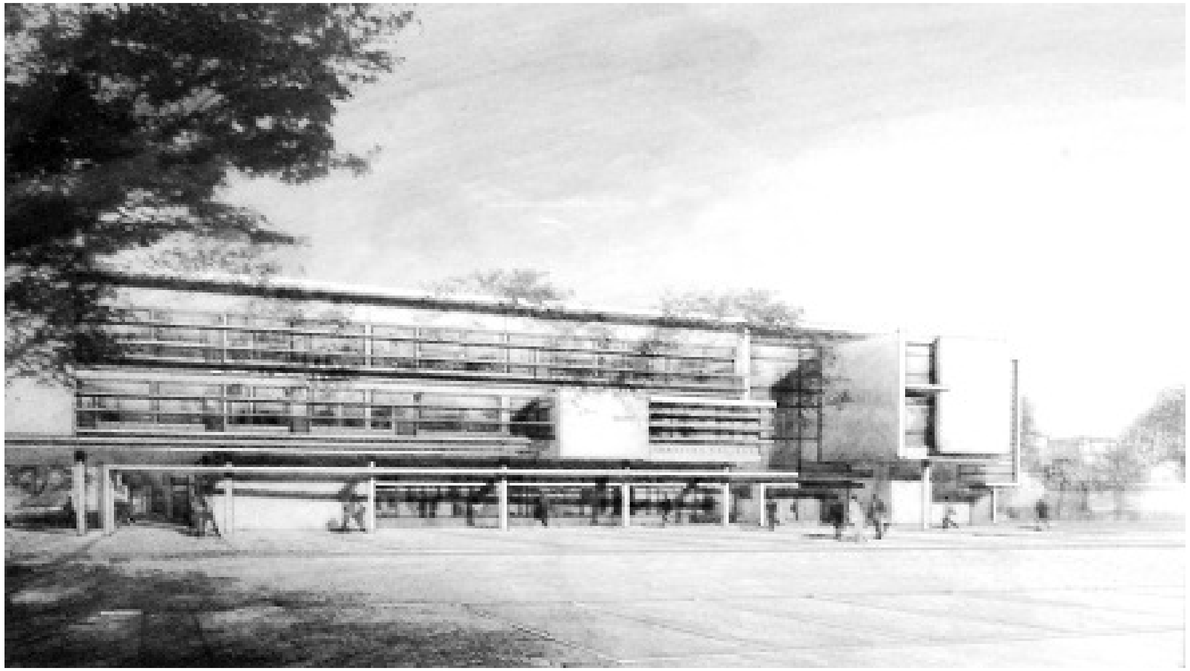
Institut Galilée

Sciences et technologies



Licence 1$^{\text{ère}}$ année. Deuxième semestre 2016/2017

Algèbre linéaire et algorithmique

Introduction

Ce polycopié est destiné aux étudiants du tronc commun de L1 mathématiques/informatique de l'institut Galilée et concerne le cours *algèbre linéaire et algorithmique*. Le sujet principal est l'algèbre linéaire, qui généralise et formalise l'étude des systèmes linéaires. Cette théorie peut être vue comme l'étude systématique d'ensembles (les espaces vectoriels) munis de deux opérations : une loi de composition interne appelée addition et une multiplication par des nombres réels ou complexes. Ces notions de mathématiques seront illustrées et appliquées dans des cours, travaux dirigés et travaux pratiques d'algorithmique.

On commence (chapitre I) par expliquer la résolution de systèmes linéaires généraux par la méthode du pivot de Gauss. Ces systèmes linéaires peuvent s'écrire de manière plus succincte en utilisant des tableaux de nombres, ou matrices, qui sont étudiés de manière plus systématique au chapitre II. Le chapitre III ne concerne pas l'algèbre linéaire : il s'agit d'une introduction rapide au polynômes, qui seront utilisés notamment au chapitre VII et dans la partie informatique du cours. Le chapitre IV contient de bref rappel sur le langage C et sur le traitement des matrices et des polynômes dans ce langage.

L'étude de l'algèbre linéaire proprement dite commence au premier chapitre du Tome II, le chapitre V, où on introduit les espaces vectoriels et la notion cruciale de *dimension*. Au chapitre VI on étudie les applications linéaires, qui sont les applications d'un espace vectoriel dans un autre préservant la structure d'espace vectoriel. On y revient notamment sur les matrices vues précédemment. Le chapitre VII introduit le *déterminant* d'une application linéaire ou d'une matrice. Le chapitre VIII traite enfin de la complexité des algorithmes.

Cette version est celle du 23 janvier 2017 (3 corrections par rapport à la première édition 2017). Les corrections sont listées sur :

`https://www.math.univ-paris13.fr/~duyckaer/enseignement.html`

I. Systèmes linéaires

La référence principale pour ce chapitre est le livre de David C. Lay [1].

On appellera *nombre* ou *scalaire* un nombre rationnel, réel ou complexe. On posera $\mathbb{K} = \mathbb{Q}$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ l'ensemble de ces nombres. Le choix des nombres rationnels, réels ou complexes est indifférent dans ce chapitre, sauf pour les interprétations géométriques où l'on privilégiera les nombres réels.

I.1. Définitions et premiers exemples

I.1.a. Définitions

Définition I.1.1. Soit $n \geq 1$. On appelle *équation linéaire* à n inconnues $x_1, \ldots, x_n$ une équation de la forme

$$\text{(E)} \qquad \sum_{j=1}^{n} a_j x_j = b,$$

où $a_1, a_2, \ldots, a_n$ et b sont fixés dans $\mathbb{K}$. Les scalaires $a_1, a_2, \ldots, a_n$ sont appelés *coefficients* de l'équation, b est le *second membre*. Lorsque $b = 0$, on dit que l'équation est *homogène*.

Remarque I.1.2. La notation $\sum_{j=1}^{n} a_j x_j$ dans (E) signifie $a_1x_1 + a_2x_2 + \ldots + a_nx_n$. Il est impératif de maîtriser ce type de notation.

Définition I.1.3. L'*ensemble des solutions* de (E) est l'ensemble des $(x_1, x_2, \ldots, x_n)$ de $\mathbb{K}^n$ tels que $\sum_{j=1}^{n} a_j x_j = b$. C'est donc un sous-ensemble de $\mathbb{K}^n$.

Exemple I.1.4.

$$3x_1 + x_2 - 4x_3 = 2$$

est une équation linéaire non-homogène à 3 inconnues.

$$-ix_1 + (2 + i)x_2 - 4x_3 = 0$$

est une équation linéaire (complexe) homogène à 3 inconnues.

$$2x_1^2 + x_2x_3 - 4x_3^3 = 1$$

n'est pas une équation linéaire.

1. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

Définition I.1.5. Si $p \geq 1$, on appelle *système linéaire* à p équations et n inconnues un ensemble de p équations linéaires ayant les mêmes n inconnues :

$$\text{(S)} \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{p1}x_1 + a_{22}x_2 + \ldots + a_{pn}x_n = b_p \end{cases}$$

Les scalaires a_{ij}, $1 \leq i \leq p$, $1 \leq j \leq n$ sont encore appelés les *coefficients* du système. Il est d'usage d'utiliser le premier indice pour numéroter les lignes et le deuxième indice pour numéroter les colonnes. Le p-uplet $b = (b_1, \ldots, b_p)$ est appelé *second membre* du système. Lorsque tous les b_j sont nuls (on dit encore que b est nul), on dit que le système est *homogène*.

L'*ensemble des solutions de* (S) est le sous-ensemble de $\mathbb{K}^n$ formé des $(x_1, \ldots, x_n)$ qui vérifient *toutes* les équations de (S). On cherche à résoudre le système (S), c'est à dire décrire précisément cet ensemble.

Définition I.1.6. Le système (S) est dit *compatible* si il a au moins une solution.

Remarque I.1.7. Un système homogène est toujours compatible : $(0, 0, \ldots, 0)$ est solution.

Remarque I.1.8. On peut réécrire le système (S) sous forme abrégée :

$$\forall i = 1 \ldots p, \quad \sum_{j=1}^{n} a_{ij}x_j = b_i.$$

Remarque I.1.9. Lorsque le nombre d'inconnues n est petit, on note souvent x, y, z, t (au lieu de x_1, x_2,...) ces inconnues pour alléger les notations.

Donnons quelques exemples simples.

Exemples I.1.10.

$$\begin{cases} 17(x + 2y) = \sqrt{7}z + 3 \\ \dfrac{x + z}{2} = 12. \end{cases}$$

est un système linéaire non-homogène, à 2 équations et 3 inconnues, que l'on peut écrire sous la forme (S) (avec $x = x_1$, $y = x_2$, $z = x_3$) :

$$\begin{cases} 17x + 34y - \sqrt{7}z = 3 \\ \dfrac{1}{2}x + 0y + \dfrac{1}{2}z = 12. \end{cases}$$

Ici $a_{11}=17$, $a_{12}=34$, $a_{13}=-\sqrt{7}$, $a_{21}=\frac{1}{2}$, $a_{22}=0$, $a_{23}=\frac{1}{2}$, $b_1=3$ et $b_2=12$. Le système :

$$\begin{cases} xyz+7=0 \\ x+2y=3. \end{cases}$$

n'est pas linéaire (la première équation ne peut pas être mise sous la forme d'une équation linéaire). L'équation :

$$(x+y-3)^2+(2x+y+2)^2=0 \tag{I.1}$$

n'est pas linéaire. Toutefois, si l'on cherche à résoudre cette équation sur $\mathbb{R}$, elle est équivalente au système linéaire :

$$\begin{cases} x+y=3 \\ 2x+y=-2. \end{cases}$$

Remarquons que dans le cas $\mathbb{K}=\mathbb{C}$, l'équation (I.1) ne peut pas être ramenée à un système linéaire.

Exercice I.1.11. Mettre les systèmes linéaires suivants sous la forme (S). Déterminer p, n, et les paramètres a_{ij} et b_i.

$$\begin{cases} 3x+y=4z+3 \\ y=z, \end{cases} \tag{I.2}$$

$$x_1+x_2=x_2+x_3=x_3+x_4=0. \tag{I.3}$$

Les systèmes linéaires apparaissent dans tous les domaines d'applications des mathématiques (économie, industrie...) Dans les applications, p et n sont souvent très grands, et on ne peut pas résoudre le système "à la main". Il existe évidemment de nombreux logiciels informatiques qui en sont capables. Les buts de ce chapitre sont :

- savoir résoudre "à la main" un système lorsque p et n sont petits ;
- comprendre une méthode de résolution d'un système général, la méthode du pivot de Gauss ;
- en déduire quelques propriétés de la structure de l'ensemble des solutions. Cette structure sera précisée au Chapitre V, à travers la notion d'espace vectoriel.

On commence par donner des exemples de résolutions de systèmes linéaires à une ou deux équations et une ou deux inconnues, puis une notation commode (la notation matricielle) avant de dégager une méthode générale.

I.1.b. Exemples de petits systèmes linéaires

Une équation à une inconnue

On fixe $(a,b)\in\mathbb{K}^2$. On considère l'équation :

$$ax=b. \tag{I.4}$$

Alors :

- Si $a\neq 0$, $x=\frac{b}{a}$ est la seule solution de (I.4).
- Si $a=0$ et $b=0$, l'équation (I.4) s'écrit $0=0$ et tous les $x\in\mathbb{K}$ sont solutions.
- Si $a=0$ et $b\neq 0$, l'équation (I.4) s'écrit $b=0$, il n'y a donc pas de solution.

Remarque I.1.12. L'ensemble des solutions est ou bien vide, ou bien réduit à un seul élément, ou bien infini. Nous verrons plus tard (théorème I.2.23 p. 11) que cette propriété persiste dans le cas d'un système linéaire général.

Une équation, deux inconnues

On considère maintenant l'équation linéaire :

$$ax+by=c. \tag{I.5}$$

Supposons d'abord $(a,b)\neq(0,0)$. Si $\mathbb{K}=\mathbb{R}$, on sait que (I.5) est l'équation d'une droite du plan $\mathbb{R}^2$ et il y a une infinité de solutions. Si $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$, on peut résoudre le système algébriquement :

- si $b\neq 0$, on peut réécrire (I.5) $y=\frac{c}{b}-\frac{a}{b}x$, et l'ensemble des solutions est

  $$\left\{\left(x,\frac{c}{b}-\frac{a}{b}x\right),\ x\in\mathbb{K}\right\}.$$

  On dit qu'on a paramétré cet ensemble (x est le paramètre).
- si $b=0$ et $a\neq 0$, l'équation s'écrit $x=c/a$ et l'ensemble des solutions est donné par $\left\{\left(\frac{c}{a},y\right),\ y\in\mathbb{R}\right\}$. On a de nouveau paramétré l'ensemble des solutions. Cette fois, le paramètre est y.

Lorsque $(a,b)\neq(0,0)$ il y a donc une infinité de solutions. Si $(a,b)=(0,0)$, l'équation (I.5) s'écrit simplement $0=c$: il y a une infinité de solutions (tous les couples $(x,y)\in\mathbb{K}^2$) si $c=0$, et aucune solution si $c\neq 0$.

Remarque I.1.13. Quelles que soient les valeurs de a,b et c, l'équation (I.5) a ou bien une infinité de solutions, ou bien pas de solution du tout (mais jamais un nombre fini, non nul de solutions : comparer avec la remarque I.1.12). La suite du chapitre expliquera cette observation.

Deux équations, deux inconnues. Opérations sur les lignes

On considère maintenant un système de la forme :

$$\begin{cases} a_{11}x+a_{12}y=b_1 \\ a_{21}x+a_{22}y=b_2. \end{cases} \tag{I.6}$$

Lorsque $\mathbb{K}=\mathbb{R}$, $(a_{11},a_{12})\neq(0,0)$ et $(a_{21},a_{22})\neq(0,0)$, (I.6) décrit l'ensemble des points d'intersection de deux droites $\mathcal{D}_1$ et $\mathcal{D}_2$ du plan $\mathbb{R}^2$. On a donc trois cas :

- Les droites $\mathcal{D}_1$ et $\mathcal{D}_2$ ne sont pas parallèles : elles ont un unique point d'intersection, et (I.6) n'a qu'une seule solution.

— Les droites $\mathcal{D}_1$ et $\mathcal{D}_2$ sont parallèles. Le système (I.6) n'a pas de solution, sauf si $\mathcal{D}_1$ et $\mathcal{D}_2$ sont confondues, auquel cas le système a une infinité de solutions (l'ensemble des coordonnées des points de la droite $\mathcal{D}_1 = \mathcal{D}_2$).

On donne maintenant trois exemples que l'on va résoudre algébriquement, sans utiliser d'argument géométrique.

Exemple I.1.14.

$$\begin{cases} 3x + 2y = 8 & (L_1) \\ 2x + 5y = 9 & (L_2). \end{cases} \tag{I.7}$$

Eliminons l'inconnue "x" de la deuxième ligne à l'aide de la première ligne :

$$\begin{cases} 3x + 2y = 8 \\ 0 + \left(5 - \frac{4}{3}\right) y = \left(9 - \frac{16}{3}\right) & (L_2) - \frac{2}{3}(L_1). \end{cases} \tag{I.8}$$

La notation $(L_2) - \frac{2}{3}(L_1)$ à la deuxième ligne (on écrira parfois $(L_2) \leftarrow (L_2) - \frac{2}{3}(L_1)$) signifie que l'on a remplacé la deuxième ligne par la différence de la deuxième ligne et du produit de $\frac{2}{3}$ par la première ligne.

$$\begin{cases} 3x + 2y = 8 \\ y = 1 & \scriptstyle\frac{3}{11}(L_2). \end{cases} \tag{I.9}$$

Une fois connue la valeur de y il suffit de la substituer dans (L_1) pour obtenir la valeur de x : on obtient $3x = 8 - 2 = 6$, i.e. $x = 2$. Le système (I.7) a donc pour unique solution $(2, 1)$ (en d'autres termes, l'ensemble des solutions est le singleton $\{(2, 1)\}$).

On peut aussi conclure, à partir de (I.9), en utilisant des opérations sur les lignes :

$$\begin{cases} 3x = 6 & \scriptstyle(L_1)-2(L_2) \\ y = 1 \end{cases} \quad \text{puis} \quad \begin{cases} x = 2 & \scriptstyle\frac{1}{3}(L_1) \\ y = 1 \end{cases} \tag{I.10}$$

Exemple I.1.15.

$$\begin{cases} 3x + 2y = 8 \\ 6x + 4y = 20. \end{cases} \tag{I.11}$$

On élimine x comme dans l'exemple I.1.14

$$\begin{cases} 3x + 2y = 8 \\ 0 = 4 & \scriptstyle(L_2)-2(L_1). \end{cases} \tag{I.12}$$

Il n'y a pas de solution.

Exemple I.1.16.

$$\begin{cases} 3x + 2y = 8 \\ 6x + 4y = 16. \end{cases} \tag{I.13}$$

On utilise la même stratégie :

$$\begin{cases} 3x + 2y = 8 \\ 0 = 0 & \scriptstyle(L_2)-2(L_1). \end{cases} \tag{I.14}$$

La deuxième équation est toujours vraie. Le système (I.13) est donc équivalent à l'équation, $3x + 2y = 8$. Il y a, comme pour l'équation (I.5) lorsque $(a, b) \neq (0, 0)$, toute une droite de solutions, l'ensemble

$$\left\{(x, 4 - \frac{3}{2}x),\ x \in \mathbb{K}\right\}.$$

I.1.c. Notation matricielle

Une *matrice* $p \times n$ est un tableau de nombre à p lignes et n colonnes. Nous étudierons les matrices de manière plus systématique dans le chapitre II de ce cours.

Lors des manipulations sur les lignes des exemples précédents, on peut gagner du temps en décidant de ne pas noter les variables :

Définition I.1.17. On appelle *matrice des coefficients* du système (S) la matrice $p \times n$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \vdots & & \vdots \\ a_{p1} & a_{22} & \dots & a_{pn} \end{bmatrix}$$

et *matrice augmentée* la matrice $p \times (n + 1)$:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ & \vdots & & \vdots & \\ a_{p1} & a_{22} & \dots & a_{pn} & b_p \end{array}\right]$$

(le trait vertical, facultatif, permet de séparer les coefficients du système du second membre).

On peut reprendre les opérations précédentes en notation matricielle. Par exemple, les manipulations sur les lignes de l'exemple I.1.14 s'écrivent :

$$\left[\begin{array}{cc|c} 3 & 2 & 8 \\ 2 & 5 & 9 \end{array}\right] \qquad \left[\begin{array}{cc|c} 3 & 2 & 8 \\ 0 & 5-\frac{4}{3} & 9-\frac{16}{3} \end{array}\right]_{(L_2)-\frac{2}{3}(L_1)}$$

$$\left[\begin{array}{cc|c} 3 & 2 & 8 \\ 0 & 1 & 1 \end{array}\right]_{\frac{3}{11}(L_2)} \qquad \left[\begin{array}{cc|c} 3 & 0 & 6 \\ 0 & 1 & 1 \end{array}\right]^{(L_1)-2(L_2)} \qquad \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right]^{\frac{1}{3}(L_1)}$$

ce qui signifie bien $x = 2$, $y = 1$ (cf (I.10)).

Exercice I.1.18. Résoudre le système

$$\text{(I.15)} \qquad \begin{cases} 2x + y = 5 \\ x - y = 1 \end{cases}$$

en utilisant la notation matricielle.

I.2. Méthode du pivot

On formalise et généralise maintenant la méthode précédente pour résoudre des systèmes linéaires quelconques. Plus précisément :

— on définit en I.2.a des opérations sur les lignes des systèmes linéaires qui ne changent pas l'ensemble des solutions : les *opérations élémentaires*. Ce sont exactement les opérations apparaissant dans les exemples précédents ;

— on introduit en I.2.b une famille de systèmes pour lesquels il est très facile de décrire l'ensemble des solutions, les systèmes *sous forme échelonnée réduite* ;

— on montre en I.2.c comment ramener, par une série d'opérations élémentaires, tout système linéaire, à un système sous forme échelonnée réduite : c'est la méthode du pivot proprement dite.

Cette partie se termine par l'étude d'une classe de système importante, les systèmes de Cramer (I.2.d).

I.2.a. Systèmes équivalents. Opérations élémentaires

Définition I.2.1. Deux systèmes linéaires ayant le même nombre d'inconnues sont *équivalents* lorsqu'ils ont le même ensemble de solutions.

Exemple I.2.2. Les systèmes suivants sont équivalents :

$$\text{(I.16)} \qquad \begin{cases} x + y = 0 \\ \quad y = -1 \end{cases} \quad \text{et} \quad \begin{cases} x + y = 0 \\ \quad x = 1 \end{cases}$$

En effet, l'ensemble des solutions de ces deux systèmes est $\{(1, -1)\}$. Les deux systèmes suivants ne sont pas équivalents :

$$\text{(I.17)} \qquad \begin{cases} x + y = 0 \\ \quad y = -1 \end{cases} \quad \text{et} \quad \begin{cases} x + y = 0 \\ \quad x = -1 \end{cases}$$

En effet, l'ensemble des solutions du premier système est $\{(1, -1)\}$, l'ensemble des solutions du deuxième système est $\{(-1, 1)\}$.

Définition I.2.3. On appelle *opération élémentaire* sur un système (ou sur les lignes d'une matrice) l'une des trois opérations suivantes :

i. L'*échange* de deux lignes (L_i) et (L_j), noté $(L_i) \leftrightarrow (L_j)$.

ii. La multiplication d'une ligne par un scalaire $k \in \mathbb{K}$ *non nul*, appelée *cadrage*, notée $(L_i) \leftarrow k(L_i)$.

iii. L'ajout à une ligne du multiple d'une autre ligne par un scalaire k, appelé *remplacement*, noté $(L_i) \leftarrow (L_i) + k(L_j)$.

Le lecteur est invité à vérifier que toutes les opérations réalisées dans les exemples de la partie I.1 sont des opérations élémentaires au sens de la définition I.2.3. Par exemple, l'opération conduisant à (I.8) est un remplacement, celle conduisant à (I.9) un cadrage.

Remarque I.2.4. Il revient bien sûr au même de faire des opérations élémentaires sur un système d'équations linéaires, ou sur les lignes de la matrice augmentée de ce système.

Les opérations élémentaires ne changent pas l'ensemble des solutions :

Proposition I.2.5. *Soient (S) et (S') deux systèmes linéaires ayant le même nombre d'inconnues et d'équations. On suppose que (S') peut être obtenu à partir de (S) par une série d'opérations élémentaires. Alors (S) et (S') sont équivalents.*

Démonstration. Par une récurrence simple, il suffit de montrer qu'une seule opération élémentaire ne change pas l'ensemble des solutions.

C'est évident si cette opération élémentaire est l'échange de deux lignes.

Si k est fixé, non nul, on peut diviser les deux membres d'une égalité par k. On a donc :

$$k(a_{i1}x_1 + \ldots + a_{in}x_n) = k\, b_i \iff a_{i1}x_1 + \ldots + a_{in}x_n = b_i,$$

ce qui montre que le cadrage (multiplication d'une ligne par un scalaire non nul) ne change donc pas l'ensemble des solutions.

Il reste à traiter le cas du remplacement, i.e. l'opération $(L_j) \leftarrow (L_j) + k(L_i)$, où $i \neq j$, $k \in \mathbb{K}$. En ignorant les lignes autres que la i-ième et la j-ième, qui ne changent pas, on est amené à montrer que les deux systèmes

$$\begin{cases} a_{i1}x_1 + \ldots + a_{in}x_n = b_i & (L_i) \\ a_{j1}x_1 + \ldots + a_{jn}x_n = b_j & (L_j) \end{cases}$$

et

$$\begin{cases} a_{i1}x_1 + \ldots + a_{in}x_n = b_i & (L_i) \\ a_{j1}x_1 + \ldots + a_{jn}x_n + k(a_{i1}x_1 + \ldots + a_{in}x_n) = b_j + kb_i & (L'_j) \end{cases}$$

sont équivalents, c'est à dire que $x = (x_1, \ldots, x_n)$ vérifie (L_i) et (L_j) si et seulement si il vérifie (L_i) et (L'_j).

On suppose d'abord que x vérifie (L_i) et (L_j). Alors

$$\underbrace{a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n}_{=b_j \text{ par } (L_j)} + k\underbrace{(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n)}_{=b_i \text{ par } (L_i).} = b_j + kb_i,$$

ce qui montre (L'_j).

Supposons réciproquement que x vérifie (L_i) et (L'_j). Alors

$$\begin{aligned} &a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n \\ &\quad = \underbrace{a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n + k\left(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n\right)}_{=b_j+kb_i \text{ par } (L'_j)} \\ &\qquad\qquad - k\underbrace{\left(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n\right)}_{=b_i \text{ par } (L_i)} = b_j, \end{aligned}$$

d'où (L_j). □

Avertissement I.2.6. Pour passer d'un système linéaire à un système équivalent, il est très fortement conseillé de s'en tenir à des opérations élémentaires au sens de la définition (I.2.3). Des opérations sur les lignes mal choisies peuvent changer l'ensemble des solutions. Une erreur typique est de réaliser *simultanément* deux remplacements de la forme $(L_i) \leftarrow (L_i) + k(L_j)$ et $(L_j) \leftarrow (L_j) + k'(L_i)$. Par exemple :

$$\text{(I.18)} \qquad \begin{cases} x + 2y = 2 \\ x - y = 0 \end{cases} \text{ donne : } \begin{cases} 3x = 2 & (L_1) + 2(L_2) \\ \dfrac{3}{2}x = 1 & (L_2) + \dfrac{1}{2}(L_1). \end{cases}$$

Les deux systèmes ci-dessus ne sont pas équivalents : le premier a pour unique solution $\left(\frac{2}{3}, \frac{2}{3}\right)$, le deuxième a une infinité de solutions : $\left\{\left(\frac{2}{3}, y\right),\ y \in \mathbb{K}\right\}$. Remarquons que les deux opérations simultanées $(L_2) \leftarrow (L_2) + \frac{1}{2}(L_1)$ et $(L_1) \leftarrow (L_1) + 2(L_2)$ ne peuvent pas s'écrire comme deux remplacements successifs. Le lecteur est invité à méditer cet exemple : pratiquement toutes les erreurs dans les résolutions de systèmes, en dehors des fautes de calcul, sont de cette forme là.

I.2.b. Forme échelonnée

On définit maintenant un ensemble de matrice correspondant à des systèmes dont on peut décrire l'ensemble des solutions sans aucun calcul supplémentaire. Le but de la méthode du pivot sera de se ramener à ce type de matrice.

Définitions

Définition I.2.7. Soit A une matrice $p \times n$. Une *ligne nulle* de A est une ligne de A formée uniquement de zéros. On appelle *élément de tête* d'une ligne non nulle de A l'élément non nul le plus à gauche de cette ligne. On dit que A est *sous forme échelonnée* (ou simplement *échelonnée*) lorsque les deux propriétés suivantes sont vérifiées :

i. Toutes les lignes non nulles sont situées au-dessus des lignes nulles.
ii. L'élément de tête de chaque ligne non nulle se trouve dans une colonne (strictement) à droite de l'élément de tête de la ligne précédente.

On dit que A est *sous forme échelonnée réduite* (ou simplement que A est *échelonnée réduite*) quand de plus les deux propriétés suivantes sont vérifiées :

iii. L'élément de tête de chaque ligne non nulle vaut 1.
iv. L'élément de tête de chaque ligne non nulle est le seul coefficient non nul de sa colonne.

Remarque I.2.8. Le point ii implique que tous les éléments de la matrice situés sous un élément de tête sont nuls.

Définition I.2.9. On dit qu'un système linéaire est *sous forme échelonnée* (respectivement *échelonnée réduite*) quand sa matrice augmentée est sous forme échelonnée (respectivement échelonnée réduite).

Exercice I.2.10. Ecrire un algorithme ou une fonction en C qui détermine si une matrice est échelonnée (respectivement échelonnée réduite).

Exemples I.2.11. La matrice $A = \left[\begin{smallmatrix} 0 & 0 \\ 2 & 3 \end{smallmatrix}\right]$ n'est pas échelonnée (i n'est pas vérifiée).

La matrice $B = \left[\begin{smallmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{smallmatrix}\right]$ n'est pas échelonnée (cette fois, ii est faux).

La matrice $C = \left[\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ est échelonnée, mais pas échelonnée réduite (iii et iv sont faux).

La matrice $D = \left[\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ est échelonnée, mais pas échelonnée réduite (iii est faux).

La matrice $E = \left[\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ est échelonnée réduite.

Le système $\begin{cases} x + 2z = 1 \\ y + 3z = 2 \end{cases}$ est sous forme échelonnée réduite.

Exemples I.2.12. Le premier système de (I.10) est sous forme échelonnée non réduite, le dernier système de (I.10) est sous forme échelonnée réduite.

Le système (I.12) est sous forme échelonnée non réduite.

Le système (I.14) est sous forme échelonnée non réduite. Après multiplication de la première ligne par $\frac{1}{3}$, on obtient la forme échelonnée réduite équivalente :

$$\begin{cases} x + \dfrac{2}{3}y = \dfrac{8}{3} \\ 0 = 0. \end{cases}$$

Dans ce système la deuxième ligne $0 = 0$ est bien sûr superflue.

Exercice I.2.13. Déterminer si les systèmes suivants sont sous forme échelonnée (respectivement sous forme échelonnée réduite) :

i. Un système à une seule équation.

ii.

$$\begin{cases} x_2 + x_3 = 2 \\ x_1 + x_3 = 3 \end{cases}, \text{ puis } \begin{cases} x_1 + x_3 = 2 \\ x_2 + x_3 = 3 \end{cases}$$

iii.

$$\left[\begin{array}{cccc|c} 2 & 1+2i & 5 & 4i & 1 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 2 \\ 1 & 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right]$$

$$\left[\begin{array}{cccccc|c} 0 & 0 & 1 & 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 1 & 0 & -3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

iv.

$$(S_1)\begin{cases} x = 1 \\ y = 2 \\ z = 3 \\ 0 = 0 \end{cases}, \qquad (S_2)\begin{cases} 2x + y + z = 1 \\ e^{i\pi/3}y + 2z = -1 \\ z = -4 \\ 0 = 0 \end{cases}, \qquad (S_3)\begin{cases} 2x + y + z = 1 \\ x + y + 2z = -1 \\ x - z = -4 \end{cases}$$

(cf correction p. 13).

Application aux systèmes linéaires

On rappelle qu'un système linéaire est dit *compatible* quand il a au moins une solution. On peut décrire facilement l'ensemble des solutions d'un système linéaire (S) dont la matrice augmentée A est sous forme échelonnée réduite. On rappelle que si (S) est un système à p équations et n inconnues, la matrice A est une matrice $p \times (n+1)$ (i.e. un tableau de nombres avec p lignes et $n+1$ colonnes).

On distingue deux cas.

— *Premier cas.* Si la colonne la plus à droite (la $n+1$-ème colonne) de A contient un élément de tête (forcément égal à 1) cela se traduit par une équation $0 = 1$, tautologiquement fausse, ce qui montre que le système n'a pas de solution. Le système (S) est n'est pas compatible.

Supposons par exemple que le système a pour matrice augmentée

$$\left[\begin{array}{ccc|c} 1 & 3 & 4 & 5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & \mathbf{2} \end{array}\right]$$

Cette matrice est sous forme échelonnée (non réduite). Le 2 en bas à droite est un élément de tête situé sur la dernière colonne : le système n'a pas de solution, car la dernière ligne se lit $0 = 2$.

— *Deuxième cas.* Supposons que la colonne de droite de A ne contient aucun élément de tête. Dans ce cas, le système est compatible. Donnons une méthode pour décrire l'ensemble des solutions. L'élément de tête de chaque ligne non nulle, situé sur une des n premières colonnes, correspond donc à une des inconnues $x_1, \ldots x_n$. On appelle *variable de base* du système toute inconnue x_j telle que la j-ième colonne contient un élément de tête non nul. On appelle *variables libres* ou *paramètres* les autres inconnues. Chaque ligne non nulle donne une expression d'une des variables de base en fonction des paramètres. En faisant varier les paramètres dans $\mathbb{K}$, on obtient exactement l'ensemble des solutions du système. On dit que l'on a obtenu une *description paramétrique* de l'ensemble des solutions. Dans ce cas, le nombre p' de lignes non nulles est exactement le nombre de variable de bases. Le nombre de paramètres est donc $n - p'$. Donnnons deux exemples.

Supposons d'abord que la matrice augmentée du système (S) est

$$\left[\begin{array}{cccc|c} \mathbf{1} & 0 & 3 & 0 & 4 \\ 0 & \mathbf{1} & 2 & 0 & 3 \\ 0 & 0 & 0 & \mathbf{1} & 5 \end{array}\right]$$

C'est une matrice de forme échelonnée réduite, dont les éléments de tête sont les "1" en gras. La colonne de droite ne contient aucun élément de tête : le système est compatible. Les variables de base (correspondant aux numéros des colonnes des éléments de tête) sont x_1, x_2 et x_4. Le seul paramètre est x_3. On obtient la description paramétrique suivante de l'ensemble des solutions :

$$x_1 = 4 - 3x_3, \; x_2 = 3 - 2x_3, \; x_4 = 5, \; x_3 \in \mathbb{K}.$$

En d'autres termes, l'ensemble des solutions de (S) est :

$$\{(4 - 3x_3, 3 - 2x_3, x_3, 5), \; x_3 \in \mathbb{K}\}.$$

Supposons que le système a pour matrice augmentée :

$$\left[\begin{array}{cccc|c} \mathbf{1} & 2 & 0 & -1 & 4 \\ 0 & 0 & \mathbf{1} & 2 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Cette matrice est sous forme échelonnée réduite, et le système est compatible. Les variables de base sont x_1 et x_3, et les variables libres x_2 et x_4. L'ensemble des solutions est

$$\Big\{(4 - 2x_2 + x_4, x_2, -5 - 2x_4, x_4), \; (x_2, x_4) \in \mathbb{K}^2\Big\}.$$

On peut maintenant donner une définition plus précise de la résolution d'un système linéaire : résoudre le système (S), c'est donner une description paramétrique de l'ensemble des solutions.

Remarque I.2.14. Un cas particulier de système compatible est donné par un système dont la matrice sous forme échelonnée réduite a autant de lignes non nulles qu'il y a d'inconnues. Dans ce cas, toutes les variables sont des variables de base, et il n'y a qu'une seule solution, dont les coordonnées $x_1, x_2, \dots, x_n$ sont données par les valeurs de la colonne de droite. On dit que le système est un *système de Cramer* (cf §I.2.d pour la définition générale et l'étude des systèmes de Cramer).

Remarque I.2.15. Il n'est pas nécessaire de mettre un système sous forme échelonnée réduite pour savoir si il est compatible ou non : une forme échelonnée non réduite convient tout aussi bien. Par le même raisonnement que précédemment, un système sous forme échelonnée est compatible si et seulement si la colonne de droite de sa matrice augmentée ne contient aucun élément de tête, i.e. si le système ne contient aucune ligne de la forme $0 = c$, où c est une constante non nulle.

Lorsque le système est compatible, la forme échelonnée réduite est commode pour décrire l'ensemble des solutions.

Exercice I.2.16. Dire si les systèmes de l'exercice I.2.13, lorsqu'ils sont sous forme échelonnée réduite, sont compatibles. Donner alors une description paramétrique de l'ensemble des solutions.

(cf correction p. 14).

La proposition suivante permet de compter le nombre de paramètres d'un système sous forme échelonnée réduite. Elle découle des arguments précédents.

Proposition I.2.17. *Un système compatible sous forme échelonnée réduite avec n inconnues et p' équations non nulles se décrit avec $n - p'$ paramètres. En d'autres termes, si le système est sous forme échelonnée réduite, on a :*

nombre d'inconnues − nombre d'équations = nombre de degrés de liberté.

Cette propriété persiste lorsque le système est sous forme échelonnée non réduite (cf Remarque I.2.24).

Exercice I.2.18. Dire si chacune des matrices suivantes est sous forme échelonnée (respectivement sous forme échelonnée réduite). Le système dont c'est la matrice augmentée est-il compatible ? Si c'est le cas, donner une description paramétrique de l'ensemble des solutions.

$$\text{a)} \left[\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

$$\text{b)} \left[\begin{array}{cccc|c} 1 & 2 & 3 & -4 & 4 \\ 0 & 1 & 1 & 2 & -5 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right], \quad \text{c)} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -3 & 1 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(cf correction p. 14).

I.2.c. Méthode du pivot de Gauss

Nous venons de voir qu'un système linéaire sous forme échelonnée réduite peut être résolu très facilement. Nous allons maintenant montrer :

Théorème I.2.19. *Soit (S) un système linéaire. Alors (S) peut être transformé, par des opérations élémentaires sur les lignes, en un système sous forme échelonnée réduite.*

Il est en fait plus commode de travailler en notation matricielle, et de prouver le théorème équivalent :

Théorème I.2.20. *Soit A une matrice. Alors A peut être transformée, par des opérations élémentaires sur les lignes, en une matrice échelonnée réduite.*

Notons que le théorème I.2.20, appliqué à la matrice augmentée du système (S), implique le théorème I.2.19.

Pour montrer le théorème I.2.20, on utilise une méthode appelée *méthode du pivot* ou *du pivot de Gauss*, qui suit exactement la stratégie ébauchée en I.1 pour la résolution des systèmes à deux équations, deux inconnues. Le nom de cette méthode est un hommage au mathématicien allemand Carl Friedrich Gauss (1777-1855), mais elle était déjà connue des mathématiciens chinois du 1er siècle de notre ère. D'après Wikipedia “Elle est référencée dans l'important livre chinois *Jiuzhang suanshu* ou *Les Neuf Chapitres sur l'art mathématique*, dont elle constitue le huitième chapitre, sous le titre *Fang cheng* (la disposition rectangulaire)”.

La méthode du pivot permet de démontrer les théorèmes I.2.19 et I.2.20, mais elle donne aussi et surtout une méthode générale de résolution des systèmes linéaires. Elle doit donc être parfaitement comprise.

Soit A une matrice (p, N) (p lignes, N colonnes). Pour montrer le théorème I.2.20, on veut transformer A en une matrice échelonnée réduite par une série d'opérations élémentaires sur les lignes. La méthode est divisée en une phase de descente (permettant d'obtenir une matrice échelonnée qui n'est pas forcément réduite) et une phase de remontée. Pour illustrer cette méthode, on l'applique à la matrice

$$A = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 2 & 4 & 8 & 14 \\ 1 & 0 & -2 & 3 \end{array}\right].$$

L'utilisation de la notation matricielle n'est pas obligatoire : on peut aussi résoudre un système en réalisant les opérations décrite plus bas directement sur ce système plutôt que sur sa matrice augmentée.

Phase de descente

Cette phase permet de transformer une matrice quelconque en une matrice échelonnée. Elle est divisée en 4 étapes, que l'on doit éventuellement réaliser plusieurs fois.

Etape 1 : choix du pivot. On appelle *colonne pivot* la première colonne non nulle [2]. On choisit sur cette colonne un élément non nul, appelé pivot. Dans notre exemple, la colonne pivot est la première colonne. On peut choisir comme élément pivot le 1 en haut à gauche (en gras).

$$A = \left[\begin{array}{ccc|c} \mathbf{1} & 2 & 4 & 7 \\ 2 & 4 & 8 & 14 \\ 1 & 0 & -2 & 3 \end{array}\right].$$

Etape 2. On échange la première ligne avec la ligne de l'élément pivot. Le pivot devient ainsi l'élément de tête de la première ligne. Dans notre exemple, l'élément pivot est déjà situé sur la première ligne : cette étape ne modifie pas la matrice.

Etape 3. En ajoutant aux autres lignes un multiple adéquat de la première ligne, on annule tous les coefficients de la colonne pivot autre que le pivot. Dans notre exemple, cela donne

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & -4 \end{array}\right] \begin{array}{l} \\ {\scriptstyle (L_2)-2(L_1)} \\ {\scriptstyle (L_3)-(L_1)} \end{array}.$$

Etape 4. Si la matrice obtenue est sous forme échelonnée, la phase de descente est terminée. Sinon, on applique les étapes 1 à 4 à la matrice à laquelle on a enlevé la première ligne.

Revenons à notre exemple. La matrice A a 3 lignes. On applique les étapes 1 à 4 à la matrice $A' = \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & -4 \end{array}\right]$ obtenue à partir de A en enlevant la première ligne. En pratique, on continue à écrire cette première ligne, que l'on ignore.

Étape 1' : On considère donc la matrice $\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{-2} & -6 & -4 \end{array}\right]$. La colonne pivot (première colonne non nulle lorsqu'on ignore la première ligne) est la deuxième colonne. On doit choisir comme pivot le -2, situé à la troisième ligne de cette colonne.

Étape 2' : on échange la 2ème et la 3ème ligne, la matrice obtenue est

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & \mathbf{-2} & -6 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{I.19}$$

Étape 3' : les coefficients de la colonne pivot (la deuxième colonne) autre que le pivot sont déjà nuls, il n'y a donc rien à faire. Rappelons que l'on ignore la première ligne. Il n'y a donc que le coefficient situé à la troisième ligne de cette deuxième colonne à considérer. Ce coefficient est bien égal à zéro.

Étape 4' : la matrice obtenue est échelonnée : on arrête ici la phase de descente.

Si A est une matrice échelonnée $p \times N$, et $(a_0, \ldots, a_N)$ sont $N+1$ scalaires tels que $a_0 \neq 0$, la matrice

$$B = \left[\begin{array}{cccc} a_0 & a_1 & \ldots & a_N \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{array}\right],$$

(ou tout autre matrice obtenue à partir de B en rajoutant des colonnes de zéros à sa gauche) est une matrice échelonnée. De cette remarque, et du fait que toute matrice ayant une seule ligne est échelonnée, on déduit que la phase de descente de la méthode du pivot aboutit bien, après un certain nombre [3] de passages par les étapes 1 à 4, à une matrice échelonnée.

Pour obtenir une matrice échelonnée réduite, on a besoin de deux étapes supplémentaires, qui constituent la phase de remontée.

Phase de remontée

Etape 5 : cadrage. On multiplie chaque ligne non nulle par l'inverse de son élément de tête, de telle manière que l'élément de tête de la nouvelle ligne vaut 1. Dans notre exemple :

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ {\scriptstyle -\frac{1}{2}(L_2)} \\ \\ \end{array}$$

Etape 6 : on ajoute à chaque ligne un multiple de la dernière ligne non nulle, pour que la colonne au-dessus de l'élément de tête de la dernière ligne ne soit composée que de zéros. On répète cette opération avec l'avant-dernière ligne, etc, jusqu'à la deuxième ligne. Sur notre exemple :

$$\left[\begin{array}{ccc|c} 1 & 0 & -2 & 3 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} {\scriptstyle (L_1)-2(L_2)} \\ \\ \\ \end{array} \tag{I.20}$$

Par construction, la matrice obtenue à l'issue de la phase de remontée est bien une matrice échelonnée réduite : on a transformé en 1 tous les éléments de tête, et en 0 tous les coefficients situés au dessus d'un élément de tête.

Ceci termine la démonstration du théorème I.2.20 (et son analogue sur les systèmes, le théorème I.2.19) et la description de la méthode du pivot.

Mentionnons que l'on peut en fait démontrer l'unicité de la forme échelonnée réduite :

2. c'est à dire la colonne la plus à gauche qui n'est pas composée uniquement de zéros

3. au plus $p-1$

Théorème I.2.21. *Soit A et B deux matrices échelonnées réduites. Supposons que B puisse être obtenue à partir de A par des opérations élémentaires sur les lignes. Alors $A = B$.*

On omet la démonstration, cf par exemple David C. Lay [4], annexe A.

Remarque I.2.22. On déduit de l'exemple donné une description paramétrique de l'ensemble des solutions du système :

$$\begin{cases} x + 2y + 4z = 7 \\ 2x + 4y + 8z = 14 \\ x - 2z = 3 \end{cases}$$

dont la matrice augmentée est la matrice A. Le système est compatible : il n'y a pas d'élément de tête sur la dernière colonne de la matrice échelonnée réduite obtenue en (I.20). Il y a deux variables de base, x et y, et une variable libre z. L'ensemble des solutions est donné par

$$\{(3 + 2z, 2 - 3z, z),\ z \in \mathbb{K}\}.$$

Remarquons que la compatibilité du système peut se vérifier à la fin de la phase de descente, sur la forme échelonnée non réduite (I.19).

En combinant la méthode du pivot de Gauss avec la description de l'ensemble des solutions d'un système sous forme échelonnée réduite (cf la proposition I.2.17), on obtient la propriété suivante, qui généralise la remarque I.1.12 à tous les systèmes linéaires :

Théorème I.2.23. *L'ensemble des solutions d'un système linéaire est vide, ou réduit à un seul élément, ou infini.*

Remarque I.2.24. La phase de remontée du pivot de Gauss montre que tout système sous forme échelonnée est équivalent à un système sous forme échelonnée réduite *avec le même nombre de lignes non nulles*. Si (S) est un système *compatible* sous forme échelonnée à n inconnues et p' équations non nulles, l'ensemble des solutions se décrit donc avec exactement $n - p'$ paramètres.

Exercice I.2.25. Combien de paramètres faut-il pour décrire chacun des systèmes suivants ?

$$(S_1) \begin{cases} x + y + z = 1 \\ y + 2z = 2 \end{cases} \qquad (S_2) \begin{cases} 2x + 3y = 4 \\ y + z = 1 \\ z = 2 \end{cases} \qquad (S_3) \begin{cases} x + (1+i)y + z = 2 \\ x + y = 1 \\ iy + z = 1 \end{cases}$$

(cf correction p. 14).

4. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

Remarque I.2.26. Il y a plusieurs variantes (parfaitement équivalentes) de la méthode du pivot que nous venons de décrire. On peut par exemple échanger les étapes 5 et 6. On peut aussi réaliser l'étape de cadrage 5 pendant la phase de descente. Dans tous les cas il est important de n'utiliser que des opérations élémentaires sur les lignes (cf l'avertissement I.2.6). Ceci est particulièrement important lorsque l'on cherche à résoudre des systèmes avec paramètres (cf §I.3 p. 13).

Récapitulons la méthode générale de résolution d'un système linéaire décrite dans ce chapitre :

— Appliquer la phase de descente de la méthode du pivot de Gauss à la matrice augmentée du système. On obtient une matrice échelonnée.

— Déterminer si le système est compatible : si la colonne de droite contient un élément de tête, le système n'est pas compatible (i.e. il n'y a pas de solution). Sinon il est compatible.

— Si le système est compatible, appliquer la phase de remontée du pivot de Gauss. On obtient une matrice échelonnée réduite. On peut alors donner une description paramétrique de l'ensemble des solutions à l'aide de cette matrice échelonnée réduite.

Donnons un autre exemple, cette fois avec des coefficients complexes. On considère le système :

$$(S) \begin{cases} 4z_1 - (3+i)z_2 - (9+3i)z_3 = 5 - 3i \\ 2z_1 - 2z_2 - 6z_3 = 2 - 2i \\ 4z_1 - (2+2i)z_2 - (6+6i)z_3 = 6 - 2i \end{cases}$$

Appliquons la méthode du pivot à sa matrice augmentée :

$$\left[\begin{array}{rrr|r} 4 & -3-i & -9-3i & 5-3i \\ \mathbf{2} & -2 & -6 & 2-2i \\ 4 & -2-2i & -6-6i & 6-2i \end{array}\right]$$

Phase de descente

Etape 1 : choix du pivot. La colonne pivot est la première colonne. Le pivot est le 2 en gras.

Etape 2 : on place la ligne du pivot en première ligne (opération $(L_1) \leftrightarrow (L_2)$)

$$\left[\begin{array}{rrr|r} \mathbf{2} & -2 & -6 & 2-2i \\ 4 & -3-i & -9-3i & 5-3i \\ 4 & -2-2i & -6-6i & 6-2i \end{array}\right]$$

Etape 3 :

$$\begin{array}{c} \left[\begin{array}{rrr|r} 2 & -2 & -6 & 2-2i \\ 0 & 1-i & 3-3i & 1+i \\ 0 & 2-2i & 6-6i & 2+2i \end{array}\right] \end{array} \begin{array}{l} \\ {\scriptstyle (L_2)-2(L_1)} \\ {\scriptstyle (L_3)-2(L_1)} \end{array}$$

Etape 4 : on repasse à l'étape 1 en ignorant la première ligne.

Etape 1' : la colonne pivot est la deuxième colonne. On choisit le $1-i$ sur la deuxième ligne de cette colonne comme élément pivot.

Etape 2' : la ligne pivot est la deuxième ligne donc la "première" si on ignore la première ligne. Il n'y a rien à faire.

Etape 3' :

$$\left[\begin{array}{ccc|c} 2 & -2 & -6 & 2-2i \\ 0 & 1-i & 3-3i & 1+i \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \begin{array}{l} \\ \\ {\scriptstyle (L_3)-2(L_2)} \end{array}$$

Etape 4' : la matrice obtenue est échelonnée. La phase de descente est terminée. On remarque que la colonne de droite ne contient aucun élément de tête : le système est compatible. On passe à la phase de remontée pour obtenir une matrice échelonnée réduite.

Phase de remontée

Etape 5 :

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & 1-i \\ 0 & 1 & 3 & i \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \begin{array}{l} {\scriptstyle \frac{1}{2}(L_1)} \\ {\scriptstyle \frac{1}{1-i}(L_2)} \\ \\ \end{array}$$

Etape 6 : on utilise la ligne (L_2) pour annuler l'élément de la deuxième colonne au-dessus de l'élément de tête :

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 3 & i \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \begin{array}{l} {\scriptstyle (L_1)+(L_2)} \\ \\ \\ \end{array}$$

La matrice obtenue est sous forme échelonnée réduite. Il y a une variable libre, z_3, et deux variables de base, z_1 et z_2. Une description paramétrique de l'ensemble des solutions est donnée par :

$$\{(1, i-3z, z),\ z \in \mathbb{C}\}.$$

Exercice I.2.27. Résoudre le système

$$\begin{cases} 2x-2y-4z=-12 \\ -x+y+3z=9 \\ x-y+z=3. \end{cases}$$

I.2.d. Système de Cramer

Les coefficients (a_{ij}) d'un système linéaire (S) étant fixés, la compatibilité de (S) dépend en général de son second membre (b_i). On présente ici un cas particulier important, appelé *système de Cramer* pour lequel le système est compatible quelques soient les b_i.

Proposition I.2.28. *Soit (S) un système linéaire de n équations à n inconnues. Supposons que (S) a une et une seule solution. Alors tout système obtenu à partir de (S) en changeant seulement le second membre a une seule solution.*

Preuve de la proposition I.2.28. Par des opérations élémentaires sur les lignes, on peut, d'après la méthode du pivot de Gauss, se ramener à un système sous forme échelonnée réduite (S') ayant le même ensemble de solutions que (S). Soit p' le nombre de lignes non nulles de ce système. Le système étant compatible, le nombre de paramètres permettant de décrire l'ensemble des solutions est, d'après la proposition I.2.17, $n-p'$. Mais on ne peut pas avoir $n-p' \geq 1$, sinon l'ensemble des solutions serait infini. On a donc $n=p'$: la forme échelonnée réduite du système a n lignes non nulles, aucune ligne nulle (car il y a n lignes en tout), et s'écrit donc :

$$\begin{cases} x_1 = b'_1 \\ x_2 = b'_2 \\ \\ x_n = b'_n. \end{cases} \tag{I.21}$$

Lorsque l'on change le membre de droite du système (S) sans toucher au membre de gauche, on ne change que le membre de droite du système (I.21), puisque ce dernier est obtenu par des opérations élémentaires sur les lignes qui ne mélangent jamais le membre de gauche et le membre de droite des équations. Ceci montre que tout système obtenu à partir de (S) en ne changeant que le membre de droite n'a qu'une seule solution. □

Définition I.2.29. Un système (S) vérifiant les hypothèses de la proposition I.2.28 est appelé *système de Cramer*. Un système de Cramer est donc par définition un système linéaire ayant autant d'inconnues que d'équations et une unique solution. De manière équivalente, c'est un système à n équations et n inconnues qui a une forme échelonnée réduite du type (I.21).

Remarque I.2.30. D'après la proposition I.2.28, le fait d'être un système de Cramer ne dépend pas du second membre du système, mais seulement de ses coefficients.

Exercice I.2.31. Dire *avec le moins de calculs possibles* si les systèmes suivants ont une unique solution, pas de solution ou une infinité de solutions. On identifiera en particulier les systèmes homogènes et les systèmes de Cramer. Les systèmes (S_1), (S_2) et (S_3) ont pour inconnues x, y et z. Les systèmes (S_4) et (S_5) x, y, z et t.

$$(S_1)\begin{cases} x+3y+z=2 \\ x+y+2z=-5 \\ x+3y+2z=2 \end{cases}, \quad (S_2)\begin{cases} x+3y+z=4 \\ x+y+2z=1 \\ x+3y+2z=3 \end{cases}, \quad (S_3)\begin{cases} x+3y+z=-17 \\ x+y+2z=2 \\ x+3y+2z=5 \end{cases}$$

$$(S_4)\begin{cases} x+2y+z+t=0 \\ x+y-z+5t=0 \\ 2x+y+2z+4t=0 \end{cases}, \quad (S_5)\quad x=2y+2t=3z+4(x+y)=6x+y+z+t.$$

(cf correction p. 14).

I.3. Systèmes avec paramètres

On considère parfois une famille de systèmes (S_λ) dépendant d'un paramètre λ. Le but est de résoudre le système selon les valeurs de λ. Il faut traiter ce type de problème avec précaution. Une erreur classique est de diviser une équation par une quantité, dépendant de λ, qui s'annule pour une certaine valeur de λ. On donne ici deux exemples de résolutions détaillées.

Exercice I.3.1. Résoudre, selon la valeur de $\lambda \in \mathbb{R}$, le système suivant :

$$\left\{\begin{aligned} -x+(2\lambda-6)y-2z&=-7\\ -x+(4\lambda-12)y-2z&=-11\\ x+(3-\lambda)y+2z&=5 \end{aligned}\right. \tag{I.22}$$

Correction. On applique la méthode du pivot de Gauss.

$$\left\{\begin{aligned} x+(3-\lambda)y+2z&=5 && (L_3)\\ -x+(4\lambda-12)y-2z&=-11\\ -x+(2\lambda-6)y-2z&=-7 && (L_1) \end{aligned}\right.$$

$$\left\{\begin{aligned} &x+(3-\lambda)y+2z=5\\ &\qquad(-9+3\lambda)y=-6 && (L_2)+(L_1)\\ &\qquad(-3+\lambda)y=-2 && (L_3)+(L_1) \end{aligned}\right.$$

On remarque que les lignes (L_2) et (L_3) du dernier système sont équivalentes. Précisément, (L_2) vaut exactement $3(L_3)$. Le système (I.22) est donc équivalent à

$$\left\{\begin{aligned} &x+(3-\lambda)y+2z=5\\ &\qquad(-3+\lambda)y=-2 \end{aligned}\right. \tag{I.23}$$

Ce système est sous forme échelonnée. En regardant le coefficient de y dans la deuxième ligne, on voit qu'il faut distinguer deux cas.

1er cas : $\lambda=3$. La deuxième ligne de (I.23) s'écrit $0=-2$. Le système n'a pas de solution.

2ème cas : $\lambda\neq 3$ On obtient :

$$\left\{\begin{aligned} x+2z&=3 && (L_1)+(L_2)\\ (-3+\lambda)y&=-2 \end{aligned}\right. .$$

En prenant x et y comme variable de base et z comme paramètre, on obtient que l'ensemble des solutions est $\left\{(3-2z,\frac{2}{3-\lambda},z),\ z\in\mathbb{R}\right\}$.

Exercice I.3.2. Résoudre, selon la valeur de $\lambda\in\mathbb{R}$, le système de matrice augmentée :

$$\left[\begin{array}{rr|r} 1 & -(1+\lambda) & \lambda-4\\ -2 & 2+2\lambda & 12-4\lambda\\ 2 & -2\lambda & -4+2\lambda \end{array}\right] \tag{I.24}$$

Correction.

$$\left[\begin{array}{rr|r} 1 & -(1+\lambda) & \lambda-4\\ 0 & 0 & 4-2\lambda\\ 0 & 2 & 4 \end{array}\right] \quad \begin{array}{l} \\ {\scriptstyle (L_2)+2(L_1)}\\ {\scriptstyle (L_3)-2(L_1)} \end{array}$$

Le système obtenu est presque sous forme échelonnée (il suffirait d'échanger les lignes 2 et 3). La ligne 2 se lit $4-2\lambda=0$. On distingue deux cas :

1er cas : $\lambda\neq 2$. La ligne 2 est contradictoire. Le système n'est pas compatible.

2ème cas : $\lambda=2$. La ligne 2 se lit $0=0$. En notant x et y les variables, le système est équivalent à :

$$x-3y=-2 \text{ et } 2y=4,$$

soit $(x,y)=(4,2)$.

Remarquons que l'on peut parfois ramener un système non-linéaire à un système linéaire avec paramètre :

Exercice I.3.3. En utilisant les exercices précédents, résoudre les systèmes non-linéaires :

$$(S_1)\left\{\begin{aligned} -x_1+(2x_2-6)x_3-2x_4&=-7\\ -x_1+(4x_2-12)x_3-2x_4&=-11\\ x_1+(3-x_2)x_3+2x_4&=5 \end{aligned}\right. \qquad (S_2)\left\{\begin{aligned} x-(1+z)y&=z-4\\ -2x+(2+2z)y&=12-4z\\ 2x-2zy&=-4+2z \end{aligned}\right.$$

(cf correction p. 14).

I.4. Réponse à certains exercices

Exercice I.2.13

i. Une matrice à une ligne est toujours sous forme échelonnée. Elle est sous forme échelonnée réduite si et seulement si son premier coefficient non nul vaut 1. L'équation (ou le "système" à une équation) correspondant(e) est sous forme échelonnée réduite si et seulement si le coefficient de la première variable qui apparaît dans le système vaut 1.

ii. La matrice du premier système est $\left[\begin{array}{ccc|c} 0 & 1 & 1 & 2\\ 1 & 0 & 1 & 3 \end{array}\right]$, celle du deuxième système est $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2\\ 0 & 1 & 1 & 3 \end{array}\right]$. Le premier n'est pas sous forme échelonnée. Le deuxième est sous forme échelonnée (réduite). Remarquons que la deuxième matrice est obtenue à partir de la première en échangeant les 2 premières colonnes. Cela revient, sur le système correspondant, à échanger les variables x_1 et x_2.

iii. échelonnée non réduite/ non échelonnée / échelonnée réduite.

iv. échelonné réduit/ échelonné non réduit / non échelonné.

Exercice I.2.16

ii. L'ensemble des solutions du deuxième système est donné par : $\Big\{(2-x_3, 3-x_3, x_3),\ x_3 \in \mathbb{K}\Big\}$.

iii. L'ensemble des solutions du système dont la matrice augmentée est la troisième matrice est donné par $\Big\{(x_1, x_2, 1-2x_5-3x_6, -2+3x_6, x_5, x_6),\ (x_1, x_2, x_5, x_6) \in \mathbb{K}^4\Big\}$.

iv. Le système (S_1) a évidemment pour unique solution $(1, 2, 3)$.

Exercice I.2.18

La matrice a) est sous forme échelonnée. La dernière colonne ne contient pas d'élément de tête : le système est compatible. Pour décrire l'ensemble des solutions, ont peut facilement mettre la matrice sous forme échelonnée réduite (par le remplacement $(L_1) \leftarrow (L_1) + (L_2)$). On obtient la matrice (en omettant la dernière ligne, inutile, qui signifie $0 = 0$) :

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right].$$

Les variables de base sont x_1 et x_4, les variables libres x_2 et x_3. Une description paramétrique de l'ensemble des solutions est donnée par :

$$\Big\{(4-2x_2-3x_3, x_2, x_3, 2),\ (x_2, x_3) \in \mathbb{K}^2\Big\}.$$

La matrice b) n'est pas sous forme échelonnée. La dernière ligne de cette matrice signifie $3 = 0$: le système n'est pas compatible.

La matrice c) est échelonnée réduite. L'ensemble des solutions est :

$$\Big\{(1+3x_4, 2-3x_4, -4x_4, x_4),\ x_4 \in \mathbb{K}\Big\}.$$

Exercice I.2.25

Les systèmes (S_1) et (S_2) sont sous forme échelonnée et compatibles (ils ne contiennent pas de ligne de la forme $0 = c$ avec c non nulle). Le système (S_1) a 3 inconnues et 2 lignes non nulles, on décrit l'ensemble des solutions avec $3 - 2 = 1$ paramètre. Le système (S_2) a 3 inconnues et 3 équations : on décrit l'ensemble des solutions avec $3 - 3 = 0$ paramètre : il a une unique solution (en l'occurrence $(x, y, z) = (7/2, -1, 2)$).

Le système (S_3) n'est pas sous forme échelonnée. De fait, la ligne (L_1) est la somme des lignes (L_2) et (L_3). Il est équivalent au système (S_3') obtenu en retirant la ligne (L_3) au système (S_3). Ce système (S_3') est échelonné, on peut décrire l'ensemble des solutions de (S_3) avec $3 - 2 = 1$ paramètre.

Exercice I.2.27 : on montre en utilisant la méthode du pivot de Gauss qu'il y a une infinité de solutions et que l'ensemble des solutions peut s'écrire $\{(x, x, 3),\ x \in \mathbb{K}\}$.

Exercice I.2.31

Par les remplacements $(L_2) \leftarrow (L_2) - (L_1)$ et $(L_3) \leftarrow (L_3) - (L_1)$, on voit que le système (S_1) est équivalent au système

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

Ce dernier système est un système échelonné compatible à trois équations non nulles pour trois inconnues : c'est donc un système de Cramer, qui a une unique solution. De plus, par la Proposition I.2.28 sur les systèmes de Cramer, tout système obtenu à partir de (S_1) en ne changeant que le second membre sont aussi des systèmes de Cramer : on en déduit que (S_2) et (S_3) sont des systèmes de Cramer (et ont donc chacun une et une seule solution).

Le système (S_4) est un système homogène avec 3 équations et 4 inconnues. Il a donc une infinité de solutions. Le système (S_5) est équivalent à :

$$\begin{cases} x - (2y + 2t) = 0 \\ x - 3z - 4(x + y) = 0 \\ x - (6x + y + z + t) = 0 \end{cases}$$

C'est donc aussi un système homogène à 3 équations et 4 inconnues : il a une infinité de solutions.

Exercice I.3.3

Les deux systèmes proposés *ne sont pas* des systèmes linéaires, mais se ramènent à des systèmes linéaires en fixant une des variables.

Si l'on fixe x_2, le système (S_1) est un système linéaire d'inconnues (x_1, x_3, x_4). Plus précisément, c'est exactement le système (I.22) avec $x = x_1$, $y = x_3$, $z = x_4$ et le paramètre $\lambda = x_2$. L'ensemble des solutions est donné par la résolution du système (I.22) :

$$\left\{\left(3 - 2x_4, x_2, \frac{2}{3 - x_2}, x_4\right),\ x_2 \in \mathbb{R} \setminus \{3\},\ x_4 \in \mathbb{R}\right\}.$$

Si l'on fixe z, le système (S_2) est un système linéaire. Plus précisément, c'est le système d'inconnue (x, y), dont la matrice augmentée est donnée par (I.24) avec $z = \lambda$. On déduit de la résolution de ce système que l'unique solution du système (S_2) est $(x, y, z) = (4, 2, 2)$.

II. Introduction aux matrices

Référence : Liret-Martinais [1], chapitre 4.

Nous avons déjà rencontré des tableaux de nombres, ou matrices. Nous allons étudier ici ces matrices de manière plus systématique, en définissant notamment des opérations (additions, multiplications...) sur des ensembles de matrices. Les motivations de ce chapitre sont d'une part de mieux comprendre les systèmes linéaires, qui peuvent être vus comme des équations matricielles, d'autre part d'introduire des notions et des méthodes utiles pour l'algèbre linéaire proprement dite qui sera l'objet de la suite du cours (à partir du chapitre V sur les espaces vectoriels).

Comme dans le chapitre précédent, on notera $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

II.1. Définitions. Opérations sur les matrices

II.1.a. Définitions

Définition II.1.1. Soient p et n deux entiers ≥ 1. Une matrice $p \times n$ à coefficients dans $\mathbb{K}$ est un tableau de nombres (c.à.d. d'éléments de $\mathbb{K}$) à p lignes et n colonnes, que l'on note :

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{bmatrix},$$

ou $A = [a_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}$, ou encore, quand il n'y a pas d'ambiguité sur p et n, $A = [a_{ij}]$.
Les $a_{ij} \in \mathbb{K}$ sont appelés *coefficients* (ou *éléments*) de A. Le coefficient a_{ij} est situé à la $i-$ème ligne et $j-$ième colonne (le premier indice indique toujours la ligne, le deuxième la colonne).

L'ensemble des matrices $p \times n$ est noté $\mathcal{M}_{p,n}(\mathbb{K})$, ou plus simplement $\mathcal{M}_{p,n}$. Lorsque $p = n$, on dit que la matrice est *carrée*. On note simplement $\mathcal{M}_n(\mathbb{K})$ (ou $\mathcal{M}_n$) au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$ l'ensemble des matrices carrées de taille n. On dit que deux matrices sont *de même taille*, ou *de même dimension* lorsqu'elles ont le même nombre de lignes et le même nombre de colonnes.

Exemples II.1.2. La *matrice nulle* de $\mathcal{M}_{p,n}(\mathbb{K})$ est la matrice dont tous les coefficients sont nuls. Elle est notée 0, ou plus précisément $0_{p,n}$ quand on veut préciser le nombre de lignes et de colonnes.

$\begin{bmatrix} 1 & i & -5 \\ 3 & 4 & 7+i \end{bmatrix}$ est une matrice complexe 2×3 (*attention ici i est le nombre complexe tel que $i^2 = -1$, ce n'est pas l'indice des lignes !*)

$\begin{bmatrix} -1 & 0 \\ 7 & 3,1 \\ 2 & \sqrt{3} \end{bmatrix}$ est une matrice réelle 3×2.

Définition II.1.3. On dit que deux matrices *de même taille* $A = [a_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}$ et $B = [b_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}$ sont égales lorque tous leur coefficients sont égaux, i.e. $\forall i \in \{1,\dots,p\}$, $\forall j \in \{1,\dots,n\}$, $a_{i,j} = b_{i,j}$. On note alors $A = B$.

Exemples II.1.4. Les matrices $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ et $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ ne sont pas égales.

Les matrices $[2i+j]_{\substack{1\leq i\leq 3\\1\leq j\leq 2}}$ et $\begin{bmatrix} 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}$ sont égales.

Définition II.1.5. Une *matrice colonne* (ou un vecteur colonne) à p coefficients est un élément de $\mathcal{M}_{p,1}(\mathbb{K})$.

Une *matrice ligne* (ou un vecteur ligne) à n coefficients est un éléments de $\mathcal{M}_{1,n}(\mathbb{K})$.

La j-ième colonne de la matrice $[a_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}$ est la matrice colonne $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{pj} \end{bmatrix}$, sa i-ème ligne la matrice ligne $\begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{bmatrix}$.

II.1.b. Multiplication par un scalaire et additions

On définit maintenant deux opérations qui se font *coefficient par coefficient*.

Définition II.1.6. Soit $\lambda \in \mathbb{K}$ et $A \in \mathcal{M}_{p,n}(\mathbb{K})$. Le produit de A par λ, noté λA, est la matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ obtenue en multipliant chaque coefficient de A par λ : si A est la matrice $[a_{ij}]$, λA est la matrice $[\lambda a_{ij}]$.

Définition II.1.7. Soient $A, B \in \mathcal{M}_{p,n}(\mathbb{K})$ deux matrices. La somme de A et B, notée $A + B$ est la matrice de $\mathcal{M}_{p,n}$ obtenue en sommant les coefficients de A et de B deux à deux : si $A = [a_{ij}]$ et $B = [b_{ij}]$, $A + B$ est la matrice $[a_{ij} + b_{ij}]$.

Avertissement II.1.8. La somme de deux matrices n'est définie que si ces matrices sont *de même taille*.

1. François Liret et Dominique Martinais. *Algèbre 1re année - Cours et exercices avec solutions.* Dunod, deuxième édition, 2003

Exemples II.1.9.

$$17\begin{bmatrix}0 & -2\\0 & 1\end{bmatrix}=\begin{bmatrix}0 & -34\\0 & 17\end{bmatrix},\qquad (3+i)\begin{bmatrix}1 & 3-i\\i & -1\\0 & -i\end{bmatrix}=\begin{bmatrix}3+i & 10\\3i-1 & -3-i\\0 & 1-3i\end{bmatrix}$$

$$\begin{bmatrix}1 & -1\\2 & 3\end{bmatrix}+\begin{bmatrix}4 & 7\\-2 & 2\end{bmatrix}=\begin{bmatrix}5 & 6\\0 & 5\end{bmatrix},\qquad \begin{bmatrix}1 & -1\\2 & 3\end{bmatrix}+\begin{bmatrix}4 & 7\\2 & -2\\3 & 0\end{bmatrix}\text{ n'est pas définie.}$$

$$[i+j]_{\substack{1\leq i\leq 3\\1\leq j\leq 2}}+[2i-j]_{\substack{1\leq i\leq 3\\1\leq j\leq 2}}=[3i]_{\substack{1\leq i\leq 3\\1\leq j\leq 2}}.$$

Les propriétés suivantes découlent immédiatement des propriétés (commutativité, associativité, distributivité) de l'addition et de la multiplication sur $\mathbb{K}$.

Proposition II.1.10. *Soient $A, B\in\mathcal{M}_{p,n}(\mathbb{K})$, $\lambda,\mu\in\mathbb{K}$.*

i. *(commutativité)* $A+B=B+A$.

ii. *(associativité)* $(A+B)+C=A+(B+C)$ *(on note leur valeur commune $A+B+C$).*

iii. $\lambda(A+B)=\lambda A+\lambda B$.

iv. $\lambda(\mu A)=(\lambda\mu)A$ *(on note la valeur commune $\lambda\mu A$).*

v. $0+A=A$.

vi. $A+(-1)A=0$.

Exercice II.1.11. Démontrer les propriétés de la proposition.

Notation II.1.12. On note $-A$ la matrice $(-1)A$ et $A-B$ la somme $A+(-B)$, appelée *différence* de A et B.

II.1.c. Transposition

Définition II.1.13. La transposée de la matrice $A=[a_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}\in\mathcal{M}_{p,n}(\mathbb{K})$ est la matrice $[a_{ji}]_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ de $\mathcal{M}_{n,p}(\mathbb{K})$. On la note ${}^tA$. Les coefficients de la i-ème ligne de ${}^tA$ sont ceux de la i-ème colonne de A, et inversement, les coefficients de la j-ème colonne de ${}^tA$ sont ceux de la j-ème ligne de A.

On rencontre aussi, en particulier dans les ouvrages en anglais, la notation A^T au lieu de ${}^tA$.

Avertissement II.1.14. Lorsqu'on transpose une matrice, on inverse le nombre de lignes et le nombre de colonnes. Par exemple, la transposée d'une matrice ligne est une matrice colonne et la transposée d'une matrice colonne est une matrice ligne. La transposée d'une matrice 2×4 est une matrice 4×2 etc...

Exemples II.1.15.

$${}^t\begin{bmatrix}1 & -2\\3 & 4\\0 & 5\end{bmatrix}=\begin{bmatrix}1 & 3 & 0\\-2 & 4 & 5\end{bmatrix},\qquad {}^t\begin{bmatrix}2\\3\\4\end{bmatrix}=\begin{bmatrix}2 & 3 & 4\end{bmatrix}\qquad {}^t\begin{bmatrix}1 & 3\\3 & 2\end{bmatrix}=\begin{bmatrix}1 & 3\\3 & 2\end{bmatrix}.$$

Le dernier exemple de matrice A est une matrice carrée qui vérifie $A={}^tA$: on dit que A est *symétrique*.

On déduit immédiatement de la définition de la transposée :

Proposition II.1.16. *Soient $A,B\in\mathcal{M}_{p,n}(\mathbb{K})$, $\lambda\in\mathbb{K}$.*

$${}^t({}^tA)=A,\qquad {}^t(A+B)={}^tA+{}^tB,\qquad {}^t(\lambda A)=\lambda\,{}^tA.$$

II.1.d. Multiplication des matrices

La définition de la multiplication est plus délicate que celle des opérations précédentes, et nous allons la diviser en deux étapes. Soulignons que contrairement à l'addition, il ne s'agit pas d'une opération coefficient par coefficient.

Multiplication d'une matrice ligne par une matrice colonne

Définition II.1.17. Soit $A=[a_i]_{1\leq i\leq n}\in\mathcal{M}_{1,n}(\mathbb{K})$ une matrice ligne et $B=[b_j]_{1\leq j\leq n}\in\mathcal{M}_{n,1}(\mathbb{K})$ une matrice colonne *ayant le même nombre de coefficients*. Le *produit* AB de A et B est le scalaire :

$$AB=\sum_{j=1}^{n}a_jb_j=a_1b_1+a_2b_2+\ldots+a_nb_n.$$

Avertissement II.1.18. On ne peut pour l'instant multiplier qu'un vecteur ligne et un vecteur colonne *ayant le même nombre de coefficients*, et dans cet ordre.

Exemple II.1.19.

$$\begin{bmatrix}3 & 4 & 5\end{bmatrix}\begin{bmatrix}-1\\i\\2\end{bmatrix}=-3+4i+10=7+4i.$$

Cas général

Définition II.1.20. Soit $A=[a_{ij}]_{\substack{1\leq i\leq p\\1\leq j\leq n}}\in\mathcal{M}_{p,n}(\mathbb{K})$ et $B=[b_{ij}]_{\substack{1\leq i\leq n\\1\leq j\leq q}}\in\mathcal{M}_{n,q}(\mathbb{K})$ deux matrices telles que le nombre de colonnes de A est égal au nombre de lignes de B. Le *produit* de A et B, noté AB, $A.B$ ou $A\times B$ est la matrice de $\mathcal{M}_{p,q}(\mathbb{K})$ dont le coefficient (i,j) est le produit (au sens de la définition II.1.17) de la i-ème ligne de A par la j-ième colonne de B. Le produit $A\times A$ d'une matrice carrée A est noté A^2 et appelé *carré* de A. On définit de même la puissance n-ième $A^n=A\times\ldots\times A$ (n fois) d'une matrice carrée A.

Remarque II.1.21. Il faut savoir déduire de la définition du produit matricielle une fonction en C retournant le produit de deux matrices données (cf les travaux dirigés et travaux pratiques d'algorithmique).

Remarque II.1.22. On déduit immédiatement de la définition la formule du produit matriciel suivante : si $AB = [c_{ij}]_{\substack{1\leq i\leq p\\ 1\leq j\leq q}}$, alors

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

Avertissement II.1.23. La matrice AB n'est définie que lorsque le nombre de colonnes de A est égal au nombre de lignes de B. Le nombre de lignes de AB est le nombre de lignes de A. Le nombre de colonnes de AB est le nombre de colonnes de B.

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}.$$

$$(\underline{3}, \mathbf{4}) \qquad (\mathbf{4}, 2) \qquad (\underline{3}, 2).$$

Exemple II.1.24. Pour toute matrice A, $0.A = A.0 = 0$. Ici 0 désigne n'importe quelle matrice nulle de dimension appropriée. Attention : les trois 0 ne désignent pas forcément les mêmes matrices nulles ! On peut écrire plus précisément, si $A \in \mathcal{M}_{p,n}$

$$0_{q,p}A = 0_{q,n}, \quad A0_{n,q} = 0_{p,q}.$$

Exemples II.1.25.

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 2 & -1 \end{bmatrix} \text{ n'est pas défini.}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 1 \\ 2 & -3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -9 & 0 & 2 \\ 6 & -15 & 0 & 4 \\ -9 & 3 & 0 & 7 \end{bmatrix}.$$

$$\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} 2 & 1 & -7 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & -21 & 3 \\ 8 & 4 & -28 & 4 \\ 10 & 5 & -35 & 5 \end{bmatrix}$$

En pratique, pour calculer le produit $[c_{ij}]$ de deux matrices A et B, on peut les disposer de telle manière que le coefficient c_{ij} à calculer soit aligné sur la i-ème ligne de A et la j-ième colonne de B :

$$\begin{array}{cc} & \begin{bmatrix} -1 & 0 & 0 & 1 \\ 2 & -3 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} & \begin{bmatrix} 4 & -9 & 0 & 2 \\ 6 & -15 & 0 & 4 \\ -9 & 3 & 0 & 7 \end{bmatrix}. \end{array}$$

Exercice II.1.26. On considère les matrices

$$A = \begin{bmatrix} 1 & 2 \\ -3 & i \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1+i & -3 \\ 2 & 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ -3 & -6 \\ -2 & -4 \end{bmatrix}.$$

Donner les valeurs des produits et des carrés de ces matrices (16 opérations potentielles) lorsqu'ils sont définis.

(cf réponses p. 26).

Soit n un entier ≥ 1. On note I_n la matrice de $\mathcal{M}_n(\mathbb{K})$ dont la diagonale principale est composée de 1 et dont tous les autres coefficients sont nuls :

$$I_n = [\delta_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}, \quad \delta_{ii} = 1, \quad i \neq j \Rightarrow \delta_{ij} = 0. \tag{II.2}$$

Le symbole δ_{ij} est appelé *symbole de Kronecker.*

Exemple II.1.27.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Proposition II.1.28. *Soient $A \in \mathcal{M}_{p,n}(\mathbb{K})$, alors*

$$AI_n = I_pA = A.$$

Exercice II.1.29. Démontrer la proposition précédente à l'aide de la formule (II.1).

Définition II.1.30. La matrice I_n est appelée *matrice identité.*

On donne maintenant les propriétés de base de la multiplication matricielle :

Théorème II.1.31. *Soient A, B et C des matrices.*

i. Associativité : si $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$ et $C \in \mathcal{M}_{q,r}$,

$$(AB)C = A(BC).$$

On note simplement ABC le produit des trois matrices.

ii. Distributivité à gauche : si $A \in \mathcal{M}_{p,n}$ et $B, C \in \mathcal{M}_{n,q}$,

$$A(B + C) = AB + AC.$$

iii. Distributivité à droite : si $A, B \in \mathcal{M}_{p,n}$ et $C \in \mathcal{M}_{n,q}$,

$$(A + B)C = AC + BC.$$

iv. Si $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$ et $\lambda \in \mathbb{K}$,

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

v. *Si* $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$,
$$^t(AB) = {}^tB\,{}^tA.$$

Démonstration. On démontre (i). Les preuves (plus simples) des autres propriétés sont laissées au lecteur. Remarquons que les matrices $(AB)C$ et $A(BC)$ sont bien de même taille $p \times r$. On note

$$AB = [d_{ij}]_{\substack{1\leq i\leq p\\ 1\leq j\leq q}},\quad (AB)C = [e_{ij}]_{\substack{1\leq i\leq p\\ 1\leq j\leq r}},\quad BC = [f_{ij}]_{\substack{1\leq i\leq n\\ 1\leq j\leq r}} \text{ et } A(BC) = [g_{ij}]_{\substack{1\leq i\leq p\\ 1\leq j\leq r}}.$$

Notre but est de montrer

$$\forall i \in \{1,\ldots,p\},\ \forall j \in \{1,\ldots,r\},\quad e_{ij} = g_{ij}. \tag{II.3}$$

Par la formule du produit matriciel (II.1),

$$d_{ij} = \sum_{\ell=1}^{n} a_{i\ell}b_{\ell j},\qquad e_{ij} = \sum_{k=1}^{q} d_{ik}c_{kj} = \sum_{k=1}^{q}\left(\sum_{\ell=1}^{n} a_{i\ell}b_{\ell k}\right)c_{kj} = \sum_{k=1}^{q}\sum_{\ell=1}^{n}(a_{i\ell}b_{\ell k}c_{kj})$$
$$f_{ij} = \sum_{k=1}^{q} b_{ik}c_{kj},\qquad g_{ij} = \sum_{\ell=1}^{n} a_{i\ell}f_{\ell j} = \sum_{\ell=1}^{n} a_{i\ell}\sum_{k=1}^{q} b_{\ell k}c_{kj} = \sum_{k=1}^{q}\sum_{\ell=1}^{n}(a_{i\ell}b_{\ell k}c_{kj}),$$

d'où (II.3). □

Exercice II.1.32. Refaire le calcul précédent lorsque $p = q = r = n = 2$, sans utiliser le symbole $\sum$.

Exercice II.1.33. Démontrer les propriétés (ii), (iii), (iv), (v) du théorème.

Remarque II.1.34. Soit $A = [a]$ et $B = [b]$ des matrices 1×1, avec $a \in \mathbb{K}$ et $b \in \mathbb{K}$. Alors $A + B = [a + b]$ et $AB = ab$. L'ensemble $\mathcal{M}_1(\mathbb{K})$ muni de l'addition et de la multiplication matricielles s'identifie à $\mathbb{K}$, muni de l'addition et de la multiplication usuelles.

On termine cette partie par une mise en garde :

Avertissement II.1.35. Soient $a, b \in \mathbb{K}$. Alors

$$ab = 0 \Longrightarrow a = 0 \text{ ou } b = 0 \quad \text{(régularité de la multiplication scalaire)} \tag{II.4}$$
$$ab = ba \quad \text{(commutativité de la multiplication scalaire).} \tag{II.5}$$

Les propriétés analogues pour la multiplication matricielle sont fausses en générale (sauf pour les matrices 1×1 qui sont des scalaires !) Par exemple :

$$\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} = 0$$
$$\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \neq \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}.$$

II.1.e. Systèmes linéaires et matrices

La multiplication matricielle permet une nouvelle interprétation des systèmes linéaires. Considérons un système linéaire à p équations et n inconnues :

$$\text{(S)}\qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2\\ \quad\vdots \qquad\qquad\qquad \vdots \qquad\quad \vdots\\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n = b_p.\end{cases}$$

Soit $A = [a_{ij}]_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$ la matrice des coefficients de (S). On note

$$B = \begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_p\end{bmatrix},\quad X = \begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}.$$

Alors le système (S) est équivalent à l'équation matricielle :

$$AX = B.$$

Exemple II.1.36. Le système linéaire

$$\begin{cases} 2x_1 + 3x_2 & -x_3 = 1\\ -x_1 & +x_3 = 5\end{cases}$$

est équivalent à l'équation

$$\begin{bmatrix}2 & 3 & -1\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}1\\ 5\end{bmatrix}.$$

Exercice II.1.37. Ecrire les systèmes linéaires du chapitre I sous forme matricielle.

II.1.f. Formule du binôme

On rappelle que si a, b sont deux nombres complexes et $n \geq 1$, on a la formule du binôme :

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k},$$

où les $\binom{n}{k}$ sont les coefficients du binôme :

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Rappelons encore que ces coefficients du binôme vérifient la relation de récurrence :

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}. \tag{II.6}$$

Cette formule du binôme reste vraie pour des matrices carrées, *lorsque ces matrices commutent.*

Proposition II.1.38. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ telles que*

$$AB = BA.$$

Alors

$$(A+B)^n = \sum_{k=1}^{n} \binom{n}{k} A^k B^{n-k}. \tag{II.7}$$

La démonstration, qui est exactement la même que dans le cas scalaire, est laissée au lecteur. On peut raisonner par récurrence sur n, en utilisant la formule (II.6).

Exercice II.1.39. Trouver deux matrices 2×2 A et B, qui ne commutent pas, et telles que la formule (II.7) soit fausse avec $n = 2$.

II.2. Matrices inversibles : définitions et exemples

Le but de cette section et de la suivante est l'étude des matrices carrées inversibles pour la multiplication matricielle. Cette section est consacrée à des définitions et des exemples simples. En II.3, nous verrons une méthode de calculs des inverses et une caractérisation des matrices inversibles qui reposent largement sur la méthode du pivot de Gauss vue au chapitre I.

II.2.a. Définition

Définition II.2.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice *carrée.* La matrice A est dite *inversible* quand il existe des matrices $B, C \in \mathcal{M}_n(\mathbb{K})$ telles que

$$AB = CA = I_n,$$

où la matrice I_n est la matrice identité $n \times n$ définie en (II.2). L'ensemble des matrices $n \times n$ inversibles est noté $GL_n(\mathbb{K})$. Les étudiants de la filière *mathématiques* verront en cours d'arithmétique que $GL_n(\mathbb{K})$, muni de la multiplication, est un *groupe.*

Exemple II.2.2. La matrice nulle de $\mathcal{M}_n(\mathbb{K})$ n'est pas inversible. La matrice I_n est inversible ($I_n = I_n \times I_n$). La matrice $\left[\begin{smallmatrix} 3 & 2 \\ 1 & 1 \end{smallmatrix}\right]$ est inversible :

$$\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} = I_2.$$

(Vérifier le calcul à l'aide de la formule de la multiplication matricielle.)

Proposition II.2.3. *Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, il existe un unique $B \in \mathcal{M}_n(\mathbb{K})$ tel que*

$$AB = BA = I_n. \tag{II.8}$$

La matrice B est appelée inverse *de A, et notée A^{-1}. L'unicité signifie que si $M \in \mathcal{M}_n(\mathbb{K})$ vérifie $MA = I_n$* ou *$AM = I_n$, alors $M = A^{-1}$.*

En d'autres termes, si une matrice est inversible, l'inverse à gauche et l'inverse à droite de cette matrice sont uniques et égaux.

Démonstration. Il suffit de montrer que si $AB = I_n$ et $CA = I_n$, alors $B = C$. La matrice B donnée par la définition II.2.1 vérifiera alors (II.8), et sera bien unique au sens donné par la proposition.

Cette propriété découle de l'associativité de la multiplication matricielle. En multipliant à gauche l'égalité $AB = I_n$ par C, on obtient

$$\underbrace{CA}_{I_n} B = CI_n = C$$

et donc $B = C$. □

Voici une propriété importante des matrices inversibles (cf l'avertissement II.1.35) :

Proposition II.2.4. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice inversible. Alors :*

i. Si $M \in \mathcal{M}_{p,n}(\mathbb{K})$,

$$MA = 0 \Longrightarrow M = 0.$$

ii. Si $M \in \mathcal{M}_{n,p}(\mathbb{K})$,

$$AM = 0 \Longrightarrow M = 0.$$

Remarque II.2.5. La proposition implique que si A est inversible, $(0, 0, \dots, 0)$ est l'unique solution du système homogène $AX = 0$, $X \in \mathbb{K}^n$ qui a n solutions et n inconnues. En d'autres termes, ce système est un système de Cramer. De fait, l'unique solution de l'équation matricielle $AX = B$ est $X = A^{-1}B$. Nous verrons plus loin qu'il y a en fait équivalence : un système de n équations à n inconnues est un système de Cramer si et seulement si la matrice de ses coefficients est inversible.

Démonstration. Démontrons (i), la démonstration de (ii) est similaire. On suppose donc $MA = 0$. En multipliant à droite par A^{-1}, on obtient

$$M \underbrace{A\,A^{-1}}_{I_n} = 0\,A^{-1} = 0$$

et donc $M = MI_n = 0$. □

On peut en déduire un exemple typique de matrice non-inversible :

Proposition II.2.6. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On suppose qu'une des colonnes, ou une des lignes de A est nulle. Alors A n'est pas inversible.*

Démonstration. On suppose que la i-ème ligne de $A = [a_{ij}]$ est nulle. Soit $Y = [y_j]_{1\leq j\leq n} = [\delta_{ij}]_{1\leq j\leq n}$ la matrice ligne de $\mathcal{M}_{1,n}(\mathbb{K})$ dont tous les coefficients sont nuls, sauf le i-ème, qui vaut 1. Alors la matrice ligne YA est nulle : en effet, le ℓ-ième coefficient de cette matrice est donné par

$$\sum_{k=1}^{n} y_k a_{k\ell} = 0,$$

car $y_k = 0$ si $k \neq i$ par définition de Y, et $a_{i\ell} = 0$ car la i-ème ligne de A est nulle. On en déduit par la Proposition II.2.4 que A n'est pas inversible.

Dans le cas où la j-ième colonne de A est nulle, on fait le même raisonnement en multipliant A par la matrice colonne $X \in \mathcal{M}_n(\mathbb{K})$ dont tous les coefficients sont nuls, sauf le j-ième qui vaut 1. □

Exemple II.2.7. La matrice $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & 2i & 4 \end{bmatrix}$ n'est pas inversible. Dans ce cas, la démonstration de la proposition II.2.6 signifie que $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} A = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, contredisant la proposition II.2.4.

On donne maintenant deux exemples où il est facile de voir si une matrice est inversible et, le cas échéant, de calculer son inverse.

II.2.b. Matrices diagonales

Définition II.2.8. On appelle *matrice diagonale* une matrice carrée $[a_{ij}]_{\substack{1\leq i\leq n\\1\leq j\leq n}}$ dont les coefficients en dehors de la diagonale principale $\{i = j\}$ sont nuls. En d'autres termes :

$$i \neq j \Longrightarrow a_{ij} = 0.$$

Exemples II.2.9. Les matrices I_n et $0_{n,n}$ sont diagonales.

Considérons les matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix} \text{ et } C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

La matrice A est diagonale. Les matrices B et C ne sont pas diagonales (C n'est même pas carrée).

Remarquons que la somme de deux matrices diagonales de même taille est diagonale, et que si A est diagonale, ${}^tA = A$. On note $\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ ou $\operatorname{diag}(\lambda_j)_{1\leq j\leq n}$ la matrice diagonale dont les coefficients diagonaux sont λ_1, λ_2,…,λ_n. Par exemple

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \operatorname{diag}(1,2,3) = \operatorname{diag}(j)_{1\leq j\leq 3}, \quad I_n = \operatorname{diag}(1)_{1\leq j\leq n}.$$

On peut calculer très facilement le produit de deux matrices diagonales :

Proposition II.2.10. *Soit $A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ et $B = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)$ deux matrices diagonales de même taille. Alors*

$$AB = \operatorname{diag}(\lambda_1\mu_1, \lambda_2\mu_2, \ldots, \lambda_n\mu_n).$$

La démonstration, facile, est laissée au lecteur (utiliser (II.1)).

Corollaire II.2.11. *La matrice diagonale $D = \operatorname{diag}(\lambda_j)_{1\leq j\leq n}$ est inversible si et seulement si $\lambda_j \neq 0$ pour tout $j \in \{1, \ldots, n\}$. Dans ce cas,*

$$D^{-1} = \operatorname{diag}(1/\lambda_j)_{1\leq j\leq n}.$$

Démonstration. Si tous les λ_j sont non nuls, il est facile de vérifier, en utilisant la proposition II.2.10 que

$$\operatorname{diag}(1/\lambda_j)_{1\leq j\leq n}\, D = D \operatorname{diag}(1/\lambda_j)_{1\leq j\leq n} = I_n.$$

Supposons qu'il existe $k \in \{1, \ldots, n\}$ tel que $\lambda_k = 0$. Alors la k-ième ligne de D est nulle, et donc par la Proposition II.2.6, D n'est pas inversible. □

Exemples II.2.12. La matrice

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

est inversible, d'inverse

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1/3 \end{bmatrix}.$$

II.2.c. Inversibilité des matrices 2×2

Proposition II.2.13. *Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice 2×2. Alors la matrice A est inversible si et seulement si $ad - bc \neq 0$. Dans ce cas, on a $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.*

Remarque II.2.14. La quantité $ad-bc$ est appelée *déterminant* de A. Le déterminant se généralise à des matrices carrées de plus grande dimension mais les formules sont plus compliquées (voir le chapitre VI de ce cours).

Démonstration. Soit $B = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Par la formule du produit matriciel :

$$AB = BA = (ad - bc)I_2.$$

Lorsque $ad - bc = 0$, on obtient $AB = 0$ et la proposition II.2.4 montre que A n'est pas inversible. Lorsque $ad - bc = 0$, on obtient

$$A\,\frac{1}{ad - bc}B = \frac{1}{ad - bc}B\,A = I_2,$$

ce qui montre que A est inversible, d'inverse $\frac{1}{ad-bc}B$. □

Exercice II.2.15. Inverser les matrices suivantes

$$\begin{bmatrix}-1 & -47\\ 27 & 1268\end{bmatrix} \qquad \begin{bmatrix}-1 & 25\\ 33 & -824\end{bmatrix} \qquad \begin{bmatrix}-18 & -19\\ 701 & 740\end{bmatrix}.$$

II.2.d. Stabilité par multiplication et transposition

Proposition II.2.16. *i. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors ${}^tA$ est inversible si et seulement si A est inversible. Dans ce cas, $({}^tA)^{-1} = {}^t(A^{-1})$.*

ii. Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ deux matrices inversibles. Alors AB est inversible et

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

Démonstration. On montre d'abord (i). Supposons pour commencer que A est inversible. On transpose les égalités :

$$AA^{-1} = A^{-1}A = I_n,$$

ce qui donne (${}^t(AB) = {}^tB{}^tA$) :

$${}^t(A^{-1}){}^tA = {}^tA{}^t(A^{-1}) = {}^tI_n = I_n.$$

Donc ${}^tA$ est inversible, d'inverse ${}^t(A^{-1})$.

Réciproquement, si ${}^tA$ est inversible, $A = {}^t({}^tA)$ est inversible par ce qui précède, ce qui conclut la preuve du point (i).

Pour montrer (ii), on utilise l'associativité de la multiplication :

$$AB\,B^{-1}A^{-1} = A\,I_n\,A^{-1} = A\,A^{-1} = I_n,$$

et de même

$$B^{-1}A^{-1}\,AB = B^{-1}\,I_n\,B = B^{-1}B = I_n.$$

□

Avertissement II.2.17. Il ne faut pas se tromper dans l'ordre des facteurs dans la formule du (ii). Rappelons que la multiplication matricielle *n'est pas* commutative. On n'a donc pas, en général $(AB)^{-1} = A^{-1}B^{-1}$.

II.3. Opérations sur les lignes et inversion de matrices

Il y a plusieurs critères équivalents d'inversibilité d'une matrice. Pour le démontrer, nous allons utiliser des notions déjà vues au chapitre I du cours, consacré aux systèmes linéaires. En II.3.a, on introduit des *matrices élémentaires* correspondant aux opérations élémentaires du chapitre I. En II.3.b on reparle de matrices échelonnées réduites, et en II.3.c de la méthode du pivot de Gauss. Les matrices inversibles sont caractérisées en II.3.d, par le théorème II.3.21. Le II.3.e est consacrée à l'interprétation matricielle des systèmes de Cramer.

Les deux points les plus importants de ce chapitre sont l'inversion de matrice par la méthode du pivot de Gauss et le théorème II.3.21.

II.3.a. Matrices élémentaires

On va montrer que les opérations élémentaires sur les lignes d'une matrice, rencontrées dans le chapitre sur les systèmes linéaires, reviennent à des multiplications à gauche par des matrices bien particulières, appelées matrices élémentaires. On commence par définir ces matrices.

On fixe $n \geq 2$.

Définition II.3.1. On appelle *matrice élémentaire* une matrice carrée $n \times n$ d'un des trois types suivants.

Soient $\lambda \in \mathbb{K} \setminus \{0\}$ et $k \in \{1, \dots, n\}$. La matrice de *dilatation* $D_k(\lambda)$ est la matrice diagonale de $\mathcal{M}_n(\mathbb{K})$ dont le k-ième coefficient diagonal vaut λ et les autres coefficients diagonaux valent 1.

Soient $k, \ell \in \{1, \dots, n\}$ avec $k \neq \ell$, et $\lambda \in \mathbb{K}$. La matrice de *transvection* $T_{k\ell}(\lambda)$ est la matrice dont les coefficients diagonaux sont égaux à 1, le coefficient (k, ℓ) vaut λ, et les autres coefficients sont nuls.

Si $k, \ell \in \{1, \dots, n\}$ avec $k \neq \ell$, on note $R_{k\ell}$ la matrice dont les coefficients diagonaux valent 1, sauf les coefficients (k, k) et (ℓ, ℓ), qui valent 0, les coefficients (k, ℓ) et (ℓ, k) valent 1, et les autres coefficients sont nuls. La matrice $R_{k\ell}$ est donc obtenue, à partir de I_n, en échangeant la ligne k et la ligne ℓ. Remarquons que $R_{\ell k} = R_{k\ell}$. Les matrices $R_{k\ell}$ sont parfois appelées *matrices de transposition* (à ne pas confondre avec la transposée ${}^tA$).

Remarquons que les matrices élémentaires dépendent de n. Nous n'avons pas indiqué cette dépendance pour ne pas alourdir les notations.

Exemples II.3.2. On suppose $n = 3$. Alors

$$D_1(\lambda) = \begin{bmatrix}\lambda & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad D_2(\lambda) = \begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & 1\end{bmatrix}, \quad D_3(\lambda) = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \lambda\end{bmatrix},$$

$$T_{12}(\lambda) = \begin{bmatrix}1 & \lambda & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad T_{31}(\lambda) = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ \lambda & 0 & 1\end{bmatrix}, \quad R_{23} = \begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix},$$

$$R_{13} = \begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}, \quad R_{12} = \begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Exercice II.3.3. Ecrire $D_2(\lambda)$, $T_{24}(\lambda)$, $T_{42}(\lambda)$, R_{13} quand $n = 4$.

Proposition II.3.4. *Soit $q \geq 1$ et $A \in \mathcal{M}_{n,q}(\mathbb{K})$.*

i. Soient $k \in \{1, \dots, n\}$ et $\lambda \neq 0$. La matrice $D_k(\lambda)A$ est obtenue à partir de A en multipliant la k-ième ligne de A par λ. La multiplication à gauche par $D_k(\lambda)$ correspond donc à l'opération élémentaire, appelée cadrage *et notée $(L_k) \leftarrow \lambda(L_k)$ dans le chapitre précédent du cours.*

ii. *Soient $k, \ell \in \{1, \dots, n\}$, $k \neq \ell$ et $\lambda \neq 0$. La matrice $T_{k\ell}(\lambda)A$ est obtenue à partir de A en ajoutant à la k-ième ligne de A le produit de λ et la ℓ-ième ligne de A. La multiplication à gauche par $T_{k\ell}(\lambda)$ correspond donc à l'opération élémentaire, appelée* remplacement *et notée $(L_k) \leftarrow (L_k) + \lambda(L_\ell)$ dans le chapitre précédent du cours.*

iii. *Soient $k, \ell \in \{1, \dots, n\}$. La matrice $R_{k\ell}A$ est obtenue à partir de A en échangeant les lignes k et ℓ. La multiplication à gauche par $R_{k\ell}$ correspond donc à l'opération élémentaire notée $(L_k) \leftrightarrow (L_\ell)$ dans le chapitre précédent.*

Remarque II.3.5. On peut montrer [2] que la multiplication à droite par les matrices élémentaires correspond à des opérations élémentaires sur les colonnes.

Exercice II.3.6. Calculer en utilisant la formule du produit matriciel :

$$T_{21}(\lambda)\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

et vérifier que la résultat est cohérent avec la proposition II.3.4.

Preuve de la proposition II.3.4. On ne démontre que le point (ii). La démonstration des autres points est laissée au lecteur.

On note $A = [a_{ij}]_{\substack{1\leq i\leq n \\ 1\leq j\leq q}}$, $T_{k\ell}(\lambda) = [b_{ij}]_{\substack{1\leq i\leq n \\ 1\leq j\leq n}}$ et $T_{k\ell}(\lambda)A = [c_{ij}]_{\substack{1\leq i\leq n \\ 1\leq j\leq q}}$. On a donc

$$b_{ii} = 1,\ i = 1 \dots n, \qquad b_{k\ell} = \lambda, \qquad b_{ij} = 0 \text{ si } i \neq j \text{ et } (i,j) \neq (k,\ell).$$

Soient $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, q\}$. La formule du produit matriciel (II.1) donne :

$$c_{ij} = \sum_{r=1}^{n} b_{ir} a_{rj}. \tag{II.9}$$

Si $i \neq k$, $b_{ir} = 0$ pour $r \neq i$, et $b_{ii} = 1$. La formule précédente donne donc $c_{ij} = a_{ij}$. La i-ème ligne de $T_{k\ell}(\lambda)A$ est donc exactement la i-ème ligne de A.

On considère maintenant le cas $i = k$. On a $b_{kr} = 0$ pour $r \notin \{k, \ell\}$, $b_{kk} = 1$, et $b_{k\ell} = \lambda$. La formule (II.9) avec $i = k$ s'écrit donc

$$c_{kj} = a_{kj} + \lambda a_{\ell j},$$

et la k-ième ligne de $T_{k\ell}(\lambda)A$:

$$[c_{k1}, \dots, c_{kq}] = [a_{k1}, \dots, a_{kq}] + \lambda[a_{\ell 1}, \dots, a_{\ell q}] = (L_k) + \lambda(L_\ell),$$

en notant (L_k) et (L_ℓ) la k-ième et la ℓ-ième ligne de A. Le point (ii) est démontré. □

Exercice II.3.7. En s'inspirant de la démonstration précédente, montrer les points (i) et (iii) de la proposition II.3.4.

2. Par exemple en appliquant la proposition II.3.4 aux matrices transposées.

Exercice II.3.8. Soit

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & -1 & 3 \\ 2 & 4 & 2 & 5 \\ 3 & 6 & 7 & 8 \\ 4 & 9 & 10 & 11 \end{bmatrix}.$$

En utilisant la proposition II.3.4, calculer AC, BC, AB, BA.

Proposition II.3.9. *Les matrices $D_k(\lambda)$ ($\lambda \neq 0$), $T_{k\ell}(\lambda)$ ($k \neq \ell$) et $R_{k\ell}$ sont inversibles et :*

i. $(D_k(\lambda))^{-1} = D_k\left(\frac{1}{\lambda}\right)$;

ii. $(T_{k\ell}(\lambda))^{-1} = T_{k\ell}(-\lambda)$;

iii. $R_{k\ell}^{-1} = R_{k\ell}$.

Démonstration. Les matrices de dilatation étant diagonales, le point (i) découle immédiatement du Corollaire II.2.11 (on peut aussi utiliser la proposition II.3.4 comme dans ce qui suit).

D'après la proposition II.3.4, la matrice

$$T_{k\ell}(\lambda)T_{k\ell}(-\lambda) = T_{k\ell}(\lambda)T_{k\ell}(-\lambda)\, I_n$$

est obtenue à partir de la matrice I_n par les opérations :

$$(L_k) \leftarrow (L_k) - \lambda L_\ell,$$

puis

$$(L_k) \leftarrow (L_k) + \lambda L_\ell.$$

Puisque $k \neq \ell$, on a donc bien [3] $T_{k\ell}(\lambda)T_{k\ell}(-\lambda) = I_n$ et de même $T_{k\ell}(-\lambda)T_{k\ell}(\lambda) = I_n$, ce qui montre le point (ii).

D'après la proposition II.3.4, la matrice $R_{k\ell}R_{k\ell} = R_{k\ell}R_{k\ell}I_n$ est obtenue à partir de I_n en échangeant deux fois les lignes ℓ et k. C'est donc bien la matrice I_n, ce qui montre le point (iii). □

Exercice II.3.10. Calculer lorsque $n = 4$, $R_{24}(17)R_{24}(-17)$ en utilisant la formule du produit matriciel et vérifier le point (II.3.4).

Exercice II.3.11. Démontrer la proposition II.3.9 en utilisant seulement la formule du produit matriciel, mais pas la proposition II.3.4.

3. l'hypothèse $k \neq \ell$ montre que la ligne (L_ℓ) n'a pas changé après la première opération.

II.3.b. Matrices échelonnées réduites carrées

On renvoie au chapitre précédent ou à David C. Lay [4] pour la définition d'une matrice échelonnée et d'une matrice échelonnée réduite. Le but de cette partie est de caractériser les matrices échelonnées réduites carrées inversibles.

Définition II.3.12. La matrice carrée $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ est dite *triangulaire supérieure* si tous les coefficients de A en dessous de la diagonale principale sont nuls. En d'autres termes :

$$1 \leq j < i \leq n \Longrightarrow a_{ij} = 0.$$

Exemples II.3.13. Les matrices diagonales sont triangulaires supérieures. En particulier, la matrice I_n et la matrice $0 \in \mathcal{M}_n(\mathbb{K})$ sont triangulaires supérieures. Considérons les matrices :

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}.$$

La matrice A est triangulaire supérieure. Les matrices B et C ne le sont pas (B n'est pas carrée. Le coefficient $(2,1)$ de C est non nul).

Proposition II.3.14. *Une matrice échelonnée carrée est triangulaire supérieure.*

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée. On note p le nombre de lignes non nulles de A. On a donc $0 \leq p \leq n$, $p = n$ si toutes les lignes de A sont non nulles, $p = 0$ si $A = 0$. De plus, si $1 \leq p \leq n$, A étant échelonnée, les lignes 1, 2,..., p de A sont non nulles, et les lignes $p+1, \ldots, n$ sont nulles.

On suppose $A \neq 0$, i.e. $p \geq 1$ (sinon $A = 0$ est triangulaire supérieure et la démonstration est finie).

Pour $1 \leq i \leq p$, on note $J(i)$ la colonne de l'élément de tête (le coefficient non nul le plus à gauche) de la i-ème ligne de A. Par propriété des matrices échelonnées, on a $J(k+1) \geq J(k)+1$ pour $k = 1 \ldots p-1$ et donc, puisque $J(1) \geq 1$, $J(i) \geq i$ pour tout i entre 1 et p. On en déduit

$$(1 \leq i \leq p \text{ et } j < i) \Longrightarrow (1 \leq i \leq p \text{ et } j < J(i)) \Longrightarrow a_{ij} = 0,$$

ce qui montre que la matrice A est triangulaire supérieure. □

Exemple II.3.15. Pour illustrer la proposition et sa démonstration, considérons la matrice échelonnée 5×5 :

$$\begin{bmatrix} 1 & 1+i & 3 & 2 & 0 \\ 0 & 0 & 2 & 3 & i \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

4. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

C'est bien une matrice triangulaire supérieure. Le nombre de lignes non nulles est $p = 4$; $J(1) = 1$, $J(2) = 3$, $J(3) = 4$ et $J(4) = 5$.

Théorème II.3.16. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée réduite. Les conditions suivantes sont équivalentes :*

i. A est inversible ;

ii. aucune ligne de A n'est nulle ;

iii. $A = I_n$.

La seule matrice échelonnée réduite inversible est donc la matrice identité.

Démonstration. La matrice I_n est inversible, donc (iii)$\Longrightarrow$(i). De plus, on sait déjà (i)$\Longrightarrow$(ii) (cf proposition II.2.6).

Il reste à démontrer (ii)$\Longrightarrow$(iii). Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée réduite sans ligne nulle. Par la proposition II.3.14, elle est triangulaire supérieure. Par hypothèse, elle a exactement n lignes non nulles. On reprend la notation $J(i)$ de la démonstration de la proposition II.3.14. Montrons pour commencer

$$\forall i \in \{1, \ldots, n\}, \quad J(i) = i. \tag{II.10}$$

Puisque la matrice est triangulaire supérieure, on a $i \leq J(i) \leq n$ pour tout $i \in \{1, \ldots, n\}$. La matrice A étant échelonnée, on a aussi $J(i) \leq J(i+1) - 1$ pour tout $i \in \{1, \ldots, n-1\}$. Puisque $J(n) \leq n$, on obtient $J(i) \leq i$ pour tout i, par une récurrence descendante sur i. D'où (II.10).

Par (II.10), les éléments de tête sont tous sur la diagonale $i = j$. La matrice A étant échelonnée réduite, ces éléments de tête sont tous égaux à 1, et les coefficients au-dessus de chaque élément de tête sont nuls, ce qui montre $A = I_n$. □

II.3.c. Inversions de matrices par la méthode du pivot de Gauss

Soit A une matrice carrée. Par la méthode du pivot de Gauss (cf chapitre précédent du cours), on peut ramener A à une matrice échelonnée réduite A' par un certain nombre (disons p) de transformations élémentaires sur les lignes de A. Ces transformations élémentaires correspondant à des multiplications par des matrices élémentaires (cf §II.3.a et Proposition II.3.4), on a donc :

$$E_p \ldots E_1 A = A',$$

où les matrices E_j correspondent aux p transformations élémentaires appliquées à A. Puisque les matrices élémentaires sont inversibles, on a :

$$A = E_1^{-1} \ldots E_p^{-1} A'.$$

D'où (les matrices E_j^{-1} étant elles aussi des matrices élémentaires) :

Théorème II.3.17. *Toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est produit de matrices élémentaires et d'une matrice échelonnée réduite A'. Elle est inversible si et seulement si $A' = I_n$, c'est à dire si et seulement si elle est produit de matrices élémentaires.*

(le dernier point découle du théorème II.3.16).

Application : calcul de l'inverse d'une matrice

Donnons maintenant une méthode pratique pour étudier l'inversibilité de A, et, lorsque A est inversible, calculer son inverse. On commence par écrire sur deux colonnes la matrice A et la matrice I_n. On ramène ensuite, par la méthode du pivot de Gauss, la matrice A à une matrice échelonnée réduite, tout en appliquant les mêmes opérations élémentaires sur la matrice I_n.

— Si A est inversible, on obtient sur la colonne de gauche la matrice $I_n = E_p \dots E_1 A$ et sur la colonne de droite la matrice $E_p \dots E_1$. L'égalité $I_n = E_p \dots E_1 A$ montre que la matrice obtenue sur la colonne de droite est exactement A^{-1}.

— Si A n'est pas inversible, on obtient sur la colonne de gauche une matrice échelonnée réduite avec au moins une ligne nulle. Dans ce cas, si le but est seulement d'étudier l'inversibilité de A, la colonne de droite est inutile, et on peut arrêter la méthode du pivot dès que l'on a obtenu une ligne nulle.

Cette méthode sera implémentée informatiquement dans la partie "algorithmique" de ce cours.

Exemple II.3.18. On veut inverser la matrice $A = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 3 & 0 \\ -6 & 8 & -1 \end{bmatrix}$.

$$\begin{bmatrix} 1 & -1 & 1 \\ -2 & 3 & 0 \\ -6 & 8 & -1 \end{bmatrix} \qquad \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{bmatrix} \qquad \begin{matrix} (L_2) \leftarrow (L_2) + 2(L_1) \\ (L_3) \leftarrow (L_3) + 6(L_1) \end{matrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 6 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \qquad (L_3) \leftarrow (L_3) - 2(L_2) \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{matrix} (L_1) \leftarrow (L_1) - (L_3) \\ (L_2) \leftarrow (L_2) - 2(L_3) \end{matrix} \qquad \begin{bmatrix} -1 & 2 & -1 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (L_1) \leftarrow (L_1) + (L_2) \qquad \begin{bmatrix} -3 & 7 & -3 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix}.$$

Donc A est inversible, d'inverse

$$A^{-1} = \begin{bmatrix} -3 & 7 & -3 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix}.$$

Exemple II.3.19. Considérons la matrice $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 4 \\ 1 & 3 & 3 \end{bmatrix}$. Par les opérations $(L_2) \leftarrow (L_2) - 2(L_1)$, $(L_3) \leftarrow (L_3) - (L_1)$ puis $(L_3) \leftarrow (L_3) - (L_2)$, on obtient la matrice $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. La troisième ligne de cette matrice étant nulle, on en déduit que A n'est pas inversible.

On termine cette partie par une remarque qui découle facilement de la méthode précédente :

Proposition II.3.20. *Soit A une matrice triangulaire supérieure. Alors A est inversible si et seulement si ses coefficients diagonaux sont tous non-nuls.*

En effet, si aucun des coefficients diagonaux de A n'est nul, A est échelonnée, et la phase de remontée de la méthode du pivot permet d'obtenir I_n à partir de A par des opérations élémentaires sur les lignes.

En revanche, si un des coefficients diagonaux est nul, on peut transformer, par des opérations élémentaires sur les lignes, la matrice A en une matrice A' ayant une ligne nulle, ce qui montrera que A n'est pas inversible (sinon A' serait un produit de matrice inversible, donc inversible). Pour obtenir la ligne nulle, on note (i, i) les coordonnées du dernier coefficient diagonal nul, de tel sorte que $a_{ii} = 0$, et $a_{kk} \neq 0$ pour $k > i$. Exactement de la même façon que dans la phase de remontée de la méthode du pivot, en utilisant que les éléments de tête des lignes (L_k), $k > i$ sont non nuls, on transforme A, par une série de remplacements, en une matrice dont la i-ième ligne est nulle.

II.3.d. Caractérisation des matrices inversibles

Le théorème fondamental suivant, qui découle de ce qui précède, donne plusieurs critères pour reconnaître une matrice inversible.

Théorème II.3.21. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les conditions suivantes sont équivalentes :*

i. *A est inversible ;*

ii. *$\exists B \in \mathcal{M}_n(\mathbb{K})$ t.q. $BA = I_n$;*

iii. *$\exists C \in \mathcal{M}_n(\mathbb{K})$ t.q. $AC = I_n$;*

iv. *l'équation $AX = 0$, d'inconnue $X \in \mathcal{M}_{n,1}(\mathbb{K})$, a pour seule solution $X = 0$;*

v. *pour tout $E \in \mathcal{M}_{n,1}(\mathbb{K})$, l'équation $AX = E$, a une seule solution $X \in \mathcal{M}_{n,1}(\mathbb{K})$;*

Démonstration. On commence par montrer que les points (i), (ii), (iv) sont équivalents. Par définition de l'inversibilité, (i)⇒(ii). Par ailleurs, si (ii) est vrai et $AX = 0$, alors $X = BAX = B0 = 0$ et donc (ii)⇒(iv).

Supposons (iv). On veut montrer que A est inversible. Par la méthode du pivot de Gauss, réinterprétée en terme d'opérations élémentaires sur les lignes (cf §II.3.c), $E_p \dots E_1 {}^tA = R$ où R est une matrice échelonnée réduite et les E_j des matrices élémentaires. On veut montrer que $R = I_n$, ce qui impliquera que ${}^tA$ est inversible (car produit des matrices inversibles $E_1^{-1} \dots E_p^{-1}$) et donc que A est inversible. On raisonne par l'absurde : si $R \neq I_n$, par le théorème II.3.16, la dernière ligne de R est une ligne de 0. Soit $Y' = \begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$. Alors

$$Y' E_p \dots E_1 \, {}^tA = Y'R = 0_{1,n}$$

et donc

$$Y \, {}^tA = 0_{1,n}$$

avec $Y = Y' E_p \dots E_1 \in \mathcal{M}_{1,n}(\mathbb{K})$. Les matrices $E_1, \dots, E_p$ étant inversible, Y est non nul (Proposition II.2.4). En passant à la transposée, on obtient $AX = 0$ avec $X = {}^tY$, qui est un vecteur colonne non nul. Ceci contredit (iv). Donc A est inversible, ce qui conclut la preuve de (iv)⇒(i).

On a évidemment (v)⇒(iv), (i)⇒(v), et puisque (iv)⇒(i), le point (v) est équivalent à tous les points précédents.

Enfin, (i) implique (iii) par définition. Réciproquement, supposons (iii). En transposant l'égalité $AC = I_n$, on obtient ${}^tC {}^tA = I_n$. Donc ${}^tA$ vérifie (ii). Par ce qui précède, ${}^tA$ est inversible, ce qui implique, par la proposition II.2.16 sur l'inversibilité des matrices transposées, que A est inversible. On a montré (iii)⇒(i), ce qui conclut la preuve du théorème. □

Exercice II.3.22. Soit

$$A = \begin{bmatrix} -7 & 15 & -6 \\ -4 & 9 & -4 \\ -3 & 6 & -2 \end{bmatrix}, \quad X = \begin{bmatrix} -6 \\ -4 \\ -3 \end{bmatrix}.$$

Calculer AX. La matrice A est-elle inversible ? (Correction p. 26).

Remarque II.3.23. On peut également montrer que l'inversibilité de A est équivalente au fait que l'équation $YA = 0$, d'inconnue $Y \in \mathcal{M}_{1,n}(\mathbb{K})$, a pour seule solution $Y = 0$, ou que pour tout $F \in \mathcal{M}_{1,n}(\mathbb{K})$, l'équation $YA = F$, a une seule solution $Y \in \mathcal{M}_{1,n}(\mathbb{K})$. La démonstration de ces équivalences est laissée au lecteur.

II.3.e. Système de Cramer et matrice inversible

Soit (S) un système linéaire à n équations et n inconnues, et A la matrice des coefficients. La matrice A est donc une matrice carrée $n \times n$. Le système (S) s'écrit

$$AX = B,$$

où B est la matrice colonne ($B \in \mathcal{M}_{n,1}(\mathbb{K})$) formé du second membre de l'équation, et $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ est la matrice inconnue.

Le point (i) $\iff$ (v) du théorème II.3.21 signifie exactement :

$$(S) \text{ est un système de Cramer } \iff A \text{ est inversible.}$$

On distingue deux cas :

— La matrice A est inversible et le système (S) est un système de Cramer : il a une unique solution $X = A^{-1}B$, quel que soit le second membre B. Le calcul de A^{-1} permet de résoudre rapidement le système quel que soit B.

— Si A n'est pas inversible, le système homogène $AX = 0$ a une infinité de solutions : en effet, le point (iv) est faux, il y a donc une solution non nulle X et tous les λX, $\lambda \in \mathbb{K}$ sont aussi solutions. Le système (S) a ou bien aucune solution, ou bien une infinité de solutions. Dans ce cas, A étant fixée, la compatibilité du système $AX = B$ dépend du second membre B.

Exemple II.3.24. Résoudre les systèmes :

$$\begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = 1 \\ -x_1 + 2x_2 + 3x_3 + x_4 = 2 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 0 \\ -3x_2 - 3x_3 - 4x_4 = 1 \end{cases}, \quad \begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = 2 \\ -x_1 + 2x_2 + 3x_3 + x_4 = -1 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 1 \\ -3x_2 - 3x_3 - 4x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = -2 \\ -x_1 + 2x_2 + 3x_3 + x_4 = 0 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 0 \\ -3x_2 - 3x_3 - 4x_4 = 1 \end{cases}$$

La matrice des coefficients de ces trois systèmes est

$$A = \begin{bmatrix} 1 & -3 & -5 & -3 \\ -1 & 2 & 3 & 1 \\ -1 & 3 & 4 & 2 \\ 0 & -3 & -3 & -4 \end{bmatrix}.$$

On montre par la méthode du pivot que A est inversible et

$$A^{-1} = \begin{bmatrix} -1 & -9 & 7 & 2 \\ 1 & -1 & 2 & 0 \\ -1 & -3 & 2 & 1 \\ 0 & 3 & -3 & -1 \end{bmatrix}.$$

Les trois systèmes sont des systèmes de Cramer. Les solutions de ces systèmes sont respectivement :

$$A^{-1}\begin{bmatrix}1\\2\\0\\1\end{bmatrix}=\begin{bmatrix}-17\\-1\\-6\\5\end{bmatrix}\qquad A^{-1}\begin{bmatrix}2\\-1\\1\\0\end{bmatrix}=\begin{bmatrix}14\\5\\3\\-6\end{bmatrix}\qquad A^{-1}\begin{bmatrix}-2\\0\\0\\1\end{bmatrix}=\begin{bmatrix}4\\-2\\3\\-1\end{bmatrix}$$

Exercice II.3.25. Appliquer la méthode du pivot à l'exemple précédent pour calculer A^{-1} et vérifier le résultat annoncé.

II.4. Réponse à quelques exercices

Exercice II.1.26. $AC, AD, BA, B^2, CA, CB, DC$ et D^2 ne sont pas définis.

$$A^2=\begin{bmatrix}-5 & 2+2i\\-3-3i & -7\end{bmatrix},\quad AB=\begin{bmatrix}6 & 3+i & -3\\-6+2i & -3-2i & 9\end{bmatrix},\quad BC=\begin{bmatrix}5 & 4 & 6\\2 & 4 & 6\end{bmatrix}$$

$$BD=\begin{bmatrix}3-3i & 6-6i\\-3 & -6\end{bmatrix},\quad C^2=\begin{bmatrix}-2 & 2 & 3\\0 & 0 & 0\\-1 & -2 & -3\end{bmatrix},\quad CD=\begin{bmatrix}-12 & -24\\0 & 0\\0 & 0\end{bmatrix}$$

$$DA=\begin{bmatrix}0 & 0\\15 & -6-6i\\10 & -4-4i\end{bmatrix},\quad DB\begin{bmatrix}0 & 0 & 0\\-18 & -9-3i & 9\\-12 & -6-2i & 6\end{bmatrix}$$

Exercice II.2.15. Les matrices inverses sont $\left[\begin{smallmatrix}1268 & 47\\-27 & -1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}824 & 25\\33 & 1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}-740 & -19\\701 & 18\end{smallmatrix}\right]$.
Exercice II.3.22. On trouve

$$AX=\begin{bmatrix}0 & 0 & 0\end{bmatrix}.$$

La matrice ne vérifie donc pas le point iv du théorème II.3.21, ce qui montre qu'elle n'est pas inversible.

III. Les polynômes

Dans ce chapitre, la lettre $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. Les éléments de $\mathbb{K}$ sont appelés “nombres” ou “scalaires”. Ce chapitre est indépendant des autres. Il sera utilisé dans la partie “algorithmique” du cours et dans le chapitre VI de ce cours, qui est consacré aux déterminants.

III.1. Définitions

III.1.a. Polynômes comme suites finies

Un polynôme P sur $\mathbb{K}$ (ou à coefficients dans $\mathbb{K}$) est la donnée d'une suite $(a_n)_{n\geq 0}$ d'éléments de $\mathbb{K}$ telle qu'il existe un entier $p \geq 0$ avec

$$\forall n \geq p, \quad a_n = 0.$$

Les nombres a_n sont appelés les *coefficients* de P. Le plus grand nombre d tel que $a_d \neq 0$ est appelé *degré* de P et noté $\deg P$ ou $\deg(P)$. Le coefficient a_d correspondant à ce degré est appelé *coefficient dominant* de P. Le polynôme $(a_n)_{n\geq 0}$ tel que $a_n = 0$ pour tout n est appelé le *polynôme nul* et noté 0. Par convention $\deg 0 = -\infty$.

Exemple III.1.1. La suite $(1, 2, 0, 0, 0, \ldots)$ est un polynôme de degré 1 ($a_0 = 1$, $a_1 = 2$, les $\ldots$ signifient que $a_n = 0$ si $n \geq 2$).

La suite $(2^n)_{n\geq 0}$ n'est pas un polynôme.

III.1.b. Addition

Soit $P = (a_n)_{n\geq 0}$ et $Q = (b_n)_{n\geq 0}$ deux polynômes. La *somme* $P + Q$ de P et Q est par définition la suite $(a_n + b_n)_{n\geq 0}$. C'est aussi un polynôme. Ceci définit une addition sur l'ensemble des polynômes, qui est commutative et associative :

$$P + Q = Q + P, \quad (P + Q) + R = P + (Q + R).$$

Du fait de l'associativité, on peut noter sans ambiguïté $P + Q + R$ la somme de trois polynômes.

Proposition III.1.2. *Soit P et Q deux polynômes à coefficients dans $\mathbb{K}$.*

$$\deg(P + Q) \leq \max(\deg P, \deg Q).$$

Si $\deg P \neq \deg Q$,

$$\deg(P + Q) = \max(\deg P, \deg Q).$$

Démonstration. Soit $n = \deg P$, $p = \deg Q$. On a donc :

$$P = (a_j)_{j\geq 0}, \quad Q = (b_j)_{j\geq 0}$$

avec $a_j = 0$ si $j > n$, $b_j = 0$ si $j > p$, $a_n \neq 0$, $b_p \neq 0$. Par définition de l'addition des polynômes,

$$P + Q = (a_j + b_j)_{j\geq 0}.$$

Puisque $a_j + b_j = 0$ dès que $j > \max(p, n)$, on a bien que le degré de $P + Q$ est au plus égal à $\max(p, n)$.

Supposons maintenant $n \neq p$. Pour fixer les idées, on suppose $n > p$. On a donc $b_n = 0$. Le n-ième coefficient de $P + Q$ est $a_n + b_n = a_n \neq 0$. Le polynôme $P + Q$ est donc exactement de degré $n = \max(\deg P, \deg Q)$. Le cas $n < p$ se traite de la même manière. $\square$

III.1.c. Indéterminée

On s'empresse d'adopter une notation plus commode que la notation $(a_n)_{n\geq 0}$ pour désigner les polynômes. On fixe une lettre, généralement X, appelée *indéterminée.* Soit $P = (a_n)_{n\geq 0}$ un polynôme de degré d. On note ce polynôme :

$$P = a_d X^d + a_{d-1} X^{d-1} + \ldots + a_1 X + a_0 = \sum_{j=0}^{d} a_j X^j. \tag{III.1}$$

Exemple III.1.3. Le polynôme $(1, 2, 0, 0, 0, \ldots)$ est noté $2X + 1$. Le polynôme

$$(0, 0, 0, 9, 4, 0, 3, 0, 0, 0, \ldots)$$

est noté $3X^6+0X^5+4X^4+9X^3+0X^2+0X+0$ ou plus simplement $3X^6+4X^4+9X^3$.

Dans (III.1), les puissances sont rangées dans l'ordre décroissant. On peut aussi les ranger dans l'ordre croissant. De fait, par la définition et la commutativité de l'addition des polynômes, on a :

$$a_d X^d + a_{d-1} X^{d-1} + \ldots + a_1 X + a_0 = a_0 + a_1 X + \ldots + a_{d-1} X^{d-1} + a_d X^d.$$

Notation III.1.4. L'ensemble des polynômes à coefficients dans $\mathbb{K}$ et d'indéterminée X est noté $\mathbb{K}[X]$. Un polynôme P est parfois noté $P(X)$ lorsqu'on veut insister sur le fait que la lettre X désigne l'indéterminée.

III.1.d. Multiplication

Soit $P=\sum_{j=0}^{d} a_j X^j$ et $Q=\sum_{j=0}^{d'} b_j X^j$ deux polynômes, $d=\deg P$, $d'=\deg Q$. Par définition, le produit PQ de P et Q est le polynôme

$$PQ=\sum_{j=0}^{d+d'}\left(\sum_{k+\ell=j} a_k b_\ell\right) X^j. \tag{III.2}$$

Il s'obtient en développant l'expression $\sum_{k=0}^{d} a_k X^k \sum_{\ell=0}^{d'} b_\ell X^\ell$ et en utilisant les règles de calcul usuelles sur l'addition et la multiplication des scalaires, et la règle : $X^k X^\ell = X^{k+\ell}$.

Remarque III.1.5. Dans l'expression (III.2), $\sum_{k+\ell=j}$ signifie que la somme porte sur tous les indices (entiers naturels) k et ℓ tels que $k+\ell=j$. Par exemple :

$$\sum_{k+\ell=2} a_k b_\ell = a_0 b_2 + a_1 b_1 + a_2 b_0.$$

Il est très important de maîtriser ce genre de notation.

La formule

$$\deg(PQ)=\deg P+\deg Q \tag{III.3}$$

découle immédiatement de la définition de la multiplication des polynômes. Le coefficient dominant de PQ est $a_d b_{d'}$. La multiplication des polynômes vérifient aussi les règles de calcul usuelles :

Proposition III.1.6. *La multiplication des polynômes est commutative, associative, et distributive par rapport à l'addition : si $P,Q,R\in\mathbb{K}[X]$,*

$$PQ=QP,\quad (PQ)R=P(QR)\ \text{et}\ P(Q+R)=PQ+PR.$$

Démonstration. On ne démontre que l'associativité, la preuve des autres propriétés est laissée au lecteur. Soit $P=\sum_{k=0}^{d_1} a_k X^k$, $Q=\sum_{\ell=0}^{d_2} b_\ell X^\ell$ et $R=\sum_{m=0}^{d_3} c_m X^m$ trois polynômes. On doit vérifier

$$(PQ)R=P(QR). \tag{III.4}$$

En utilisant la définition de la multiplication, on obtient $PQ=\sum_{j=0}^{d_1+d_2}\left(\sum_{k+\ell=j} a_k b_\ell\right) X^j$, puis

$$(PQ)R=\sum_{r=0}^{d_1+d_2+d_3}\sum_{j+m=r}\left(\sum_{k+\ell=j} a_k b_\ell\right) c_m X^r \underset{\bigstar}{=} \sum_{r=0}^{d_1+d_2+d_3}\sum_{k+\ell+m=r} a_k b_\ell c_m X^r.$$

De même, $QR=\sum_{s=0}^{d_2+d_3}\sum_{\ell+m=s} b_\ell c_m X^s$ et donc

$$P(QR)=\sum_{r=0}^{d_1+d_2+d_3}\sum_{k+s=r} a_k\left(\sum_{\ell+m=s} b_\ell c_m\right) X^r \underset{\bigstar}{=} \sum_{r=0}^{d_1+d_2+d_3}\sum_{k+\ell+m=r} a_k b_\ell c_m X^r.$$

D'où $(PQ)R=P(QR)$. □

Exercice III.1.7. Se convaincre que les deux égalités ★ ci-dessus sont bien impliquées par la distributivité de la multiplication sur l'addition. On pourra commencer par écrire explicitement ces égalités dans le cas $d_1=d_2=d_3=1$.

En pratique, pour calculer le produit de deux polynômes, on n'utilise pas directement la formule (III.2)**, mais simplement les règles de calcul usuelles.** Par exemple :

$$\begin{aligned}(X^3+2X^2-1)(X^2+7)&=X^3X^2+7X^3+2X^2X^2+14X^2-X^2-7\\&=X^5+7X^3+2X^4+13X^2-7.\end{aligned}$$

III.2. Premières propriétés

III.2.a. Division euclidienne

Théorème III.2.1. *Soient A et B deux polynômes de $\mathbb{K}[X]$, avec B non nul. Il existe un unique polynôme Q et un unique polynôme R dans $\mathbb{K}[X]$ tels que :*

$$A=QB+R \qquad \deg(R)<\deg(B).$$

Définition III.2.2. Lorsque $R=0$ dans le théorème ci-dessus, c'est à dire lorsqu'il existe Q dans $\mathbb{K}[X]$ tel que $A=QB$, on dit que B *divise* A, que A est *divisible* par B, ou que B est un *diviseur de* A.

Démonstration. Existence.

Lorsque $A=0$, on peut choisir $Q=0$ et $R=0$. Supposons A non nul, et notons $A=a_nX^n+a_{n-1}X^{n-1}+\ldots+a_1X+a_0$, $B=b_pX^p+b_{p-1}X^{p-1}+\ldots+b_1X+b_0$, avec $n=\deg A$, $p=\deg B$. Les polynôme Q et R se construisent par récurrence, en utilisant une "division euclidienne partielle", résumée dans le lemme suivant, qui consiste à diviser seulement les termes de plus haut degré.

Lemme III.2.3. *Soit $G\in\mathbb{K}[X]$ tel que $\deg G\geq\deg B$. On note d le degré de G et $G=\sum_{k=0}^{d} g_k X^k$. Soit $S=\frac{g_d}{b_p}X^{d-p}$. Alors*

$$\deg(G-B\,S)<\deg G. \tag{III.5}$$

Preuve du lemme. Le polynôme $B\,S$ est de degré $d-p+p=d$, et son coefficient dominant (le coefficient de X^d) est $\frac{g_d}{b_p}\times b_p=g_d$. On en déduit comme annoncé que $G-B\,S$ est au plus de degré $d-1$. □

Pour montrer l'existence de Q et R, on construit par récurrence deux suites finies $(A_j)_{j=0\ldots J}$ et $(Q_j)_{j=1\ldots J}$, avec $A_0=A$ et telles que :
pour tout $j=0,\ldots,J$,

$$A=A_j+(Q_1+Q_2+\ldots+Q_j)B \tag{III.6}$$

$$d^o A_j<d^o A_{j-1} \tag{III.7}$$

et

$$d^o A_J < d^o B. \tag{III.8}$$

Soit $j \geq 1$. Supposons $A_0, A_1, \ldots, A_j, B_1, \ldots, B_j$ connus et vérifient (III.6), (III.7) On distingue deux cas :

- ou bien $\deg(A_j) < \deg(B)$, on pose $J = j$ et on arrête la construction ;
- ou bien $\deg(A_j) \geq \deg(B)$. On utilise alors le lemme qui nous donne Q_{j+1}, tel que $A_{j+1} = A_j - BQ_{j+1}$ vérifie $d^o A_{j+1} < d^o A_j$. Par définition de A_{j+1} et l'hypothèse de récurrence (III.6), on a bien
$$A = A_{j+1} + (Q_1 + Q_2 + \ldots + Q_j + Q_{j+1})B.$$

Par (III.5), la suite $\deg(A_j)$ est strictement décroissante. Puisque c'est une suite d'éléments de $\mathbb{N} \cup \{-\infty\}$, la construction précédente doit s'arrêter pour un certain J. Pour lequel (III.8) est vérifié. D'après (III.6) (au rang J) et (III.8), on a bien obtenu une division euclidienne de A par B, de quotient $Q = Q_1 + \ldots + Q_J$ et de reste A_J.

Unicité.

Supposons que $A = Q\,B + R = S\,B + T$, où les degrés de R et T sont strictement inférieurs à celui de B. On en tire :

$$R - T = (S - Q)B \quad \text{et} \quad \deg(R - T) < \deg(B). \tag{III.9}$$

On montre $Q = S$ par l'absurde. Si $Q \neq S$, alors $\deg(S - Q) \geq 0$. Donc

$$\deg((S - Q)\,B) = \deg(S - Q) + \deg(B) \geq \deg(B),$$

ce qui contredit (III.9).

Donc $Q = S$, d'où $R = T$ par (III.9). □

En pratique, pour calculer le quotient et le reste de la division euclidienne d'un polynôme par un autre, on calcule les suites A_j, Q_j de la démonstration précédente en posant la division euclidienne.

Exemple III.2.4. Division de $X^3 + 2X^2 + X + 1$ par $X^2 + 1$

$$\begin{array}{rrrrrrr|l}
 & X^3 & + & 2X^2 & + & X & + & 1 & X^2+1 \\
\hline
- & X^3 & & & - & X & & & X+2 \\
\hline
 & & & 2X^2 & & & + & 1 & \\
 & & - & 2X^2 & & & - & 2 & \\
\hline
 & & & & & & - & 1 &
\end{array}$$

Le résultat s'écrit : $X^3 + 2X^2 + X + 1 = (X^2 + 1)(X + 2) - 1$.
Quotient : $X + 2$, reste : -1.
Dans cet exemple, on a $Q_1 = X$, $A_1 = 2X^2 + 1$, $Q_2 = 2$ et $A_2 = -1$.

Exercice III.2.5. Effectuer la division euclidienne de A par B dans les cas suivants :
$A = X^4 - 2X^2 - X + 1$, $B = X^2 + X$;
$A = X^6 + 4X^4 - X^2 + 1$, $B = X^2 + 1$.
(voir correction p. 33).

III.2.b. Fonctions polynomiales

Définition III.2.6. On appelle *fonction polynomiale* sur $\mathbb{K}$ toute fonction de la forme

$$\widetilde{P} : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0,$$

où les coefficients $(a_j)_{j=0\ldots n}$ sont dans $\mathbb{K}$. Le polynôme $P = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_2X^2 + a_1X + a_0$ est appelé *polynôme associé* à la fonction $\widetilde{P}$.

Remarque III.2.7. Il faut toujours avoir en mémoire la différence entre la fonction polynomiale $\widetilde{P}$ (et sa variable x) et le polynôme P (et son indéterminée X) ; même si, lorsqu'il n'y a pas de confusion possible, on omet d'écrire le symbole $\widetilde{\ }$.

Ainsi, pour tout $a \in \mathbb{K}$, on notera désormais $P(a)$ la valeur prise par la fonction $\widetilde{P}$ au point a.

Proposition III.2.8. *Le reste de la division euclidienne d'un polynôme P par $X - \alpha$ est le polynôme constant égal à $P(\alpha)$.*

Démonstration. La division euclidienne de P par $X - \alpha$ s'écrit :

$$P = (X - \alpha)Q + R \text{ avec } \deg(R) < 1.$$

Puisque $\deg(R) < 1$, le polynôme R est une constante c et $P(\alpha) = c$. □

III.2.c. Polynôme dérivé

Définition III.2.9. Soit $P = a_nX^n + \cdots + a_1X + a_0$ un polynôme sur $\mathbb{K}$. On appelle *polynôme dérivé* de P le polynôme :

$$P' = na_nX^{n-1} + (n-1)a_{n-1}X^{n-2} + \cdots + 2a_2X + a_1 = \sum_{i=0}^{n-1} (i+1)a_{i+1}X^i.$$

On note P'', P''', $P^{(4)}$, …, $P^{(k)}$ la suite des polynômes dérivés successifs. On pose enfin $P^{(0)} = P$.

Exemple III.2.10. Soit $P = 5 + iX^2 + 4X^3$. Alors $P' = 2iX + 12X^2$, $P'' = 2i + 24X$ etc...

Remarquons que la fonction polynôme associée à la dérivée [1] P' d'un polynôme réel P est exactement la dérivée [2] de la fonction polynôme associée à P.

Proposition III.2.11. *Soient $P, Q \in \mathbb{K}[X]$. Alors*

$$\deg P \geq 1 \Longrightarrow \deg(P') = \deg(P) - 1 \tag{III.10}$$

$$P' = 0 \iff P \text{ est constant} \tag{III.11}$$

$$(\lambda P + \mu Q)' = \lambda P' + \mu Q', \quad P, Q \in \mathbb{K}[X],\ \lambda, \mu \in \mathbb{K} \tag{III.12}$$

$$(PQ)' = P'Q + PQ'. \tag{III.13}$$

1. au sens de la dérivation des polynômes
2. au sens de la dérivation des fonctions réelles de la variable réelle

Démonstration. La propriété (III.10) se déduit immédiatement des définitions du degré et du polynôme dérivé.

Preuve de (III.11). Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de degré n. Alors

$$P' = 0 \iff \sum_{k=1}^{n} k a_k X^{k-1} = 0 \iff \forall k \in \{1, \dots, n\},\ a_k = 0.$$

En particulier, si $n \geq 1$, $a_n = 0$, une contradiction. Donc $n = 0$ ou $n = -\infty$, ce qui signifie exactement que P est constant.

La preuve de (III.12) est directe et laissée au lecteur.

Preuve de (III.13). On commence par prouver (III.13) lorsque $P = X^n$, $n \geq 1$. On note

$$Q = \sum_{k=0}^{p} b_k X^k, \quad p = \deg(Q).$$

Alors

$$P' = nX^{n-1}, \quad Q' = \sum_{k=1}^{p} k b_k X^{k-1}, \quad PQ = \sum_{k=0}^{p} b_k X^{k+n}.$$

et donc

$$P'Q + PQ' = \sum_{k=0}^{p} n b_k X^{k+n-1} + \sum_{k=0}^{p} k b_k X^{k+n-1} = \sum_{k=0}^{p} (k+n) b_k X^{k+n-1} = (PQ)'.$$

On démontre maintenant le cas général. On commence par remarquer que si $N \geq 1$, P_1, …, P_N sont des polynômes et λ_1,…,λ_N des scalaires, alors (III.12) et une récurrence élémentaire nous donnent :

$$\left(\sum_{k=1}^{N} \lambda_k P_k\right)' = \sum_{k=1}^{N} \lambda_k P_k'. \tag{III.14}$$

Supposons P non nul (sinon le résultat est évident). On note $P = \sum_{k=0}^{n} a_k X^k$ avec $n = \deg(P)$. Par (III.14),

$$\begin{aligned}(PQ)' &= \sum_{k=0}^{n} a_k (X^k Q)' \\ &= \sum_{k=0}^{n} a_k \left[(X^k)'Q + X^k Q'\right] = \sum_{k=1}^{n} a_k k X^{k-1} Q + \sum_{k=0}^{n} a_k X^k Q' = P'Q + PQ',\end{aligned}$$

ce qui termine la preuve. □

III.3. Racines

III.3.a. Cas général

Définition III.3.1. Soit $P \in \mathbb{K}[X]$. On dit que $\alpha \in \mathbb{K}$ est une *racine* (ou un *zéro*) de P lorsque $P(\alpha) = 0$.

Théorème III.3.2. *Soit $P \in \mathbb{K}[X]$. Un élément α de $\mathbb{K}$ est racine de P si, et seulement si, P est divisible par $X - \alpha$.*

Démonstration. D'après la proposition III.2.8 $P = Q\,(X - \alpha) + P(\alpha)$, pour un certain $Q \in \mathbb{K}[X]$. Par conséquent P est divisible par $X - \alpha$ si, et seulement si, $P(\alpha) = 0$. □

Exercice III.3.3. Soit P le polynôme sur $\mathbb{R}$ défini par $P = X^3 - X^2 - 3X + 3$.

a. Déterminer une racine évidente de P.

b. En déduire une expression de P sous la forme d'un produit d'un polynôme de degré 1 par un polynôme de degré 2.

c. En déduire l'ensemble des racines de P.

(Correction p. 33.)

Définition III.3.4. Soit $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $r \in \mathbb{N}^*$. On dit que α est une racine *d'ordre r*, ou *de multiplicité r*, de P si $P = (X - \alpha)^r Q$ avec $Q(\alpha) \neq 0$.

D'après le théorème III.3.2, une racine est toujours de multiplicité au moins 1.
Lorsque $r = 1$, on dit que la racine est simple.
Lorsque $r = 2$, on dit que la racine est double.

Exemple III.3.5. Le polynôme $P = 3(X-1)^2(X-i)(X+i)^3$ a pour racines $1, i$ et $-i$. 1 est une racine double, i est une racine simple, et $-i$ est une racine d'ordre 3.

Un trinôme complexe du second degré de discriminant non nul a deux racines simples. Un trinôme du second degré de discriminant nul a une racine double.

Définition III.3.6. On dit que le polynôme P a exactement n racines *comptées avec leur ordre de multiplicité* (ou *avec multiplicité*) lorsque la somme des multiplicités de ses racines est exactement n.

Exemple III.3.7. Le polynôme P de l'exemple III.3.5 a 3 racines distinctes, mais 6 racines comptées avec leur ordre de multiplicité.

Avertissement III.3.8. Il y a donc deux manières de compter le nombre de racines d'un polynôme. L'expression "le polynôme P a n racines" est ambigüe et ne doit jamais être utilisée sans précision supplémentaire.

Remarque III.3.9. Si P a r racines et Q a s racines (comptées avec leur ordre de multiplicité), alors PQ a $r + s$ racines (comptées avec leur ordre de multiplicité). Cette propriété n'est plus valable lorsque l'on compte les racines distinctes des polynômes.

Exemple III.3.10. Si l'on compte les racines avec multiplicité, le polynôme $P = (X-1)^2$ a deux racines, le polynômes $Q = (X-1)$ a 1 racines, et le polynôme PQ a bien $2+1$ racines.

Mais P, Q et PQ ont chacun une seule racine distincte.

Théorème III.3.11. *Soit $P \in \mathbb{K}[X]$. La racine $\alpha \in \mathbb{K}$ de P est de multiplicité r si et seulement si, pour tout k entre 0 et $r-1$, $P^{(k)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$.*

Démonstration. Si α est racine de multiplicité r, alors $P = (X-\alpha)^r Q$ avec $Q(\alpha) \neq 0$.

En dérivant, on obtient :

$$P' = r(X-\alpha)^{r-1}Q + (X-\alpha)^r Q' = (X-\alpha)^{r-1}\left(rQ + (X-\alpha)Q'\right),$$

de la forme $(X-\alpha)^{r-1}Q_1$ avec $Q_1(\alpha) = rQ(\alpha) \neq 0$. Donc, lorsque $r > 1$, $P'(\alpha) = 0$.

En itérant ce raisonnement, on obtient :

$$\text{pour tout } k \leq r,\ P^{(k)} \text{ est de la forme } (X-\alpha)^{r-k}Q_k \text{ avec } Q_k(\alpha) \neq 0.$$

Donc lorsque $k < r$, $P^{(k)}(\alpha) = 0$. De plus $P^{(r)} = Q_r$ avec $Q_r(\alpha) \neq 0$ et donc $P^{(r)}(\alpha) \neq 0$.

Réciproquement, supposons $P(\alpha) = P'(\alpha) = \cdots = P^{(r-1)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$. Soit s la multiplicité de α. D'après ce qui précède

$$P(\alpha) = \ldots = P^{(s-1)}(\alpha) = 0 \text{ et } P^{(s)}(\alpha) \neq 0 \tag{III.15}$$

On montre $s = r$ par l'absurde :

— si $s > r$ alors, on aurait $P^{(r)}(\alpha) = 0$ par (III.15), ce qui est contraire aux hypothèses.

— si $s < r$ alors, on aurait $P^{(s)}(\alpha) = 0$, contredisant (III.15).

Donc $s = r$ et α est de multiplicité r. ☐

III.3.b. Polynômes à coefficients complexes

Théorème III.3.12 (Théorème de D'Alembert)**.** *(admis) Dans $\mathbb{C}[X]$, tout polynôme non constant admet au moins une racine.*

Corollaire III.3.13. *Tout polynôme P, de degré $n \geq 1$, de $\mathbb{C}[X]$ admet exactement n racines complexes (comptées avec leur ordre de multiplicité).*

Démonstration. Par récurrence sur n. Dans toute la démonstration, les racines sont comptées avec leur ordre de multiplicité.

— Initialisation : si $n = 1$, le résultat est immédiat.

— Hérédité : supposons que tout polynôme de degré $n-1$ de $\mathbb{C}[X]$ admette exactement $n-1$ racines complexes.
Si P est un polynôme de degré n, d'après le théorème de D'Alembert, il admet au moins une racine α.
Il existe donc Q, de degré n-1, tel que $P = (X-\alpha)Q$. D'après l'hypothèse de récurrence, Q admet n-1 racines α_1 , ..., α_{n-1}. Par conséquent, P admet les n racines α , α_1 , ..., α_{n-1}.

☐

Proposition III.3.14. *Soient f et g deux fonctions polynomiales sur $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, définies par $f(x) = a_n x^n + \cdots + a_1 x + a_0$ et $g(x) = b_p x^p + \cdots + b_1 x + b_0$, avec $a_n \neq 0$ et $b_p \neq 0$. Si*

$$\forall x \in \mathbb{K}, \quad f(x) = g(x),$$

alors $p = n$ et $a_i = b_i$ pour tout i.

Démonstration. S'il existait i tel que $a_i \neq b_i$, la fonction polynomiale $f - g$ serait de degré $k \geq i$. Elle aurait au plus k racines complexes, donc au plus k racines dans $\mathbb{K}$, et ne serait pas identiquement nulle, ce qui est contraire à l'hypothèse. ☐

III.3.c. Polynômes à coefficients réels

Puisque $\mathbb{R} \subset \mathbb{C}$, un polynôme à coefficients réels peut être considéré comme un polynôme à coefficients complexes, et ses racines réelles sont aussi des racines complexes. En particulier, un polynôme réel a au plus n racines réels. De plus, pour un tel polynôme P, si $\alpha \in \mathbb{C}$, on a $\overline{P(\alpha)} = P(\overline{\alpha})$ ($\overline{z}$ désigne le nombre complexe conjugué de z).

La proposition suivante donne les propriétés des racines complexes d'un polynôme réel :

Proposition III.3.15. *Soit $P \in \mathbb{R}[X]$. Si $\alpha \in \mathbb{C}$ est racine de P, alors $\overline{\alpha}$ l'est aussi. De plus, α et $\overline{\alpha}$ ont même ordre de multiplicité.*

Démonstration. Soit r l'ordre de multiplicité de α. On a donc $P^{(k)}(\alpha) = 0$ pour tout k entre 0 et $r-1$, et $P^{(r)}(\alpha) \neq 0$. Donc, $P^{(k)}(\overline{\alpha}) = \overline{P^{(k)}(\alpha)} = \overline{0} = 0$ pour tout $k \leq r-1$ et $P^{(r)}(\overline{\alpha}) = \overline{P^{(r)}(\alpha)} \neq 0$. ☐

Corollaire III.3.16. *Tout polynôme P, de degré impair de $\mathbb{R}[X]$ admet au moins une racine réelle.*

Démonstration. En effet, un nombre complexe α est réel si et seulement si $\alpha = \overline{\alpha}$. Il découle donc de la proposition III.3.15 que les racines complexes non réelles de P peuvent se ranger par paires de racines de même multiplicité. Il y a donc un nombre pair de racines complexes non réelles (comptées avec leur ordre de multiplicité). Puisque, par le corollaire III.3.13, P a exactement $\deg(P)$ racines, le nombre de racines réelles de P est impair, et donc non nul. ☐

Exercice III.3.17. Donner une autre démonstration du corollaire III.3.16, en étudiant la fonction polynôme associée à P et en utilisant le théorème des valeurs intermédiaires.

Exercice III.3.18.

a. Montrer que i est racine double du polynôme $P = X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1$.

b. Déterminer les réels a et b tels que le polynôme $P = X^5 + aX^4 + bX^3 - bX^2 - aX - 1$ admette 1 comme racine de plus grande multiplicité possible.

III.4. Complément : polynômes irréductibles

III.4.a. Cas général

Définition III.4.1. Un polynôme P non constant qui vérifie la condition :

si P est produit de deux polynômes de $\mathbb{K}[X]$, l'un des deux est constant

est dit *irréductible* dans $\mathbb{K}[X]$.

Par convention, les polynômes constants ne sont pas irréductibles.

Exemple III.4.2. Un polynôme de degré 1 est irréductible.

Théorème III.4.3. *Dans $\mathbb{K}[X]$, tout polynôme P non constant se décompose en produit de polynômes irréductibles*

Démonstration. Par récurrence sur le degré n de P :

Initialisation. Si $n = 1$ alors le polynôme est irréductible.

Hérédité. Supposons que tout polynôme de degré $< n$ soit produit de polynômes irréductibles.
Soit P un polynôme de degré n.

- Si P est irréductible, le résultat est obtenu.
- Si P n'est pas irréductible, $P = P_1P_2$ avec $\deg(P_1) \geq 1$ et $\deg(P_2) \geq 1$.
  $\deg(P_1) + \deg(P_2) = \deg(P) = n$ donc $\deg(P_1) < n$ et $\deg(P_2) < n$. Par hypothèse de récurrence P_1 et P_2 sont tous les deux produits de polynômes irréductibles donc P est aussi produit de polynômes irréductibles.

□

III.4.b. Polynômes irréductibles de $\mathbb{C}[X]$

Théorème III.4.4. *Les polynômes de degré 1 sont les seuls polynômes irréductibles de $\mathbb{C}[X]$.*

Démonstration. D'après le théorème de d'Alembert, tout polynôme P, de degré ≥ 2, admet au moins une racine $\alpha \in \mathbb{C}$. P est donc divisible par $(X - \alpha)$, et il n'est pas irréductible. □

Corollaire III.4.5. *(Décomposition dans $\mathbb{C}[X]$). Soit P un polynôme de degré $n \geq 1$ de $\mathbb{C}[X]$. Sa décomposition en produit de facteurs irréductibles est de la forme :*

$$P = \lambda\,(X - \alpha_1)^{r_1} \cdots (X - \alpha_p)^{r_p} \quad \text{avec} \quad r_1 + ... + r_p = n,$$

où $\alpha_1, \ldots \alpha_p$ sont les racines distinctes *de P, $r_1, \ldots, r_p$ leurs multiplicités et λ le coefficient dominant de P.*

III.4.c. Polynômes irréductibles de $\mathbb{R}[X]$

Théorème III.4.6. *Les seuls polynômes irréductibles de $\mathbb{R}[X]$ sont*

- *les polynômes de degré 1*
- *les polynômes de degré 2, dont le discriminant est strictement négatif*

Démonstration. — Si P est de degré 1, il est irréductible.

- Si P est un polynôme de degré 2 avec $\Delta < 0$, il est irréductible. Sinon, il se décomposerait en produit de deux polynômes, chacun de degré 1 :$P = (aX + b)(cX + d)$. Il aurait deux racines (distinctes ou confondues), et son discriminant serait ≥ 0. Contradiction.
- Si P est un polynôme de degré 2 avec $\Delta \geq 0$, il admet deux racines réelles (distinctes ou confondues) et s'écrit $P = a\,(X - \alpha_1)\,(X - \alpha_2)$. Il n'est pas irréductible.
- Si P est un polynôme de degré $n > 2$, d'après le théorème de d'Alembert, il admet au moins une racine $\alpha \in \mathbb{C}$
  - Ou bien $\alpha \in \mathbb{R}$, P est divisible par $(X - \alpha)$, et il n'est pas irréductible.
  - Ou bien $\alpha \notin \mathbb{R}$ alors $\overline{\alpha}$ est aussi racine de P.
    $(X - \alpha)(X - \overline{\alpha}) = \left(X^2 - 2\,\mathrm{Re}\,(\alpha)X + |\alpha|^2\right)$ est un polynôme à coefficients réels.
    On fait la division euclidienne dans $\mathbb{R}[X]$: $P = (X^2 - 2\,\mathrm{Re}\,(\alpha)X + |\alpha|^2)Q + R$ avec $\deg R \leq 1$. Or $R(\alpha) = R(\overline{\alpha}) = 0$ et $\alpha \neq \overline{\alpha}$. Donc $R = 0$.
    Par conséquent, P n'est pas irréductible dans $\mathbb{R}[X]$ car divisible par un polynôme de degré 2.

□

Corollaire III.4.7. *Décomposition dans $\mathbb{R}[X]$*

Soit P un polynôme de degré $n \geq 1$ de $\mathbb{R}[X]$. Sa décomposition en produit de facteurs irréductibles est de la forme :

$$P = \lambda\,(X - \alpha_1)^{r_1} \cdots (X - \alpha_p)^{r_p} (X^2 + \beta_1 X + \gamma_1)^{s_1} \cdots (X^2 + \beta_k X + \gamma_k)^{s_k}$$

$$\text{avec } r_1 + \cdots + r_p + 2(s_1 + \cdots + s_k) = n \text{ et } \beta_i^2 - 4\gamma_i < 0 \text{ pour } i = 1, \cdots k,$$

et les entiers naturels k et n sont non tous deux nuls.

Exemple III.4.8. $X^3 + X = X(X^2+1) = X(X+i)(X-i)$
La décomposition dans $\mathbb{C}[X]$ est $X(X+i)(X-i)$
La décomposition dans $\mathbb{R}[X]$ est $X(X^2+1)$.

Exercice III.4.9. Décomposer en produit de facteurs irréductibles dans $\mathbb{R}[X]$ le polynôme $P = X^4 - 2\,X^3 + 14\,X^2 - 18\,X + 45$ sachant qu'il admet $1+2\,i$ comme racine.

Correction. P est divisible par $(X-1-2\,i)(X-1+2\,i) = X^2 - 2X + 5$. En effectuant la division euclidienne on obtient $P = (X^2-2X+5)(X^2+9)$ qui est bien le produit de deux polynômes irréductibles de $\mathbb{R}[X]$. La décomposition dans $\mathbb{C}[X]$ s'en déduit immédiatement :

$$P = (X-1-2\,i)(X-1+2\,i)(X-3i)(X+3i).$$

Exercice III.4.10. Décomposer en produit de facteurs irréductibles le polynôme X^3+1 dans $\mathbb{R}[X]$ puis dans $\mathbb{C}[X]$.

III.5. Réponse à quelques exercices

Exercice III.2.5.

$$
\begin{array}{rrrrrrrrrr|l}
 & X^4 & & & - & 2X^2 & - & X & + & 1 & X^2+X \\ \hline
- & X^4 & - & X^3 & & & & & & & X^2-X-1 \\ \cline{1-10}
 & & - & X^3 & - & 2X^2 & - & X & + & 1 & \\
 & & & X^3 & + & X^2 & & & & & \\ \cline{1-10}
 & & & & - & X^2 & - & X & + & 1 & \\
 & & & & & X^2 & + & X & & & \\ \cline{1-10}
 & & & & & & & & & 1 &
\end{array}
$$

Le quotient de la division euclidienne de $X^4 - 2X^2 - X + 1$ par $X^2 + X$ est donc $X^2 - X - 1$, le reste est 1.

De même, on trouve que le quotient de la division euclidienne de $X^6+4X^4-X^2+1$ par X^2+1 est X^4+3X^2-4, et le reste 5.

Exercice III.3.3

a. 1 est une racine de P.

b. En effectuant la division euclidienne de P par $X-1$, on obtient $P = (X-1)(X^2-3)$.

c. L'équation $x^2 = 3$ a deux solutions, $\sqrt{3}$ et $-\sqrt{3}$. On a donc $P(x) = 0 \iff x \in \{-\sqrt{3}, 0, \sqrt{3}\}$. Les trois racines de P sont 0, $-\sqrt{3}$ et $\sqrt{3}$.

IV. Matrices et polynômes en C

Le but de ce court chapitre est de donner les bases pour utiliser matrices et polynômes en programmation. Le langage retenu est le C. Les opérations élémentaires vues au chapitres précédents (multiplication, évaluation *etc.*) seront programmées en TD/TP. La notion de complexité d'un algorithme est introduite en fin de chapitre et sera étudiée plus en détails dans un chapitre ultérieur.

IV.1. Rappels de base en C

IV.1.a. Code source et exécutable

Votre *code source* ou *programme* est un fichier texte, traditionnellement avec une extension `.c`, comme `code.c`. Vous créez et éditez ce fichier texte avec votre éditeur préféré (`gedit`, `emacs`...), qu'il faut ensuite le *compiler*, *i.e.*, le transformer en un fichier exécutable par la machine :

```
gcc code.c -o code
```

L'exécutable produit s'appelle ici `code` (le "o" de l'option `-o` signifie "output") et on lance l'exécution par la commande

```
./code
```

IV.1.b. Structure type d'un programme

```
#include <stdlib.h>
#include <stdio.h>

int blublu (int);

int main
{
 int i=3;
 int j=blublu(3);
 return EXIT_SUCCESS;
}

int blublu (int i)
{
 return i+1;
}
```

IV.1.c. Structures de contrôle

Elles contrôlent l'enchaînement des instructions de votre code (le "flot"). Il y a le *test* d'une expression (ici `i==0`) :

```
if (i==0)
 printf("i est nul");
else
 printf("i n'est pas nul");
```

La boucle `while` (tant que) :

```
while (i>0)
 {
 i=i-1;
 }
```

Enfin, très utile pour les matrices ou polynômes qui sont des structures de taille donnée, la boucle `for` (pour) itère une partie du code en incrémentant une variable (compteur) à chaque passage (`i++`) tant qu'une certaine condition est vérifiée (`i<p`). On peut les *imbriquer*, mais attention à utiliser des compteurs différents !

```
  for (i=0; i<p; i++)
  {
    for (j=0; j<q; j++)
    {
     ...
    }
  }
```

IV.1.d. Variables

Les variables (y compris les compteurs de boucle `for`) se déclarent dans la fonction où elles sont utilisées. Il faut préciser leur *type* : entier `int`, flottant `float`... Elles peuvent être affectées au moment de leur déclaration ou plus tard. On peut éventuellement convertir leur type (*cast*) dans la limite du raisonnable (typiquement : de flottant à entier).

```
int i;
float f=3.14;
i=(int)f;
```

Attention : l'ensemble des flottants n'est qu'une approximation du corps des réels (en particulier, il est fini et n'est pas toujours stable par addition ou multiplication) : ceci peut parfois être une source difficile à détecter d'erreurs de calcul.

IV.1.e. Pointeurs

La mémoire d'un ordinateur est comme une longue rue : la *valeur* d'une variable y est rangée dans une case mémoire comme un habitant dans sa maison, et pour accéder à cette valeur il faut son *adresse*, laquelle *pointe* sur cette valeur. L'adresse d'une variable `i` est `&i`, et le contenu de la case à l'adresse `P` est `*P`.

```
int i=5;   //la variable i est initialisée à la valeur 5
int* P;    //P est une adresse qui pointera sur des entiers (type int*)
P=&i;      //on fait pointer P sur i : P est l'adresse de i
int j=*P;  //j est initialisé par l'entier pointé par P et vaut donc 5
```

Dans l'exemple ci-dessus on aurait pu se passer de `P` et faire simplement

```
int j=i;
```

Mais l'intérêt de déclarer un pointeur (une adresse) sur des entiers et d'accéder à plusieurs entiers voisins dans la mémoire (ce qui sera bien utile pour les matrices) :

```
int k=*(P+1); // k est initialisé au contenu (inconnu) du voisin de i
```

Attention cependant à ne pas modifier la mémoire n'importe où, *i.e.*, en dehors des cases que vous avez explicitement demandées au système (soit en déclarant une variable, soit via un `malloc` - *memory allocation*), sinon c'est le classique et fâcheux `segmentation fault`.

IV.1.f. Divers

Des *commentaires* peuvent être mis dans le code du programme, soit entre `/* */`, soit sur une ligne après `//`. Ils sont *précieux* pour qui lira votre programme, vous au premier chef, n'en soyez pas avare.

Pour construire des exemples en TD/TP, il sera utile se générer des nombres aléatoires (ici un entier entre 0 et 10) :

```
int i =(rand() / (double)RAND_MAX * 10);
```

La commande unix `printf` est également très utile (ne serait-ce que pour un debuggage artisanal). On passe à la ligne suivante avec le caractère `\n`.

```
printf("Un entier %d, un nombre flottant %f\n",i,x);
printf("Le même flottant avec deux chiffres significatifs %.2f",x);
```

IV.2. Matrices et Polynômes

Une matrice est un tableau à deux dimensions stocké dans la mémoire de l'ordinateur. Un polynôme $a_n X^n + \ldots + a_0$ donné par ses coefficients peut être stocké en mémoire comme une matrice avec une ligne et $n+1$ colonnes, la i-ème colonne donnant le coefficient a_i du monôme de degré i.

IV.2.a. Allocation statique

Une première façon de manipuler matrices ou polynômes est de les déclarer en début de fonction, comme ici un tableau `t` de deux entiers et un tableau 2×3 (une matrice) `m` de flottants :

```
int t[2];
float m[2][3];
```

On écrit ou lit dans ces tableaux de façons très naturelle :

```
t[0]=1;
m[1][2]=t[1];
```

Remarquons que la numérotation commence à 0 (un tableau de taille n est donc numéroté de 0 à $n-1$).

On peut aussi initialiser les tableaux dès leur création :

```
int t[2]={1,2};
int m[2][3]={{1,2,3},{4,5,6}};
```

Cependant, la taille d'un tableau ainsi défini doit être définie explicitement dans le code, *i.e.*, avant compilation, et non par une variable (même si elle est connue à l'exécution). En outre, comme toute variable définie dans une fonction, un tel tableau est "oublié" à la fin de la fonction (la mémoire correspondante est allouée à l'appel de la fonction et libérée à son retour). La fonction suivante, qui semble pourtant naturelle, est donc doublement incorrecte :

```
int t[] suite(int n)
{
 int t[n];
 for (i=0; i<n; i++)
 {
  t[i]=i;
 }
 return t;
}
```

IV.2.b. Allocation dynamique

On remédie au problème précédent par une *allocation dynamique* : on demande au système de nous *allouer* (`malloc`) un bloc mémoire (dont la taille peut être déterminée à l'exécution) qui servira à stocker notre tableau. La fonction `suite` précédente devient alors possible :

```
int* suite(int n)
{
 int* P = malloc(n*sizeof(int)); // place pour n entiers à l'adresse P
 int i;     // déclaration nécessaire pour la boucle for
 for (i=0; i<n; i++)
 {
  *(P+i)=i; // valeur i dans le i-ème entier après l'adresse P
  P[i]=i;   // écriture simplifiée équivalente
 }
 return P;
}
```

On n'oubliera pas de *libérer* (`free`) la mémoire allouée quand on ne s'en sert plus (le compilateur retient tout seul la taille du bloc à libérer - au moins une chose que C sait faire tout seul). La fonction précédente s'utilisera donc comme suit :

```
int* P=suite(100);
...
free(P);
```

IV.3. Complexité

Un *algorithme* est une méthode automatique (donc programmable) pour résoudre un problème donné (par exemple, multiplier deux matrices entre elles). Il peut y avoir différents algorithmes pour résoudre un même problème. La *complexité* d'un algorithme cherche à mesurer son temps d'exécution en fonction de la taille des données qu'il traite (par exemple, la taille des matrices à multiplier). Pour mesurer cela, on compte le nombre d'*opérations élémentaires* (additions, multiplications, comparaisons. . .).

Plutôt que le nombre exact d'opérations, on veut avoir une idée *générale* de ce qui se passe pour des *grandes données*. Formellement, étant donnée une fonction $f : \mathbb{N} \to \mathbb{N}$, on note

$$O(f(n)) := \{g : \mathbb{N} \to \mathbb{N} \mid \exists k,\ \forall n,\ g(n) \leq kf(n)\},$$

et on cherche une fonction f "simple" telle que la complexité soit dans $O(f(n))$. Par exemple, si le nombre exact d'opérations est $5n^2 + 2\log(n) - 5$, on se contentera de de dire qu'il y a $O(n^2)$ opérations : la complexité est quadratique.

Appendice : alphabet grec

Minuscule	Majuscule	Nom	Minuscule	Majuscule	Nom
α	A	alpha	ν	N	nu
β	B	bêta	ξ	Ξ	xi
γ	Γ	gamma	o	O	omicron
δ	Δ	delta	π	Π	pi
ε	E	epsilon	ρ	P	rhô
ζ	Z	zêta	σ	Σ	sigma
η	H	êta	τ	T	tau
θ	Θ	thêta	υ	Υ	upsilon
ι	I	iota	ϕ	Φ	phi
κ	K	kappa	χ	X	khi
λ	Λ	lambda	ψ	Ψ	psi
μ	M	mu	ω	Ω	oméga

Table des matières de la première partie