R})$ n'est pas connexe mais $GL_n^+(\mathbb{R})$ et $SL_n(\mathbb{R})$ sont connexes par arcs.*

Commençons par $GL_n(\mathbb{R})$

(a) Montrer que $GL_n(\mathbb{R})$ n'est pas connexe.

Continuons avec $SL_n(\mathbb{R})$ et $SL_n(\mathbb{C})$.

(b) Pour chaque matrice élémentaire de transvection $E = E_{ij}(\alpha)$ (avec $\alpha \in \mathbb{R}$ ou $\alpha \in \mathbb{C}$) exhiber un chemin joignant I_n à E dans $SL_n(\mathbb{R})$ ou $SL_n(\mathbb{C})$ selon le cas.

(c) En utilisant que $SL_n(K)$ est engendré par ses matrices élémentaires de transvection, déduire que $SL_n(\mathbb{R})$ et $SL_n(\mathbb{C})$ sont connexes par arcs.

On montre maintenant la connexité par arcs de $GL_n^+(\mathbb{R})$ et $GL_n(\mathbb{C})$ en utilisant celle de $SL_n(\mathbb{R})$ et $SL_n(\mathbb{C})$.

(d) Etant donnée $A \in GL_n^+(\mathbb{R})$, exhiber un chemin contenu dans $GL_n^+(\mathbb{R})$ et reliant A à une matrice dans $SL_n(\mathbb{R})$, et conclure.

(e) Faire de même dans $GL_n(\mathbb{C})$.

Il y a une autre preuve, plus facile, du fait que $SL_n(\mathbb{C})$ et $GL_n(\mathbb{C})$ sont connexes par arcs. Cette preuve n'utilise pas les générateurs mais ne fonctionne pas sur $\mathbb{R}$.

Exercice 10. Connexité par arcs de $SL_n(\mathbb{C})$ et $GL_n(\mathbb{C})$, autre preuve

Commençons par la connexité de $GL_n(\mathbb{C})$.

(a) Soit $A \in GL_n(\mathbb{C})$, et soit $F : \mathbb{C} \to \mathcal{M}_n(\mathbb{C})$ définie par $F(t) = tA + (1-t)I_n$. Montrer qu'il existe un ensemble fini $Z \subset \mathbb{C}$ tel que $F(t)$ est une matrice inversible pour tout $t \in \mathbb{C} \setminus Z$.

(b) En prenant un chemin dans $\mathbb{C} \setminus Z$ allant de 0 à 1, trouver un chemin reliant I_n à A dans $GL_n(\mathbb{C})$.

Montrons maintenant que $SL_n(\mathbb{C})$ est connexe par arcs.

(c) Soit $A \in SL_n(\mathbb{C})$. A partir de la fonction F ci-dessus, définir une fonction $G : \mathbb{C} \setminus Z \to SL_n(\mathbb{C})$ telle que $G(0) = I_n$ et $G(1) = A$ et conclure comme au-dessus.

Orientation d'un $\mathbb{R}$-espace vectoriel. L'orientation en dimension 2, on peut voir ça comme le sens dans lequel on tourne : sens trigo vs sens horaire. En dimension 3, c'est déjà un peu plus difficile à expliquer : règle de la main droite, du tire-bouchon... Et en dimension n ? Pourquoi peut-on parler d'orientation ? De quoi s'agit-il ?

On peut bien sûr répondre en termes de déterminant, c'est la définition habituelle :

Définition. *2 bases $\mathcal{B}, \mathcal{B}'$ d'un $\mathbb{R}$-espace vectoriel ont la même orientation si la matrice de passage entre $\mathcal{B}$ et $\mathcal{B}'$ est de déterminant positif.*

Cette définition a l'avantage de la simplicité et de la concision. Mais ça ne rend pas la chose très concrete, et n'explique pas son importance. Une autre réponse est que l'ensemble des bases de $\mathbb{R}^n$ a exactement 2 composantes connexes. Si 2 bases $\mathcal{B} = (v_1, \dots, v_n)$ et $\mathcal{B}' = (v'_1, \dots, v'_n)$ ont la même orientation, on peut passer continûment de l'une à l'autre en préservant à chaque instant le fait d'être une base à chaque instant ; plus précisément, il existe n applications continues $p_i : [0,1] \to \mathbb{R}^n$ telles que $(v_1, \dots, v_n) = (p_1(0), \dots, p_n(0))$, $(v'_1, \dots, v'_n) = (p_1(1), \dots, p_n(1))$, et telles qu'à chaque instant t, la famille $(p_1(t), \dots, p_n(t))$ soit une base.

Pour commencer, convainquez vous en bougeant vos doigts (ou autrement !) que c'est bien le cas en dimension 2 et 3 !

Exercice 11.

Démontrer la caractérisation suivante de l'orientation.

Théorème IV.1. *Soient $\mathcal{B} = (v_1, \dots, v_n)$, et $\mathcal{B}' = (v'_1, \dots, v'_n)$ deux bases de $\mathbb{R}^n$.*

Alors $\mathcal{B}$ et $\mathcal{B}'$ ont la même orientation si et seulement si il existe un chemin de bases reliant $\mathcal{B}$ à $\mathcal{B}'$, c'est à dire qu'il existe des applications continues $p_1, \dots, p_n : [0,1] \to E$ tels que

- *$p_1(0) = v_1, \dots, p_n(0) = v_n$, et $p_1(1) = v'_1, \dots, p_n(1) = v'_n$,*
- *pour tout $t \in [0,1]$, $(p_1(t), \dots, p_n(t))$ est une base de E.*

V. Matrices équivalentes

Rappelons que deux matrices $A, B \in \mathcal{M}_{np}(K)$ sont *équivalentes* si il existe $P \in GL_n(K)$ et $Q \in GL_p(K)$ tq $B = PAQ$. [5]

Puisque $GL_n(K)$ et $GL_p(K)$ sont engendrés par les matrices élémentaires, on obtient que A et B sont équivalentes ssi on peut passer de l'une à l'autre par une suite finie d'opérations élémentaires sur les lignes et les colonnes (il faut s'autoriser les 2 !).

Par une variante de l'algorithme du pivot, il est facile de voir qu'on peut transformer toute matrice A en une matrice par blocs de la forme $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ (ayant n lignes et p colonnes).

On rappelle que le rang d'une matrice A est le rang de l'image de A, c'est à dire la dimension de l'espace vectoriel engendré par ses vecteurs colonnes.

Théorème V.1. *Etant données $A, B \in \mathcal{M}_{np}(K)$ les énoncés suivant sont équivalents :*

(a) A et B sont équivalentes

(b) A et B ont le même rang

(c) A et B peuvent être transformée en la même matrice $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ par des opérations élémentaires

La preuve est facile à partir de la variante de l'algorithme du pivot mentionnée ci-dessus :

Exercice 12.

(a) Montrer que deux matrices équivalentes ont le même rang.

(b) L'algorithme de Gauss permet de transformer toute matrice $A \in \mathcal{M}_{np}(K)$ avec des opérations élémentaires sur les lignes et les colonnes en une matrice $\left[\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Montrer qu'alors $r = \mathrm{rg}(A)$ et en déduire que deux matrices de $\mathcal{M}_{np}(K)$ ayant le même rang sont équivalentes.

(c) Déduire aussi que $\mathrm{rg}(A) = \mathrm{rg}(A^t)$.

On peut définir l'équivalence à gauche (ou à droite) en n'autorisant la multiplication que d'un côté :

Définition. *On dit que deux matrices $A, B \in \mathcal{M}_{np}(K)$ sont* équivalentes à gauche *si il existe $P \in GL_n(K)$ tq $B = PA$.*

Elles sont équivalentes à droite *si il existe $Q \in GL_p(K)$ tq $B = AQ$.*

On peut caractériser l'équivalence à gauche des matrices de la façon suivante.

Exercice 13. Matrices équivalentes à gauche

Voir par ex Th 2.3.1 de [CG17, Chap IV].

5. On pourrait dire de manière équivalente (et c'est un point de vue important !) : A et B sont dans la même orbite pour l'action de $GL_n(K) \times GL_p(K)$ sur $\mathcal{M}_{np}(K)$ définie par $(P, Q) \cdot M = PMQ^{-1}$ pour $(P, Q) \in GL_n(K) \times GL_p(K)$ et $M \in \mathcal{M}_{np}(K)$.

Théorème. *(a) Deux matrices $A, B \in \mathcal{M}_{np}(K)$ sont équivalentes à gauche ssi elles ont le même noyau.*

(b) Toute matrice est équivalente à gauche à une matrice échelonnée réduite.

(c) Deux matrices échelonnées réduites $Ech, Ech' \in \mathcal{M}_{np}(K)$ qui sont équivalentes à gauche sont égales.

Pour le point (c), il est plus pratique de travailler sur les matrices échelonnées en colonnes et démontrer l'énoncé suivant équivalent :

(c') Deux matrices échelonnées réduites en colonnes $M, M' \in \mathcal{M}_{pn}(K)$ qui sont équivalentes à droite sont égales.

On pourra commencer par montrer que M et M' ont les même pivots en utilisant que $\mathrm{Im}(M) = \mathrm{Im}(M')$ et en montrant l'énoncé suivant :

(d) Soit $M \in \mathcal{M}_{pn}(K)$ une matrice échelonnée en colonnes et $F = \mathrm{Im}(M)$. Pour $i \leq n$, soit V_i l'ensemble des vecteurs dont les i premieres coordonnées sont nulles, c'est à dire $V_i = Vect(e_{i+1}, e_{i+2}, \ldots, e_n)$. Alors pour tout $i \leq n$, $\dim F \cap V_i$ est égal au nombre de pivots strictement au-dessous de la ligne i.

Il y a une version pour l'équivalence à droite :

Exercice 14. Matrices équivalentes à droite

En utilisant l'exo précédent, démontrer le théorème suivant (voir par ex Th 2.4.1 [CG17, Chap IV])

Théorème. *(a) Deux matrices $A, B \in \mathcal{M}_{np}(K)$ sont équivalentes à droite ssi elles ont la même image.*

(b) Toute matrice est équivalente à droite à une matrice échelonnée réduite en colonnes.

(c) Deux matrices échelonnées réduites en colonnes $Ech, Ech' \in \mathcal{M}_{np}(K)$ qui sont équivalentes à droite sont égales.

Exercice 15.

(a) Donner un algorithme [a] qui, étant donné deux matrices $A, B \in \mathcal{M}_{np}(K)$ détermine si A et B sont équivalentes

(b) Donner un algorithme qui, étant donné deux matrices $A, B \in \mathcal{M}_{np}(K)$ détermine si A et B sont équivalentes à droite

(c) même question pour l'équivalence à gauche

a. On suppose ici qu'on sait calculer de manière exacte dans K. Par exemple, $K = \mathbb{Q}$, ou K un corps fini. On ne travaille donc pas avec des flottants.

VI. Déterminant et mineurs

Définition (Déterminant d'une famille de vecteurs de K^n, d'une matrice)**.** *On définit*

le déterminant de n vecteurs $v_1, \dots, v_n \in K^n$ par la formule

$$\det(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) v_{1,\sigma(1)} \dots v_{n,\sigma(n)}$$

$(v_{i,1}, \dots, v_{i,n})$ sont les coordonnées de v_i dans la base canonique. La somme est sur toutes les permutations $\sigma \in S_n$, et $\varepsilon(\sigma) \in \{+1, -1\}$ est la signature de σ.

On définit le déterminant d'une matrice $A \in \mathcal{M}_n(K)$ comme le déterminant de ses vecteurs colonnes.

On peut utiliser le pivot de Gauss (ou les opérations élémentaires sur les matrices) dans la théorie du déterminant.

Une application classique est le fait qu'on peut calculer rapidement le déterminant (en $O(n^3)$ opérations en échelonnant la matrice), ce qui n'est pas du tout évident à partir de la formule de définition.

Remarque VI.1. On définit le *permanent* d'une matrice A par la même formule que le déterminant, mais en enlevant la signature $\varepsilon(\sigma)$ devant le produit :

$$perm(A) = \sum_{\sigma \in S_n} a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.$$

On ne connaît pas d'algorithme permettant de calculer le permanent d'une matrice en temps polynomial en n (et c'est un problème ouvert de savoir s'il existe un tel algorithme). Le permanent est donc beaucoup plus compliqué à calculer que le déterminant ! C'est donc un petit miracle que le calcul du déterminant puisse se faire de manière efficace. Ceci montre bien que les propriétés algébriques du déterminant en lien avec les opérations élémentaires sont cruciales pour son calcul.

Dans l'exercice suivant, on utilise le pivot de Gauss pour démontrer les propriétés fondamentales du déterminant (à partir de rien) : l'unicité du déterminant, et la formule $\det(AB) = \det(A)\det(B)$.

Exercice 16. Unicité et multiplicativité du déterminant,

Le but est de démontrer le théorème suivant :

Théorème. *1. L'ensemble des formes n-linéaires alternées est de dimension 1 et pour toute forme linéaire alternée f, on a $f(A) = \det(A) f(I_n)$ pour tout $A \in \mathcal{M}_n(K)$*

2. Pour toutes matrices $A, B \in \mathcal{M}_n(K)$, $\det(AB) = \det(A)\det(B)$.

(a) Vérifier que le déterminant d'une matrice triangulaire supérieure (ou inférieure) est égal au produits des coefficients diagonaux et en déduire le déterminant des matrices élémentaires.

Vérifier que det est une forme n-linéaire alternée (si deux vecteurs colonnes sont égaux, le déterminant est nul).

(b) Montrer que si f est une forme n-linéaire alternée, l'effet sur la valeur de f des opérations élémentaires sur les colonnes d'une matrice est le suivant : si A' est obtenue à partir de A en effectuant une opération du type $C_i \leftarrow C_i + \alpha C_j$, $C_i \leftarrow \lambda C_i$, ou $C_i \leftrightarrow C_j$, alors $f(A')$ est égal à $f(A)$, $\lambda f(A)$, $-f(A)$ selon le cas. En déduire que $f(AE) = f(A)\det(E)$ pour tout $A \in \mathcal{M}_n(K)$ et toute matrice élémentaire E. Noter en particulier que $\det(AE) = \det(A)\det(E)$.

(c) Cas des matrices inversibles : en utilisant que $GL_n(K)$ est engendré par les matrices élémentaires, déduire les points 1 et 2 du théorème pour des matrices A et B inversibles.

(d) Si A n'est pas inversible, remarquer que la matrice échelonnée en colonnes obtenue par le pivot de Gauss en colonnes possède une colonne nulle. En déduire que $f(A) = 0$ et en particulier que $\det(A) = 0$.

En déduire que la formule $f(A) = \det(A) f(I_n)$ est encore valable. Puisque AB n'est pas inversible si A ou B n'est pas inversible, vérifier que $\det(AB) = \det(A)\det(B)$.

La formule $\det(AB) = \det(A)\det(B)$ implique qu'une matrice est inversible ssi son déterminant est non nul.

On appelle mineur d'une matrice A le déterminant d'une sous-matrice carrée $A_{I,J}$, définie en choisissant un ensemble $I \subset \{1, \dots, n\}$ de k lignes et un ensemble $J \subset \{1, \dots, p\}$ de k colonnes de A.

Exercice 17. Rang et mineurs

Le but de l'exercice est de démontrer le théorème suivant :

Théorème. *Le rang d'une matrice $A \in \mathcal{M}_{np}(K)$ est la taille maximale d'un mineur non nul.*

(a) Vérifier que le rang d'une matrice échelonnée en colonnes est égal au nombre de colonnes non nulles, et que le théorème est vrai pour une matrice échelonnée en colonnes.

(b) Soit $A' = A$ obtenue à partir de A en effectuant une opération élémentaire sur les colonnes. Montrer que tous les mineurs de taille k de A' sont nuls ssi tous les mineurs de taille k de A sont nuls.

Rang, topologie et flottants

Exercice 18.

Voir [CG17, Chap I.] proposition 4.1.

On se place sur $K = \mathbb{R}$ ou $\mathbb{C}$. On fixe $n, p \geq 1$. On note $RG_{\leq r}$ et $RG_{=r}$ les sous-ensembles de $\mathcal{M}_{np}(K)$ formés des matrices de rang $\leq r$ et de rang $= r$ respectivement.

(a) Montrer que l'ensemble $RG_{\leq r}$ est un fermé de $\mathcal{M}_{np}(K)$.

(b) Montrer que l'ensemble des matrices de rang maximal (i.e. de rang égal à $\min(n, p)$) est un ouvert dense dans $\mathcal{M}_{np}(K)$; montrer que si $r < \min(n, p)$, $RG_{\leq r}$ est d'intérieur vide.

(c) Soit $r \leq \min(n, p)$. Montrer que $\overline{RG_{=r}} = RG_{\leq r}$: l'adhérence des matrices de rang exactement r est l'ensemble des matrices de rang $\leq r$.

Exercice 19.

Expliquer pourquoi chercher le rang d'une matrice donnée par des nombres flottants n'a pas grand sens.

Exercice 20.

On utilise les notations de l'exo précédent.

(a) Montrer que sur $K = \mathbb{C}$, $RG_{=r}$ et $RG_{\leq r}$ sont connexes pour tout $r \leq \min(n, p)$.

(b) Montrer que sur $K = \mathbb{R}$, $RG_{=r}$ et $RG_{\leq r}$ sont connexes sauf $RG_{=r}$ dans le cas $r = n = p$ (correspondant aux matrices inversibles).

Décomposition LU. On peut aussi appliquer ce type de raisonnement à l'existence de la décomposition LU. Voir [Sch98] ou [Cia98]. Une décomposition LU de A est une factorisation $A = LU$ avec U triangulaire supérieure inversible, et L triangulaire inférieure avec coefficients diagonaux égaux à 1.

Une telle factorisation est ce qu'on obtient lorsqu'on applique l'algorithme de Gauss sans jamais avoir besoin de faire des échanges de lignes (et que, suivant l'algorithme classique, on n'utilise que des opérations de la forme $L_i \leftarrow L_i + \lambda L_j$ avec $i > j$). La matrice échelonnée obtenue est la matrice triangulaire U, et L^{-1} est le produit des matrices élémentaires de transvections utilisées pour modifier A (elles sont toutes triangulaires inférieures avec des 1 sur la diagonale !).

Il y a aussi un fait remarquable : la matrice L se trouve sans calcul : pour $j > i$, le coefficient L_{ij} de L est juste le coefficient λ utilisé lors de l'opération $L_j \leftarrow L_j + \lambda L_i$ [Sch98, I. Lemme 4.3].

Exercice 21. Décomposition LU

Soit $A \in \mathcal{M}_n(K)$ une matrice inversible.

(a) Si A admet une décomposition LU, alors L et U sont uniques. Est-ce encore le cas si A n'est pas inversible ?

(b) A admet une décomposition LU ssi pour tout $k \leq n$, le mineur $\det(\Delta_k)$ est non nul où

$$\Delta_k = \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix}.$$

Remarque VI.2. La cas où A est une matrice définie positive un cas naturel qui vérifie les hypothèses.

En général, lorsqu'on travaille avec des nombres flottants, on peut avoir intérêt à faire un échange de lignes pour choisir un pivot le plus grand possible. En effet, pour la norme subordonnée à la norme 1, le conditionnement [6] de la matrice $E = E_{i,j}(\lambda)$ est égal à $(1+|\lambda|)^2$ (voir [Cia98, Thm 1.4.2]). Faire une élimination avec un λ grand (ce qui correspond à diviser par un pivot petit) risque donc d'amplifier les erreurs d'arrondi. En effet, on sait que le conditionnement de la nouvelle matrice $A' = EA$ vérifie $\mathrm{Cond}(A') \leq \mathrm{Cond}(E) \times \mathrm{Cond}(A) = (1 + |\lambda|)^2 \mathrm{Cond}(A)$, donc on a intérêt à faire en sorte que λ soit petit pour que le système intermédiaire (de matrice A') reste bien conditionné.

Remarque VI.3. Dans les cas où la décomposition LU n'existe pas, on peut utiliser une décomposition $A = LUP$ ou $A = PLU$ qui existent pour toute matrice inversible A (par contre, pas d'unicité).

VII. Complexité algorithmique

Voir [Sch98, Chap 2.5] ou [Cia98]. Le nombre d'opérations arithmétiques (additions, multiplications, divisions) nécessaire pour calculer la décomposition LU, ou pour échelonner

6. Le conditionnement d'une matrice A est égal à $||A|| \times ||A^{-1}||$. Ce nombre contrôle l'amplification des erreurs associée à la résolution d'un système linéaire de matrice A, voir [Cia98, Thm 2.2.1 et 2.2.2].

une matrice carrée $n \times n$ ou pour résoudre un système de Cramer de n équations à n inconnues est équivalent à $2n^3/3$ quand $n \to \infty$.

Rq : une fois qu'on a calculé une décomposition $A = LU$, résoudre deux systèmes triangulaires est très rapide, le nombre d'opérations nécessaires est équivalent à $2n^2$ (donc négligeable par rapport à $2n^3/3$).

Remarque VII.1. Comme l'explique bien [Sch98] on pourrait penser que le calcul de A^{-1} permet d'accélérer la résolution du système si on a plein de second membres à traiter. Mais ce n'est pas avantageux par rapport au calcul de la décomposition LU : le calcul de $A^{-1}.b$ nécessite un nombre d'opérations équivalent à $2n^2$, c'est pas mieux que la résolution des deux systèmes triangulaires.

De manière peut-être surprenante, on peut faire (théoriquement) mieux en utilisant l'algorithme de multiplication de Matrices de Strassen. Voir [AHU74], chapitre 6.

Commençons par le produit de matrices carrées $n \times n$. Si $C = AB$, la formule standard $c_{ij} = \sum_k a_{ik} b_{kj}$ demande n multiplications et $n-1$ additions pour chaque coefficients, soit $O(n^3)$ opérations arithmétiques. Il se trouve qu'on peut faire mieux.

On découpe A, B et C en blocs de taille $n/2$:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \times \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

Le calcul par blocs correspondant au calcul standard est le suivant : $C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j}$ pour $i, j \in \{1, 2\}$. Le calcul utilise donc en tout 8 multiplications de sous-matrices, deux pour chaque bloc C_{ij} (et 4 additions, une pour chaque bloc).

Il se trouve qu'on peut économiser une multiplications quitte à regrouper les calculs astucieusement (et à faire des additions en plus).

Exercice 22. Algorithme de Strassen pour la multiplication de matrices

Soient $A, B \in \mathcal{M}_n(K)$ où n est une puissance de 2, et $C = AB$.

On découpe A, B et C en blocs de taille $n/2$:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \times \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

On pose

$$\begin{aligned} M_1 &= (A_{12} - A_{22})(B_{21} + B_{22}) \\ M_2 &= (A_{11} + A_{22})(B_{11} + B_{22}) \\ M_3 &= (A_{11} - A_{21})(B_{11} + B_{12}) \\ M_4 &= (A_{11} + A_{12})B_{22} \\ M_5 &= A_{11}(B_{12} - B_{22}) \\ M_6 &= A_{22}(B_{21} - B_{11}) \\ M_7 &= (A_{21} + A_{22})B_{11} \end{aligned}$$

et on vérifie que $C_{11} = M_1 + M_2 - M_4 + M_6$, $C_{12} = M_4 + M_5$, $C_{21} = M_6 + M_7$ et $C_{22} = M_2 - M_3 + M_5 - M_7$. Donc 7 multiplications de sous-matrices (et 18 additions) permettent de calculer le produit AB.

L'algorithme de Strassen utilise cette astuce de manière récursive : pour calculer chacun des 7 produits, on découpe les matrices en 2, et on applique la même recette avec des matrices deux fois plus petites.

(a) En supposant que $n = 2^k$, montrer que le coût de cet algorithme (en nombre d'opérations arithmétiques) est $O(n^{\log_2 7}) = O(n^{2,81}) \ll n^3$.

(b) Montrer que c'est encore le cas si n n'est pas une puissance de 2 en complétant la matrice avec des zéros pour obtenir une matrice de taille 2^k.

Remarque VII.2. La version par blocs de l'algorithme standard, avec 8 multiplications de matrices de taille $n/2$, donne une complexité en $O(n^{\log_2 8}) = O(n^3)$, comme l'algorithme standard. C'est donc le fait d'avoir 7 multiplications au lieu de 8 qui permet de faire baisser l'exposant de $3 = \log_2(8)$ à $2,81 = \log_2(7)$.

On peut alors démontrer, en utilisant cet algorithme, qu'il est possible de calculer la décomposition LUP et même l'inverse d'une matrice A en utilisant $O(n^{2,81})$ opérations. Plus généralement, si on sait multiplier des matrices en $O(n^\alpha)$ avec $\alpha > 2$, alors on peut calculer l'inverse d'une matrice en $O(n^\alpha)$. Voir Th 6.5 dans [AHU74].

Attention, cet algorithme est plus efficace *asymptotiquement* que l'algorithme naïf, mais il ne devient plus rapide que pour des matrices suffisamment grandes (dépendant de l'implémentation effective des algorithmes, à partir de la taille 400 à 2000, (source : [CLRS22])). De plus cet algorithme n'est pas efficace avec des matrices creuses.

Il y a plein d'autres méthodes pour résoudre des systèmes linéaires, avec leurs avantages et leurs défauts... Voir [Cia98, Sch98].

VIII. Changement de corps

L'énoncé suivant dit que si un système linéaire à coefficients rationnels possède une solution sur $\mathbb{R}$, il en a aussi sur $\mathbb{Q}$.

Exercice 23.

Soit L un corps, et $K \subset L$ un sous-corps. Démontrer les énoncés suivants :

(a) Si un système linéaire avec second membre à coefficients dans K admet une solution dans L^n, alors il admet déjà une solution dans K^n

(b) Si un système linéaire homogène à coefficients dans K admet une solution non nulle dans L^n, alors il admet déjà une solution non nulle dans K^n

Voici des variantes de cet exercice :

Exercice 24.

Soit K un sous-corps d'un corps L.

— Une matrice $A \in \mathcal{M}_n(K)$ est inversible dans $\mathcal{M}_n(K)$ ssi elle est inversible dans $\mathcal{M}_n(L)$.

— Si $A \in \mathcal{M}_{np}(K)$, son rang ne change pas si on la considère comme une matrice dans $\mathcal{M}_{np}(L)$.

Voici une application rigolote.

Exercice 25.

On suppose qu'un rectangle est pavé par un nombre fini de carrés (dont les côtés sont parallèles à ceux du rectangle). Le but est de montrer que le rapport entre hauteur et largeur du rectangle est un nombre rationnel.

On peut supposer que la hauteur du rectangle est égale à 1, et il faut montrer que sa largeur est rationnelle. On peut supposer que le rectangle et les carrés ont leurs côtés parallèles aux axes, que les deux côtés horizontaux du rectangle sont contenus dans les droite $y = 0$ et $y = 1$.

On construit le graphe $(\mathcal{V}, \mathcal{E})$ suivant : il y a un sommet v_E pour chaque composante connexe E de la réunion des côtés horizontaux des carrés. Il y a une arête entre v_E et $v_{E'}$ pour chaque carré dont les côtés horizontaux rencontrent E et E' (on autorise donc que plusieurs arêtes joignent deux sommets). On note $\mathcal{V}$ l'ensemble des sommets et $\mathcal{E}$ l'ensemble des arêtes. Il y a deux sommets particuliers $v_0, v_1 \in \mathcal{V}$ correspondant aux deux côtés horizontaux du rectangle.

(a) Pour chaque composante E comme ci-dessus, soit y_E l'ordonnée telle que E est contenu dans la droite $y = y_E$. Soit $h : \mathcal{V} \to \mathbb{R}$ la fonction qui au sommet v_E associe y_E. Montrer que pour tout sommet $v \in \mathcal{V} \setminus \{v_0, v_1\}$,

$$\sum_{e=\{v,w\}} (h(w) - h(v)) = 0 \quad h(v_0) = 0, \quad h(v_1) = 1 \quad (*)$$

la somme étant sur les arêtes e incidentes sur v (on dit que h est harmonique sur $\mathcal{V} \setminus \{v_0, v_1\}$)

(b) En déduire qu'il existe $h' : \mathcal{V} \to \mathbb{Q}$ qui vérifie $(*)$

(c) Montrer que $(*)$ admet une unique solution $h : \mathcal{V} \to \mathbb{R}$ en montrant un principe du maximum : si h, h' sont deux solutions distinctes, on peut supposer que $h - h'$ prend une valeur strictement positive, et considérer un sommet v où $h - h'$ admet son maximum. Montrer que $h - h'$ a la même valeur sur les voisins de v, et utiliser la connexité du graphe pour en déduire une contradiction.

(d) déduire des deux questions précédentes que h est à valeurs rationnelles, donc que tous les carrés sont à cotés rationnels et conclure.

IX. Sur un anneau euclidien

Voir [NQ92, Chap III], [Zis96].

Sur un anneau euclidien R (comme $\mathbb{Z}$ ou $K[X]$ avec K un corps), on peut remplacer la division exacte dans l'algorithme du pivot de Gauss par une division euclidienne. Ceci complique un peu l'algorithme, mais on obtient des résultats analogues : la forme de Hermite et de Smith d'une matrice à coefficients dans R.

Théorème IX.1 (Forme de Hermite)**.** *Soit R un anneau euclidien, et $A \in \mathcal{M}_{np}(R)$.*

Il existe une suite d'opérations élémentaires sur les lignes qui transforme A en une matrice échelonnée.

Remarque IX.2. La notion d'opération élémentaire est la même que sur un corps sauf que les opérations de dilatation $L_i \leftarrow \lambda L_i$ doivent avoir un coefficient λ inversible dans R).

Théorème IX.3 (Forme de Smith)**.** *Soit R un anneau euclidien, et $A \in \mathcal{M}_{np}(R)$.*

Il existe une suite d'opérations élémentaires sur les lignes et les colonnes qui transforme A en une matrice "diagonale" (mais pas carrée si $n \neq p$)

$$D = \left[\begin{array}{ccc|c} d_1 & & & \\ & \ddots & & 0 \\ & & d_r & \\ \hline & 0 & & 0 \end{array}\right]$$

avec $d_1 | \ldots | d_r$.

Les d_i sont uniques à association[a] *près, et s'appellent facteurs invariants de A.*

a. 2 éléments $a, b \in R$ sont associés si $b = au$ pour un certain $u \in R^\times$

Remarque IX.4. En fait les théorèmes ci-dessus sont vrais plus généralement sur un anneau R principal.

La solution de l'exo suivant est très proche de celle l'exo 17 ci-dessus.

Exercice 26. Unicité et caractérisation des facteurs invariants.

Soit R un anneau euclidien.

(a) Soit $A \in \mathcal{M}_{n,p}(R)$ et $A' = EA$ obtenue à partir de A en effectuant une opération élémentaire sur les lignes Montrer que l'idéal de R engendré par les mineurs de taille k ne change pas quand on change A en $A' = EA$.

(b) Application aux facteurs invariants : the théorème de Smith dit que si R est un anneau euclidien (voire principal), étant donnée une matrice $A \in \mathcal{M}_{np}(R)$, on peut lui appliquer des opérations élémentaires sur les lignes et les colonnes pour obtenir une matrice "diagonale" (mais pas carrée si $n \neq p$)

$$D = \left[\begin{array}{ccc|c} d_1 & & & \\ & \ddots & & 0 \\ & & d_r & \\ \hline & 0 & & 0 \end{array}\right]$$

avec $d_1 | \dots | d_r$ (les d_i s'appellent facteurs invariants de A). Déduire de (c) que pour tout $k \leq r$, $d_1 \dots d_k$ est égal (à association près) au pgcd des mineurs de taille k de A.

Comme dans le cas des corps, ces algorithmes permettent de prouver des résultats théoriques et de faire des calculs sur des sous-modules : tout sous-module de R^n admet une base, algorithme de recherche de base, théorème de la base adaptée [7] classification des groupes abéliens de type fini... Voir [NQ92], [Art91, §14.4-14.8].

Remarque IX.5. Attention : dans ce cadre, il ne peut pas y avoir de passage de représentation paramétrique à représentation cartésienne d'un sous-module, tout simplement parce qu'il n'est pas vrai que tout sous-module peut s'écrire comme ensemble de solutions d'un système d'équations linéaires (le sous-module $M = 2\mathbb{Z}$ de $\mathbb{Z}$ est un exemple).

Forme de smith, $K[X]$-modules, et application en algèbre linéaire. Voir [NQ92, Chap III.6], [Zis96, Chap VI.3].

Soit E un K-espace vectoriel et u un endomorphisme de E. On associe à la donnée de (E, u) une structure de $K[X]$-module sur E : pour $P \in K[X]$, et $v \in E$, on définit $P.v = P(u)(v)$. En particulier, la multiplication par X correspond à appliquer l'endomorphisme u. La multiplication par le polynôme P correspond à appliquer l'endomorphisme $P(u)$.

Exercice 27. $K[X]$-modules et endomorphismes.

Montrer que les $K[X]$-modules associés à (E, u), et (E', u') sont isomorphes si et seulement si les endomorphismes u et u' sont conjugués.

7. pour tout sous-module M de R^n, il existe une base $v_1, \dots, v_n$ de R^n et $d_1|dots|d_r$ tels que $d_1 v_1, \dots, d_r v_r$ soit une base de M.

Rappelons que u et u' sont conjugués s'il existe un isomorphisme de K-espace vectoriels $\varphi : E \to E'$ tel que $u' = \varphi \circ u \circ \varphi^{-1}$, ou de manière équivalente $\varphi \circ u = u' \circ \varphi$.

Soit E, F deux K-espaces vectoriel de dimension n, et u, v des endomorphismes de E et F respectivement. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, $\mathcal{B}' = (e'_1, \ldots, e'_n)$ une base de F, et M et N les matrices de u et v dans ces bases. On voit E et F comme deux $K[X]$-modules comme dans l'exercice ci-dessus.

On va démontrer le théorème suivant :

Théorème IX.6 (conjugaison vs équivalence)**.** *Les énoncés suivants sont équivalents :*

1. *les endomorphismes u et v sont conjugués*
2. *les matrices M et N sont semblables*
3. *les matrices $M - XI_n$ et $N - XI_n$ sont équivalentes dans $M_n(K[X])$.*

L'équivalence entre 1 et 2 et l'implication $2 \Rightarrow 3$ sont faciles.

L'implication $3 \Rightarrow 1$ passe par le fait que le $K[X]$-module associé à (E, u) possède la présentation suivante : il est isomorphe à $K[X]^n/\mathrm{Im}(M - XI_n)$. ([8])

Exercice 28. Présentation du $K[X]$-module E

On note $E_1 = (1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1)$ la base canonique de $K[X]^n$. Soit $\Phi : K[X]^n \to E$ le morphisme de $K[X]$-modules défini par $\Phi(E_i) = e_i$.

(a) Vérifier que $\Phi(P_1, \ldots, P_n) = \sum_{i=1}^n P_i.e_i$.

(b) La matrice $M - XI_n \in \mathcal{M}_n(K[X])$ définit une application $K[X]$-linéaire de $K[X]^n$ dans lui même. Montrer que $\mathrm{Im}(M - XI_n) \subset \ker \Phi$.

(c) Montrer que $K[X]^n/\mathrm{Im}(M - XI_n)$ est engendré en tant que K-espace vectoriel par les images $\overline{E}_1, \ldots, \overline{E}_n$, et donc que $K[X]^n/\mathrm{Im}(M - XI_n)$ est de dimension au plus n sur K.

(d) En déduire que $\mathrm{Im}(M - XI_n) = \ker \Phi$, donc que Φ induit un isomorphisme de $K[X]$-modules entre E et $K[X]^n/\mathrm{Im}(M - XI_n)$. Cette identification de E avec ce quotient est une présentation du $K[X]$-module E.

On peut maintenant prouver le théorème :

Exercice 29. Preuve du théorème IX.6

On suppose que $M - XI_n$ et $N - XI_n$ sont équivalentes dans $M_n(K[X])$. De manière analogue à Φ, on considère $\Psi : K[X]^n \to F$ dont le noyau est $\mathrm{Im}(N - XI_n)$.

(a) Montrer qu'il existe des automorphismes P, Q de $K[X]^n$ faisant commuter le diagramme suivant

$$\begin{array}{ccccc} K[X]^n & \xrightarrow{M-XI_n} & K[X]^n & \xrightarrow{\Phi} & E \\ \downarrow P & & \downarrow Q & & \\ K[X]^n & \xrightarrow{N-XI_n} & K[X]^n & \xrightarrow{\Psi} & F \end{array}$$

(b) En déduire que Q passe au quotient en un isomorphisme de $K[X]$-modules entre E et F, et que u et v sont donc conjugués.

On déduit du théorème IX.6 l'algorithme suivant.

8. Pour les amateurs, cela signifie qu'on a une suite exacte de $K[X]$-modules $K[X]^n \xrightarrow{M-XI_n} K[X]^n \to E \to 0$.

Théorème IX.7. *Soit K un corps dans lequel on peut calculer de manière exacte (comme $\mathbb{Q}$ ou un corps fini). L'algorithme suivant permet de décider si deux matrices données $A, B \in \mathcal{M}_n(K)$ sont conjuguées :*
Algorithme.

1. *Calculer la forme de Smith de la matrice $A - XI_n$ vue comme matrice dans $\mathcal{M}_n(K[X])$*
2. *En déduire ses facteurs invariants $f_1|f_2|\dots|f_n \in K[X]$ (montrer au passage que la forme de Smith est une matrice diagonale a coefficients non nuls, il y a donc n facteurs invariants)*
3. *Calculer de même les facteurs invariants $g_1|g_2|\dots|g_n$ pour $B - XI_n$*
4. *Regarder si pour tout $i \leq n$, P_i et Q_i sont associés. Si oui, répondre que A et B sont conjuguées dans $M_n(K)$.*

Exercice 30.

Démontrer le théorème, c'est à dire que la réponse de l'algorithme est correcte.

On peut aussi en déduire l'énoncé suivant (qui marche plus généralement avec n'importe quelle extension de corps $K \subset L$) :

Exercice 31.

Démontrer le corollaire suivant :

Corollaire. *Soient $A, B \in M_n(\mathbb{Q})$. Si A et B sont conjuguées dans $M_n(\mathbb{C})$ alors A et B sont conjuguées dans $M_n(\mathbb{Q})$.*

Le même argument permet de montrer que A est conjugué à sa transposée.

Lien avec la décomposition de Frobénius La décomposition de Smith de $A - XI_n$ implique que le $K[X]$-module associé à A est isomorphe à $K[X]/\langle f_1\rangle \oplus \dots \oplus K[X]/\langle f_n\rangle$ où les f_i sont les facteurs invariants de $A - XI_n$. Si on retraduit le $K[X]$-module $K[X]/\langle f_1\rangle$ en K-espace vectoriel muni d'une application linéaire u, on voit que u est conjugué à une matrice compagnon. C'est très facile : si $d = \deg(f_1)$, on voit que $1, \bar{X}, \dots, \bar{X}^{d-1}$ une base du K-espace vectoriel $K[X]/\langle f_1\rangle$, et dans cette base, la matrice de la multiplication par X est la matrice compagnon de f. Cas très particulier, si f_1 est un polynome constant, $K[X]/\langle f_1\rangle = \{0\}$, et on peut le laisser tomber dans la somme. Revenant à la matrice A, on en déduit qu'elle est semblable à une matrice par blocs formée des matrices compagnons des f_i non constants : c'est la décomposition de Frobénius de A. Voir aussi [Rom17, §21.9] pour une approche plus terre à terre.

X. Vrai ou Faux ?

Exercice 32.

Répondre par vrai ou faux et justifier

(a) pour une matrice carrée, être échelonnée ou triangulaire supérieure, c'est la même chose

(b) si $E_1, \dots, E_n$ sont des sev de E avec $E_i \cap E_j = 0$ alors leur somme est directe

(c) tout systeme linéaire homogéne de 8 équations linéaires à 15 inconnues admet toujours au moins une solution non nulle.

(d) Dans $\mathbb{R}^{11}$ un sev de dimension 8 et un sev de dimension 6 peuvent avoir une intersection de dimension 2

(e) Dans $\mathbb{R}^{11}$ un sev de dimension 3 et un sev de dimension 7 peuvent avoir une intersection de dimension 5

(f) Dans $\mathbb{R}^{11}$ un sev de dimension 5 et un sev de dimension 7 peuvent avoir une intersection de dimension 3

(g) Pour un système linéaire homogène de n équations à p inconnues, dire s'il est possible d'avoir le nombre de solutions indiqué :

	$n > p$	$n - p$	$n < p$
0 solution			
1 solution unique			
≥ 2 solutions			

(h) Pour un système linéaire de n équations à p inconnues avec second membre, dire s'il est possible d'avoir le nombre de solutions indiqué :

	$n > p$	$n - p$	$n < p$
0 solution			
1 solution unique			
≥ 2 solutions			

(i) Soit $A \in \mathcal{M}_{np}(\mathbb{R})$ une matrice à coefficient réels. Si le système linéaire $AX = b$ possède une solution dans $\mathbb{R}^p$ pour tout $b \in \mathbb{R}^n$, alors ce même système a une solution complexe pour tout $b \in \mathbb{C}^n$.

Références

[AHU74] Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman. *The design and analysis of computer algorithms*. 1974.

[Art91] Michael Artin. *Algebra*. Prentice-Hall, Englewood Cliffs, NJ, 1991.

[CG17] Philippe Caldero and Jérôme Germoni. *Nouvelles histoires hédonistes de groupes et de géométries. Tome 1*, volume 117 of *Math. Devenir*. Paris : Calvage et Mounet, 2nd edition edition, 2017.

[Cia98] P.G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques appliquées pour la maîtrise. Dunod, 1998.

[CLRS22] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms*. Cambridge, MA : MIT Press, 4th edition edition, 2022.

[Gri15] Joseph Grifone. *Algèbre linéaire.* Toulouse : Cépaduès-Éditions, 5th edition edition, 2015.

[NQ92] Patrice Naudin and Claude Quitté. *Algorithmique algébrique.*, volume 8 of *Log. Math. Inform.* Masson, Paris, 1992.

[Per95] Daniel Perrin. *Cours d'algèbre*. CAPES-AGREG mathématiques. Ellipses, Paris, 1995.

[Rom17] J.É. Rombaldi. *Mathématiques pour l'Agrégation : Algèbre & géométrie*. LMD MATHS. De Boeck supérieur, 2017.

[Sch98] M. Schatzman. *Analyse Numerique Cours Et Exerices Pour La Licence*. Elsevier Masson, 1998.

[Zis96] Michel Zisman. *Mathématiques pour l'Agrégation : avec exercices*. Paris : Dunod, 1996.