Aix-Marseille Université

LMAT04 - Algèbre Linéaire

MI –L1

2012 – 2013

Sommaire

Introduction et bibliographie

Algèbre...

Le mot « algèbre »vient du terme arabe « al-jabr » signifiant littéralement « restauration ». Ce terme fut pour la première fois employé à propos des mathématiques par le novateur Al-Khwarizmi[1], dans son livre (Livre abrégé sur le calcul par la restauration et la comparaison) où il procède à l'étude systématique des équations du second degré et se ramène à six cas de base, en utilisant ce qu'il nomme « restauration » (al-jabr). Cette restauration se traduit essentiellement par l'ajout d'une même quantité dans les deux membres de l'équation afin d'éliminer les termes apparaissant en soustraction. Cette idée de modifier une égalité pour la rendre plus simple à résoudre est fondamentale. Il faut se rendre compte qu'à l'époque d'Al-Khwarizmi, le seul fait de penser un problème en termes d'égalité avec une grandeur inconnue était déjà une avancée considérable.

. . . linéaire

Voyons maintenant ce qu'est ce qu'un phénomène linéaire. Le prix de détail des marchandises, par exemple : quand le marchand affiche 2 euros le prix d'un kilo de pommes, il est implicite que x kilos de pommes coûteront $2x$ euros. Le prix des pommes est donc donné par la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par $x \mapsto 2x$. Le prix que vous payez est une fonction linéaire du poids. De manière plus générale, une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ est une application qui s'écrit sous la forme $x \mapsto \alpha x$, où $\alpha \in \mathbb{R}$ est donné. Finalement, dans $\mathbb{R}$, l'algèbre linéaire se réduit plus ou moins à la règle de trois. . . Le concept de linéarité peut alors s'étendre pour désigner un rapport de dépendance très simple entre plusieurs variables : la variable y *dépend linéairement des variables* $x_1, ..., x_N$ s'il existe des constantes $\alpha_1, \ldots, \alpha_N$ telles que $y = \alpha_1 x_1 + \ldots + \alpha_N x_N$. On dit encore que y s'exprime comme combinaison linéaire de $x_1, \ldots, x_N$. Par exemple, si vous achetez x_1 kilos de pommes à 2 euros le kilo et x_2 kilos de poires à 3 euros le kilo, le prix total y que vous payez est la combinaison linéaire $y = 2x_1 + 3x_2$. La notion de combinaison linéaire sera fondamentale dans la suite de ce cours.

[1]Al-Khwarizmi né vers 783, originaire de Khiva dans la région du Khwarezm qui lui a donné son nom, mort vers 850 à Bagdad, est un mathématicien, géographe, astrologue et astronome perse dont les écrits, rédigés en langue arabe, ont permis l'introduction de l'algèbre en Europe. Il est à l'origine des mots algorithme (qui n'est autre que son nom latinisé) et algèbre (issu d'une méthode et du titre d'un de ces ouvrages) ou encore de l'utilisation des chiffres arabes dont la diffusion dans le Moyen-Orient et en Europe provient d'un autre de ces livres (qui lui-même traite des mathématiques indiennes) et de l'habitude de désigner l'inconnue par la lettre x dans une équation. Il a d'ailleurs participé à la traduction de nombreux manuscrits scientifiques grecs et indiens. Son apport en mathématiques fut tel qu'il est également surnommé « le père de l'algèbre », avec Diophante dont il reprendra les travaux. En effet, il fut le premier à répertorier de façon systématique des méthodes de résolution d'équations en classant celles-ci.

Finalement, l'*algèbre linéaire* est le domaine des mathématiques qui étudie de façon systématique les propriétés associées à la dépendance linéaire. Les concepts de base sont celui de *combinaison linéaire* dont on vient de parler, et les notions d'*espace vectoriel* et d'*application linéaire*. Les espaces vectoriels les plus simples sont les ensembles $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$ lorsqu'ils sont munis de deux opérations très simples : l'addition et la multiplication par un réel, qui présentent un certain nombre de propriétés que nous verrons plus loin.

Bibliographie

- **Des livres en français.**

  D.-C. Lay *Algèbre linéaire : Théorie, exercices et applications*, 2004, De Boeck.
  R. Dalang- et A Chaabouni *Algèbre linéaire : Aide mémoire, exercices et applications*, 2005, PPUR.
  A. Denmat et F. Héaulme *Algèbre linéaire : Travaux Dirigés*, 1995, Dunod.
  F. Liret, D. Martinais *Algèbre* $1^{ère}$ *année. Cours et exercices avec solutions*, 2003, Dunod.

- **Si vous voulez en savoir plus....**

  J.-M. Monnier *Algèbre MPSI : Cours, méthodes et exercices corrigés*, 2006, Dunod.

- **Des livres en anglais.**

  G. Strang *Introduction to Linear Algebra, fourth edition* , 2009, Wellesley-Cambridge Press U.S. *La lecture de ce livre st fortement conseillé. Le premier chapitre est très largement inspiré de ce livre.*
  J. H Hubbard, B. Burke Hubbard *Vector Calculus, Linear Algebra, and Differential Forms : A Unified Approach*, 2005, Prentice Hall.

- **Quelques sites utiles (cours, exercices corrigés etc...)**

  `http://wims.unice.fr`

  `http://wims.auto.u-psud.fr/wims/wims.cgi?module=U1/algebra/docsyslin.fr`

  `http://web.mit.edu/18.06/www/Video/video-fall-99-new.html`

  `http://home.scarlet.be/~ping1339/Pvect.htm`

Chapitre 1

Algèbre linéaire dans $\mathbb{R}^2$ et $\mathbb{R}^3$

Nous allons, dans ce chapitre, introduire quelques notions nouvelles dans un cadre concret, celui de la droite, du plan et de l'espace, ce qui nous permettra aussi de réviser quelques connaissances que vous avez acquises en secondaire et au premier semestre.

1. Vecteurs de $\mathbb{R}^2$ et $\mathbb{R}^3$

1.1 Vecteurs et combinaisons linéaires

Dans ce premier chapitre, nous noterons en **gras** un vecteur $\boldsymbol{u}$ de $\mathbb{R}^2$ ou de $\mathbb{R}^3$ que vous avez peut être noté $\overrightarrow{u}$ dans vos classes précédentes. On s'affranchira de la notation en gras ou avec flèche au fur et à mesure de ce cours, en particulier lorsqu'on introduira les vecteurs comme des « éléments d'un espace vectoriel », dont nous donnerons une définition précise plus loin.

Soient $\boldsymbol{u}$ et $\boldsymbol{v}$ des vecteurs de $\mathbb{R}^2$ qui sont définis par des couples de réels (u_1, u_2) et (v_1, v_2). Dans le cadre du calcul matriciel, introduit au chapitre suivant, on notera aussi les vecteurs sous forme de *colonnes* contenant les composantes :

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

L'addition des deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ s'écrit :

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}.$$

On peut multiplier un vecteur $\boldsymbol{u} \in \mathbb{R}^2$ (ou $\mathbb{R}^3$) par un réel $\alpha \in \mathbb{R}$:

$$\alpha \boldsymbol{u} = \alpha \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \end{bmatrix}.$$

Définition 1.1

*On appelle **combinaison linéaire** de $\boldsymbol{u}$ et $\boldsymbol{v} \in \mathbb{R}^2$ tout vecteur $\boldsymbol{w}$ de la forme $\boldsymbol{w} = \alpha\boldsymbol{u} + \beta\boldsymbol{v}$, où α et β sont des réels, c.à.d. :*

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \alpha u_1 + \beta v_1 \\ \alpha u_2 + \beta v_2 \end{bmatrix}.$$

Ces propriétés s'appliquent bien sûr aussi aux vecteurs de $\mathbb{R}^3$. Un vecteur $\boldsymbol{u}$ de $\mathbb{R}^3$ est donné par ses trois composantes (u_1, u_2, u_3), et le vecteur s'écrit alors :

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

La combinaison linéaire de $\boldsymbol{u}$ et $\boldsymbol{v}$ dans $\mathbb{R}^3$ avec les coefficients α et β s'écrit

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \alpha u_1 + \beta v_1 \\ \alpha u_2 + \beta v_2 \\ \alpha u_3 + \beta v_3 \end{bmatrix}.$$

Remarque : On a identifié des couples ou des triplets de réels avec des vecteurs écrits sous forme de colonnes. Toutefois, il faudra bien faire attention de ne pas les confondre avec le *vecteur ligne* $\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$ dont les composantes sont les mêmes que celles du vecteur $\boldsymbol{u}$ mais qui n'est pas du tout le même objet mathématique. Ce vecteur ligne s'appelle « vecteur transposé » du vecteur $\boldsymbol{u}$.

On a défini précédemment une combinaison linéaire de deux vecteurs. Cette définition inclut le cas d'une combinaison linéaire d'un seul vecteur, en prenant le deuxième égal au vecteur nul. Mais on peut aussi bien sûr généraliser la définition de combinaison linéaire pour trois, quatre, ..., n vecteurs. Une question importante qui reviendra souvent pendant ce cours est justement de déterminer l'ensemble de toutes les combinaisons linéaires d'un vecteur, de deux vecteurs, de trois vecteurs ? Evidemment, le résultat dépend des vecteurs... Si par exemple on cherche toutes les combinaisons linéaires $\alpha\boldsymbol{u}$ et que le vecteur $\boldsymbol{u}$ est le vecteur nul, on obtient que l'ensemble des combinaisons linéaires est réduit au vecteur nul. Si par contre le vecteur $\boldsymbol{u}$ est non nul, l'ensemble des combinaisons linéaires est la droite de vecteur directeur $\boldsymbol{u}$. Par exemple, l'ensemble des combinaisons linéaires des deux vecteurs

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

est un plan, tandis que l'ensemble des combinaisons linéaires des deux vecteurs :

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

est une droite, car le second est lui-même combinaison linéaire du premier.

1.2 Produit scalaire, norme

Définition 1.2
Le ***produit scalaire*** *de deux vecteurs* $\boldsymbol{u} = (u_1, u_2)$ *et* $\boldsymbol{v} = (v_1, v_2)$ *de* $\mathbb{R}^2$ *est le réel :* $\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2$*. De même, le produit scalaire de deux vecteurs* $\boldsymbol{u} = (u_1, u_2, u_3)$ *et* $\boldsymbol{v} = (v_1, v_2, v_3)$ *de* $\mathbb{R}^3$ *est le réel :* $\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + u_3v_3$*.*

On rappelle que deux vecteurs sont dits *orthogonaux* si leur produit scalaire est nul.

Définition 1.3
La ***norme euclidienne*** *d'un vecteur* $\boldsymbol{u}$ *de* $\mathbb{R}^2$ *ou* $\mathbb{R}^3$ *est le réel positif ou nul* $\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$*.*

On appelle **vecteur unitaire** un vecteur $\boldsymbol{u}$ dont la norme est égale à 1 : $\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}} = 1$. Si $\boldsymbol{u}$ et $\boldsymbol{v}$ sont deux vecteurs unitaires, alors $\boldsymbol{u} \cdot \boldsymbol{v} = \cos\theta$, où θ est l'angle entre les vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$. On remarque donc que le produit scalaire de deux vecteurs unitaires est toujours compris entre -1 et 1.

Pour deux vecteurs quelconques $\tilde{\boldsymbol{u}}$ et $\tilde{\boldsymbol{v}}$, la formule donnant le cosinus de l'angle θ entre les deux vecteurs devient :

$$\cos\theta = \frac{\tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{v}}}{\|\tilde{\boldsymbol{u}}\|\|\tilde{\boldsymbol{v}}\|},$$

ce qui se démontre très facilement en posant $\boldsymbol{u} = \frac{\tilde{\boldsymbol{u}}}{\|\tilde{\boldsymbol{u}}\|}$ et $\boldsymbol{v} = \frac{\tilde{\boldsymbol{v}}}{\|\tilde{\boldsymbol{v}}\|}$: les deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ sont maintenant unitaires et on a donc $\boldsymbol{u} \cdot \boldsymbol{v} = \cos\theta$, d'où le résultat[2].

On rappelle enfin deux inégalités *absolument fondamentales* : pour tous vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ de $\mathbb{R}^2$ ou $\mathbb{R}^3$,

Inégalité de Cauchy-Schwarz :	$\lvert\boldsymbol{u} \cdot \boldsymbol{v}\rvert \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$
Inégalité triangulaire :	$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

2. Systèmes d'équations linéaires

2.1 Différentes ecritures d'un système linéaire

Commençons par un exemple dans $\mathbb{R}^2$. Considèrons le système linéaire suivant :
Chercher x_1 et x_2 , tels que :

$$\begin{cases} x_1 - x_2 = 1 \\ 3x_1 + 2x_2 = 5. \end{cases} \tag{1.1}$$

On va voir qu'on peut l'écrire sous différentes formes.
En effet on peut utiliser la notion de combinaison linéaire que nous venons de voir pour

[2] Ce raisonnement s'appelle un raisonnement par homogénéité.

écrire le système (1.1) sous la forme :
Chercher x_1 et x_2, tels que :

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

Ce qui revient à dire :
Chercher la combinaison linéaire des vecteurs $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ qui est égale au vecteur $\begin{bmatrix} 1 \\ 5 \end{bmatrix}$.
Mais on peut aussi utiliser la notion de produit scalaire que nous venons de rappeler, en écrivant :
Chercher le vecteur $\boldsymbol{x}$ de $\mathbb{R}^2$ de composantes (x_1, x_2) tel que :

$$\begin{bmatrix} 2x_1 - x_2 \\ x_1 + 2x_2 \end{bmatrix} = \begin{bmatrix} (2,-1) \cdot (x_1, x_2) \\ (1,2) \cdot (x_1, x_2) \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}.$$

Ce qui revient à dire :
Chercher le vecteur $\boldsymbol{x}$ de $\mathbb{R}^2$ tel le produit scalaire de ce vecteur avec le vecteur $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ soit égal à 1 et le produit scalaire de ce vecteur avec le vecteur $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ soit égal à 5.

Ces différentes écritures d'un système linéaire sont possibles quelque soit le nombre d'équations et le nombre d'inconnues. Par exemple dans $\mathbb{R}^3$, considérons le système linéaire suivant, où les réels b_1, b_2 et b_3 sont donnés :
Chercher x_1, x_2 et x_3 tels que :

$$\begin{cases} x_1 & = b_1 \\ -x_1 + x_2 & = b_2 \\ -x_2 + x_3 & = b_3 \end{cases} \tag{1.2}$$

En utilisant les combinaisons linéaires, on pourra écrire (1.2) sous la forme :
Chercher x_1, x_2 et x_3 tels que :

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Ou bien encore, en utilisant le produit scalaire, sous la forme :
Chercher le vecteur $\boldsymbol{x}$ de $\mathbb{R}^3$ de composantes (x_1, x_2, x_3) tel que :

$$\begin{bmatrix} (1,0,0) \cdot (x_1, x_2, x_3) \\ (-1,1,0) \cdot (x_1, x_2, x_3) \\ (0,-1,1) \cdot (x_1, x_2, x_3) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

2.2 Remarques sur les solutions d'un système linéaire

Notons que le système linéaire de trois équations à trois inconnues ci-dessus est particulièrement simple à résoudre : on a directement x_1 à partir de la première équation, on peut ensuite substituer x_1 dans la deuxième équation et trouver x_2, et enfin substituer x_1 et x_2 dans la troisième équation et trouver x_3. Et on obtient donc la solution :

$$\begin{cases} x_1 = b_1 \\ x_2 = b_1 + b_2 \\ x_3 = b_1 + b_2 + b_3. \end{cases}$$

On voit donc que quelque soit les valeurs que prennent les trois réels b_1, b_2 et b_3 ce système possède une unique solution.

Considérons, maintenant le système suivant :
Chercher x_1, x_2 et x_3 tels que :

$$\begin{cases} x_1 - x_3 & = b_1 \\ -x_1 + x_2 & = b_2 \\ -x_2 + x_3 & = b_3 \end{cases} \tag{1.3}$$

Ce système est nettement moins facile à résoudre que le précédent, et même, en fait il est impossible de trouver la solution du système, vu que le système admet une infinité de solutions, ou au contraire pas de solution du tout. Par exemple, si on prend $b_1 = b_2 = b_3 = 0$ on a la solution :

$$\overline{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

mais aussi tous les vecteurs de la forme $\alpha\overline{\boldsymbol{x}}$, $\alpha \in \mathbb{R}$. Mais par contre, si on prend $b_1 = 1$ et $b_2 = b_3 = 0$, alors le système ne possède aucune solution car la somme des trois composantes du vecteur $\begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$ vaut zéro alors que la somme des trois composantes du second membre $\boldsymbol{b}$ vaut 1. Géométriquement, ceci revient à dire qu'il n'existe pas de combinaison linéaire des trois vecteurs $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ et $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ qui donne le vecteur $\boldsymbol{b} = (1, 0, 0)$. Donc l'ensemble des combinaisons linéaires de ces trois vecteurs ne remplit pas tout l'espace $\mathbb{R}^3$. Pour que le système ait une solution, il faut que la somme des trois composantes du second membre soit nulle. En d'autres termes, l'ensemble des combinaisons linéaires de ces trois vecteurs sont dans le plan d'équation $b_1 + b_2 + b_3 = 0$, et donc ne remplissent pas tout l'espace $\mathbb{R}^3$.

3. Méthode du pivot de Gauss

3.1 Ecriture vectorielle et matricielle des systèmes linéaires

Considérons le système de deux équations à deux inconnues :

$$\begin{cases} x + 2y = 4 \\ 2x - 3y = 1 \end{cases} \tag{1.4}$$

Regardons ce qui se passe ligne par ligne. Chaque ligne donne l'équation d'une droite, qu'on représente dans la figure 1.1 Pour que le couple (x, y) vérifie les deux équations, il faut donc que le point de coordonnées x et y soit l'intersection des deux droites. C'est ce qu'on appelle la vision « par lignes » du système : la solution du système 1.4 est l'intersection de deux droites. Voyons maintenant un peu ce qui se passe si on regarde

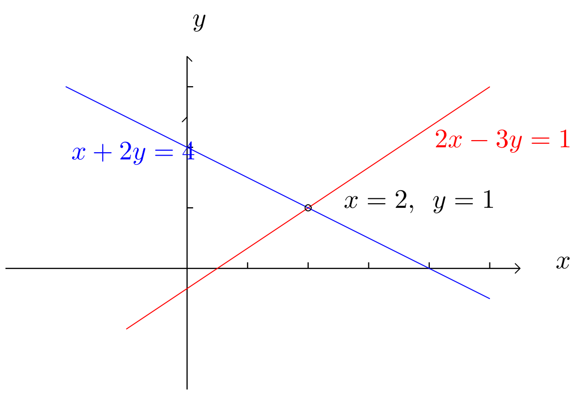


Figure 1.1 *Vision par lignes* : la solution du système est l'intersection des droites

les choses « par colonnes ». On va maintenant lire le système 1.4 comme une équation vectorielle faisant intervenir les vecteurs :

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \; \boldsymbol{v} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \text{ et } \boldsymbol{b} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

Avec ces vecteurs, on peut réécrire le système 1.4 comme une seule équation vectorielle

$$x\boldsymbol{u} + y\boldsymbol{v} = \boldsymbol{b}.$$

On cherche la « bonne » combinaison linéaire de $\boldsymbol{u}$ et $\boldsymbol{v}$ (c.à.d. les bons coefficients x et y) qui va donner $\boldsymbol{b}$, comme on le voit sur la figure 1.2.

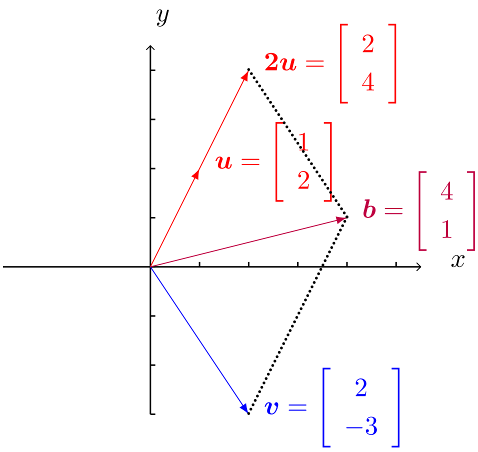


Figure 1.2 *Vision par colonnes* : $\boldsymbol{b} = 2\boldsymbol{u} + \boldsymbol{v}$

3.2 Elimination

Un exemple 2×2

Vous avez vu en secondaire comment résoudre un système par élimination et substitution. On va étudier cette année une procédure systématique d'élimination, celle qui est utilisée dans les programmes informatiques pour la résolution des systèmes linéaires, et qui est connue sous le nom de méthode de Gauss [3]. Sur le système précédent,

$$\left\{\begin{array}{l} x + 2y = 4 \\ 2x - 3y = 1 \end{array}\right. \quad \text{devient} \quad \left\{\begin{array}{ll} x + 2y = 4 & \text{(multiplier la première équation par 2)} \\ -7y = -7 & \text{(puis soustraire à la deuxième)} \end{array}\right. \tag{1.5}$$

La deuxième équation donne alors $y = 1$, puis en substituant cette valeur dans la première $x = 2$. L'étape d'élimination produit un système dont la matrice est triangulaire supérieure (les coefficients sous la diagonale sont nuls). Une fois que la matrice du système est sous forme triangulaire supérieure, il est très facile de le résoudre par substitution.

La solution du système d'origine et du système après élimination est la même.

Même si ce système est très simple à résoudre et que vous auriez très certainement réussi à le faire sans ce cours d'algèbre linéaire, cela vaut le coup d'analyser les opérations qu'on a effectuées pour comprendre la méthode et l'appliquer dans des cas plus compliqués.

Pour éliminer x dans la deuxième équation, on a multiplié la première équation par -2 et on l'a ajoutée à la deuxième. On a utilisé pour cela le fait que le premier coefficient de la première ligne est 1, et en particulier non nul : on dit que c'est le **pivot**. Comme le

[3] Johann Carl Friedrich Gauss (30 avril 1777 – 23 février 1855) est un mathématicien, astronome et physicien allemand. Il est considéré comme l'un des plus grands mathématiciens de tous les temps.

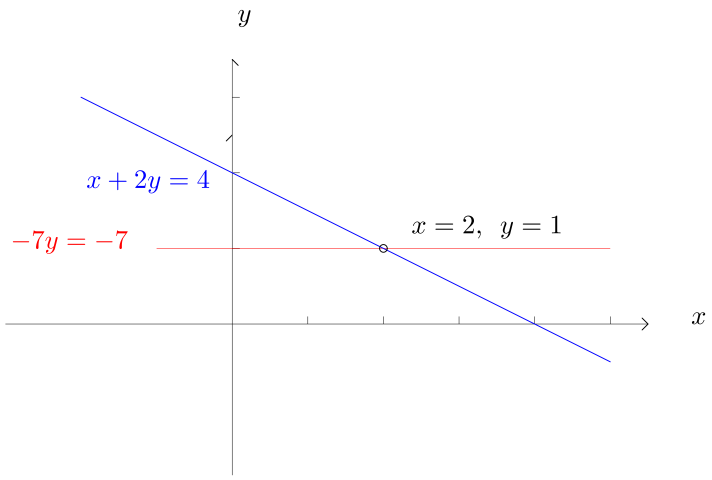


Figure 1.3 le système après élimination dans la deuxième équation : *Vision par lignes* : la solution du système est l'intersection des droites, elle n'a pas changé.

coefficient devant x dans la deuxième équation est 2, le multiplicateur pour l'élimination est donc -2.

Prenons le même système, mais où on a multiplié la première ligne par 5 : l'élimination va se faire maintenant de la manière suivante :

$$\left\{\begin{array}{lcll} 5x & + & 10y & = 20 \\ 2x & - & 3y & = 1 \end{array}\right. \tag{1.6}$$

devient

$$\left\{\begin{array}{lcllr} 5x & + & 10y & = 20 & \text{(multiplier la première équation par } -\frac{2}{5} \\ & - & 7y & = -7 & \text{(puis ajouter à la deuxième)} \end{array}\right. \tag{1.7}$$

La solution reste évidemment la même, mais maintenant, le premier coef de la première ligne est 5. Comme le coefficient devant x dans la deuxième équation reste égal 2, le multiplicateur est maintenant $-\frac{2}{5}$.

La deuxième équation a elle aussi un pivot, qui est égal à -7, et qui permet d'obtenir la solution (on l'utiliserait pour éliminer y dans la troisème équation s'il y en avait une). Pour résoudre un système de deux équations à deux inconnues, on a utilisé 2 pivots. Pour résoudre un système de n équations à n inconnues, on aura besoin de n pivots. Notez qu'à la fin de l'élimination, le système obtenu a une forme triangulaire, et que les pivots sont sur la diagonale du « triangle ».

Voyons maintenant si les choses peuvent se passer plus mal (et oui... elles le peuvent !)

3.3 Echecs possibles de l'élimination de Gauss

Si on programme l'algorithme de Gauss comme expliqué au paragraphe précédent et qu'on l'applique à n'importe quelle matrice, il est possible que le résultat soit « `NaN` » ce qui veut dire en décodé « Not a Number » : l'ordinateur vous répond que vous essayez de diviser par zéro et qu'il ne sait pas faire... petite explication sur trois exemples faciles, déduits du précédent.

1. Exemple d'échec total de l'élimination : pas de solution

Avec le système suivant :

$$\left\{\begin{array}{lcll} x & - & 2y & = 4 \\ 2x & - & 4y & = 1 \end{array}\right. \tag{1.8}$$

devient

$$\left\{\begin{array}{lcllr} x & - & 2y & = 4 & \text{(multiplier la première équation par 2)} \\ & & \mathbf{0}y & = -7 & \text{(puis soustraire à la deuxième)} \end{array}\right. \tag{1.9}$$

l'élimination ne marche pas parce qu'après élimination de x, la deuxième équation a un zéro devant le y, et on ne peut pas diviser par 0 (comme votre ordinateur vous le fera savoir si vous essayez...) ; on dit aussi qu'on a « un zéro en position pivotale », mais il est absolument interdit de dire « un pivot nul », parce qu'**un pivot n'est jamais nul**.

Cet échec peut être interprété géométriquement (vision « lignes ») par le fait que les deux équations du sytème sont les équations de deux droites parallèles et non confondues, et donc ont une intersection vide.

On peut aussi l'interpréter vectoriellement (vision « colonnes ») par le fait qu'on ne peut pas atteindre le vecteur $(4, 1)$ par une combinaison linéaire des vecteurs $(1, 2)$ et $(-2, -4)$

2. Exemple d'échec de l'élimination, mais infinité de solution

Dans l'exemple précédent, on remplace le second membre $(4, 1)$ par $(4, 8)$:

$$\left\{\begin{array}{lcll} x & + & 2y & = 4 \\ 2x & + & 4y & = 8 \end{array}\right. \tag{1.10}$$

devient

$$\left\{\begin{array}{lcllr} x & + & 2y & = 4 & \text{(multiplier la première équation par 2} \\ & & \mathbf{0}y & = 0 & \text{puis soustraire à la deuxième)} \end{array}\right. \tag{1.11}$$

Là encore on a un zéro en position pivotale, et votre ordinateur risque de ne pas aimer si vous n'avez pas mis de test dans votre programme.... Mais contrairement à l'exemple précédent, on a maintenant une solution au système, et même une infinité de solutions : en effet, tout $y \in \mathbb{R}$ satisfait la deuxième équation et une fois qu'on a choisi un y, on obtient x par la première. Dans la vision « lignes », les droites qui étaient parallèles dans l'exemple précédent sont maintenant confondues. Dans la vision colonne, le second membre $\boldsymbol{b} = (4,8)$ est maintenant sur la même droite que les vecteurs colonnes $(1,2)$ et $(2,4)$ de la matrice du système.

3. Exemple d'échec temporaire de l'élimination : deux pivots obtenus par échange de ligne

Supposons qu'on ait un zéro en position pivotale de la première ligne :

$$\begin{cases} 0x & + & 2y & = 4 \\ 2x & - & 3y & = 1 \end{cases} \tag{1.12}$$

devient, après échange des deux lignes :

$$\begin{cases} 2x & - & 3y & = 1 \\ 0x & + & 2y & = 4 \end{cases} \tag{1.13}$$

Le nouveau système est sous forme triangulaire, et il est donc prêt pour l'étape de substitution, dite aussi étape de remontée. La deuxième équation donne $y = 2$ puis la première $x = \frac{7}{2}$. Les pivots du système sont 2 et 2, mais pour les obtenir on a dû échanger les lignes.

Les deux premiers exemples sont des systèmes non inversibles (ou singuliers) : dans les deux cas on n'a pas de pivot pour la seconde équation (zéro en position pivotale). Les systèmes singuliers ont soit aucune solution, soit une infinité de solutions. Le deuxième exemple fait apparaître un système inversible (ou régulier).

3.4 Théorie générale pour les systèmes 2×2

Lemme 1.4 *Soient* $a, b, c, d, \alpha, \beta$ *des réels. On suppose* $a \neq 0$ *. Alors le système*

$$\begin{cases} ax+ & by & = \alpha \\ cx+ & dy & = \beta \end{cases}$$

est équivalent au système

$$\begin{cases} ax+ & by & = \alpha \\ & (d - \frac{bc}{a})y & = \beta - \frac{c}{a}\alpha \end{cases}$$

Démonstration. Soit $(x,y) \in \mathbb{R}^2$. Si (x,y) est solution du premier système, alors dans ce premier système, on multiplie la première équation par $-c/a$ (on a le droit car $a \neq 0$) et on l'additionne à la seconde équation. On vérifie ainsi que (x,y) est solution du second système. Réciproquement, si (x,y) est solution du second système, alors dans ce second système, on multiplie la première équation par c/a et on l'ajoute à la seconde équation. On vérifie ainsi que (x,y) est solution du premier système. On conclut que les deux systèmes ont même ensemble de solution. Ils sont donc équivalents. $\square$

Théorème 1.5 *Soient* a, b, c, d *des réels et* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Le système* $A\boldsymbol{x} = \boldsymbol{b}$ *admet une solution unique pour tout* $\boldsymbol{b} \in \mathbb{R}^2$ *si et seulement si la matrice* A *admet deux pivots lors de l'élimination de Gauss.*

Démonstration. Dire que la matrice A admet deux pivots lors de l'élimination de Gauss signifie que, quitte à échanger les deux lignes de A,

- le premier coefficient de la première ligne de A est non nul, et que
- après avoir fait apparaître un 0 en première position de la seconde ligne de A, le coefficient en deuxième position est non nul.

Par le Lemme, on peut résumer cela en

- soit $a \neq 0$ et $d - bc/a \neq 0$,
- soit $a = 0$, et dans ce cas on intervertit les deux lignes et on a $c \neq 0$ et $b - da/c \neq 0$.

Dans chacun de ces deux cas, l'étape de remontée montre par son procédé constructif qu'il existe une unique solution au système. On a ainsi montré que l'existence de deux pivots entraine l'existence et unicité de la solution du système.

Montrons maintenant que si on n'a pas deux pivots alors on n'a pas existence et unicité. Si on n'a pas deux pivots en sortie de l'élimination, on n'est donc dans aucun des cas ci-dessus, et alors on est dans l'une des situations suivantes :

- $a = 0 = c$: on est ramené au système
$$\begin{cases} by &= \alpha \\ dy &= \beta \end{cases}$$
qui a 0 ou une infinité de solution,
- $a \neq 0$ et $d - bc/a = 0$: on est ramené au système
$$\begin{cases} ax+ & by &= \alpha \\ & 0y &= \beta - \frac{c}{a}\alpha \end{cases}$$
qui a 0 ou une infinité de solution,
- $c \neq 0$ et $b - da/c = 0$: on est ramené au système
$$\begin{cases} cx+ & dy &= \beta \\ & 0y &= \alpha - \frac{a}{c}\beta \end{cases}$$
qui a 0 ou une infinité de solution.

Définition 1.6

Soient a, b, c, d *des réels. On appelle* ***déterminant*** *de la matrice* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *et on note* $\det(A)$ *le réel défini par* $\det(A) = ad - bc$.

Proposition 1.7 *Si la matrice* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *a deux pivots, le déterminant de la matrice* A *est égal au produit des pivots. Si le nombre de pivots est strictement inférieur à 2,* $\det(A) = 0$. *Le système* $A\boldsymbol{x} = \boldsymbol{b}$ *admet donc une solution unique pour tout* $\boldsymbol{b} \in \mathbb{R}^2$ *si et seulement si* $\det(A) \neq 0$.

Démonstration. Voir exercice 22. □

3.5 Un système 3×3

On va maintenant effectuer l'élimination de Gauss sur le système 3×3 suivant :

$$\left\{\begin{array}{rcrcl} 2x_1 & + & 4x_2 & - & 2x_3 = 2 \\ 4x_1 & + & 9x_2 & - & 3x_3 = 8 \\ -2x_1 & - & 3x_2 & + & 7x_3 = 10 \end{array}\right. \tag{1.14}$$

Le premier pivot est le premier coefficient non nul de la première ligne, c.à.d. 2. On utilise ce pivot pour annuler les coefficients de x_1 dans les lignes 2 et 3. On soustrait 2 fois la ligne 1 à la ligne 2, et on soustrait (-1) fois la ligne 1 à la ligne 3 :

$$\left\{\begin{array}{rcrcl} \mathbf{2}x_1 & + & 4x_2 & - & 2x_3 = 2 \\ 4x_1 & + & 9x_2 & - & 3x_3 = 8 \\ -2x_1 & - & 3x_2 & + & 7x_3 = 10 \end{array}\right. \xrightarrow[\substack{\ell_2 \text{-} \ell_2 - 2\ell_1 \\ \ell_3 \text{-} \ell_3 - (-1)\ell_1}]{} \left\{\begin{array}{rcrcl} \mathbf{2}x_1 & + & 4x_2 & - & 2x_3 = 2 \\ & & x_2 & + & x_3 = 4 \\ & & x_2 & + & 5x_3 = 12 \end{array}\right.$$

Dans les formules ci-dessus la phrase ℓ_2 - $\ell_2 - 2\ell_1$ est à comprendre par « la ligne 2 est remplacée par la ligne 2 - deux fois la ligne 1 ». On cherche maintenant le pivot de la deuxième équation, il se trouve que c'est le coefficient de x_2, égal à 1. On utilise ce pivot pour annuler le coefficient de x_2 dans la ligne 3 :

$$\left\{\begin{array}{rcrcl} \mathbf{2}x_1 & + & 4x_2 & - & 2x_3 = 2 \\ & & \mathbf{1}x_2 & + & x_3 = 4 \\ & & x_2 & + & 5x_3 = 12h \end{array}\right. \xrightarrow[\ell_3 \text{-} \ell_3 - \ell_2]{} \left\{\begin{array}{rcrcl} \mathbf{2}x_1 & + & 4x_2 & - & 2x_3 = 2 \\ & & \mathbf{1}x_2 & + & x_3 = 4 \\ & & & & 4x_3 = 8 \end{array}\right.$$

La troisième ligne comporte un pivot égal à 4 devant x_3 et donc le système est transformé par l'élimination de Gauss en un système triangulaire supérieur :

$$\left\{\begin{array}{rcrcl} 2x_1 & + & 4x_2 & - & 2x_3 = 2 \\ 4x_1 & + & 9x_2 & - & 3x_3 = 8 \\ -2x_1 & - & 3x_2 & + & 7x_3 = 10 \end{array}\right. \tag{1.15}$$

devient, après élimination de Gauss :

$$\left\{\begin{array}{rcrcrl} \mathbf{2}x_1 & + & 4x_2 & - & 2x_3 & = 2 \\ & & \mathbf{1}x_2 & + & 1x_3 & = 4 \\ & & & & \mathbf{4}x_3 & = 8. \end{array}\right. \tag{1.16}$$

où les pivots sont écrits en gras. Ceci peut encore s'écrire

$$A\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow U\boldsymbol{x} = \boldsymbol{c},$$

$$\text{avec } A = \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix},\ \boldsymbol{b} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix},\ U = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix},\ \text{ et } \boldsymbol{c} = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}.$$

On effectue ensuite une remontée pour trouver la solutions $\boldsymbol{x}$ du système :

La troisième équation $4x_3 = 8$ donne $x_3 = 2$

La deuxième équation $x_2 + x_3 = 4$ donne $x_2 = 2$

La première équation $2x_1 + 4x_2 - 2x_3 = 2$ donne $x_1 = -1$.

Dans la vision en lignes, ceci veut dire que l'intersection des trois plans dont les équations sont celles du système 1.14 est le point $(-1, 2, 2)$. Dans la vision en colonnes, ceci veut dire qu'une combinaison linéaire de vecteurs colonnes donne le second membre $\boldsymbol{b}$.

Chapitre 2

Calcul matriciel

Dans ce qui suit $\mathbb{K}$ désignera le corps des réels $\mathbb{R}$ ou bien celui des complexes $\mathbb{C}$.

1. Définitions et propriétés

On appelle *Matrice à coefficients dans* $\mathbb{K}$ un tableau de nombres représenté sous la forme suivante (par exemple) :

$$\begin{bmatrix} 1 & 0 \\ \pi & -5 \\ e & 3 \end{bmatrix} \quad \text{est une matrice } 3 \times 2$$

ou

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{est une matrice } 2 \times 2$$

ou

$$\begin{bmatrix} i & e^{i\frac{\pi}{2}} & 0 \\ -1 & e & 1 \end{bmatrix} \quad \text{est une matrice } 2 \times 3$$

ou encore

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{est une matrice } 2 \times 2$$

en notation générale

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{np} \end{bmatrix},$$

est une matrice à n lignes et p colonnes. On peut abréger l'écriture en

$$A = \begin{bmatrix} a_{ij} \end{bmatrix}_{1 \leq i \leq n; 1 \leq j \leq p}$$

La notation a_{ij} n'est pas standard on peut aussi adopter a_i^j .

On notera $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à coefficients dans $\mathbb{K}$ à n lignes et p colonnes.

Quelques matrices particulières :

- les matrices **carrées** $n \times n$, dont l'ensemble est noté $\mathcal{M}_n(\mathbb{K})$.
- les matrices **lignes** $[a_1 \;\; a_2 \ldots a_n]$,
- les matrices **colonnes** $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$
- la **matrice nulle** $O_{n,p} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$, qui peut être rectangulaire ou carrée.
- Les matrices suivantes sont toutes des matrices <u>carrées</u>, *i.e.* des éléments de $\mathcal{M}_n(\mathbb{K})$.
  - la **matrice identité** $\mathrm{Id}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$,
  - les **matrices diagonales** $A = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{n,n} \end{bmatrix}$,
  - les **matrices triangulaires supérieures** $U = \begin{bmatrix} \star & \star & \cdots & \star \\ 0 & \star & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \star \end{bmatrix}$,

    et **inférieures** $L = \begin{bmatrix} \star & 0 & \cdots & 0 \\ \star & \star & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \star & \cdots & \star & \star \end{bmatrix}$.

NB : les matrices 1×1 sont identifiées à $\mathbb{K}$, on ne distinguera pas $[a]$ et $a \in \mathbb{K}$.

1.1 Opérations sur les matrices

1) L'addition

C'est une opération interne dans $\mathcal{M}_{n,p}(\mathbb{K})$ définie, pour tous $A = \begin{bmatrix} a_{ij} \end{bmatrix}, B = \begin{bmatrix} b_{ij} \end{bmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$, par

$$A + B = \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix}$$

exemple

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

Attention : on n'ajoute que les matrices de même format.

Propriétés (évidentes) de l'addition :

- associativité
- commutativité
- il existe un élément neutre : la matrice dont tous les coefficients sont nuls elle est notée $\mathbf{0}$.
- toute matrice $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ admet une matrice opposée $-A = \begin{bmatrix} -a_{ij} \end{bmatrix}$.

2) Le produit par un scalaire

Pour une matrice $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ et pour tout $\lambda \in \mathbb{K}$, le produit est défini par

$$\lambda A = \begin{bmatrix} \lambda a_{ij} \end{bmatrix}$$

le résultat est une matrice de même format que A. Les propriétés évidentes de cette loi sont :

$$\begin{aligned} (\lambda\mu)A &= \lambda(\mu A) \\ (\lambda+\mu)A &= \lambda A + \mu A \\ \lambda(A+B) &= \lambda A + \lambda B \\ 1.A &= A \\ 0.A &= \mathbf{0} \end{aligned}$$

3) La multiplication des matrices.

Avant de la définir on va en donner un exemple. Si :

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

On pose l'opération

$$\begin{bmatrix} 3 & 1 & 2 \\ -1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ -1 & 0 & -2 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 8 & 10 \\ 1 & 3 & -10 \end{bmatrix}$$

Alors

$$AB = \begin{bmatrix} 4 & 8 & 10 \\ 1 & 3 & -10 \end{bmatrix}$$

A et B n'ont pas nécessairement le même format, mais *le nombre de lignes de* B *doit être égal au nombre de colonnes de* A.

Dans le cas général, pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et pour $B \in \mathcal{M}_{p,m}(\mathbb{K})$ on posera

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1p} \\ & & \ddots & & \\ a_{i1} & & \dots & & a_{ip} \\ & & \ddots & & \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1m} \\ & \dots & b_{2j} & \dots & b_{2m} \\ & & \ddots & & \\ & \dots & b_{pj} & \dots & b_{pm} \end{bmatrix} = \begin{bmatrix} & & \vdots & & \\ & \dots & c_{ij} & \dots & \\ & & & & \\ & & \vdots & & \end{bmatrix}$$

où

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \quad \text{on a } AB \in \mathcal{M}_{n,m}(\mathbb{K})$$

Le coefficient c_{ij} est égal au produit scalaire de la $i^{\text{ième}}$ ligne de A par la $j^{\text{ième}}$ colonne de B.

IMPORTANT : le produit AB peut être défini sans que BA le soit (voir l'exemple ci-dessus).

Premières propriétés

$$\begin{aligned} A(BC) &= (AB)C \\ A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \end{aligned}$$

Ces propriétés seront démontrées de manière simple dans le chapitre sur les applications linéaires.

On a :

$$\begin{bmatrix} a_1 \dots a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^{n} a_i b_i \ ,$$

mais :

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \begin{bmatrix} a_1 \dots a_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 & \dots & a_n b_1 \\ & \ddots & \\ a_1 b_n & \dots & a_n b_n \end{bmatrix}$$

Le produit matriciel n'est donc pas commutatif.

Quelques produits particuliers

- Le résultat du produit d'une matrice $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et d'un vecteur colonne $X = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$, est un vecteur colonne de hauteur n qui est une combinaison linéaire des colonnes de A :
$$AX = \sum_{k=1}^{p} x_k c_k(A) = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \ldots + x_p \begin{bmatrix} a_{1p} \\ \vdots \\ a_{np} \end{bmatrix} ,$$
où les $c_k(A)$ représentent les colonnes de A
- Si $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$ est un vecteur ligne, le produit XA est un vecteur ligne $Y = \begin{bmatrix} y_1 & \cdots & y_p \end{bmatrix} \in \mathcal{M}_{1,p}(\mathbb{K})$ avec
$$y_j = \sum_{k=1}^{n} x_k a_{k,j},\ j = 1, \ldots, p.$$
- Si $X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$ est un vecteur ligne et $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$ est un vecteur colonne alors le produit $XY \in \mathcal{M}_{1,1}(\mathbb{K})$ est un nombre donné par
$$x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$
On reconnaît le produit scalaire des 2 vecteurs $\boldsymbol{x} = (x_1, \ldots, x_n)$ et $\boldsymbol{y} = (y_1, \ldots, y_n)$.
- Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on note $E_j \in \mathcal{M}_{p,1}(\mathbb{K})$, $F_i \in \mathcal{M}_{1,n}(\mathbb{K})$ les matrices
$$E_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j^{\text{ème}} \text{ ligne}, \quad F_i = \begin{bmatrix} 0 & \cdots & 0 & \underset{\substack{\uparrow \\ i^{\text{ème}} \text{ colonne}}}{1} & 0 & \cdots & 0 \end{bmatrix}.$$
Alors on a $AE_j = c_j(A)$, c'est-à-dire la $j^{\text{ème}}$ colonne de la matrice A et $F_i A = \ell_i(A)$, c'est-à-dire la $i^{\text{ème}}$ ligne de la matrice A.

1.2 Algèbre des matrices carrées

On se fixe un entier n. Le produit matriciel est alors une loi interne dans $\mathcal{M}_n(\mathbb{K})$:

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

On appelle *Identité d'ordre n* la matrice de $\mathcal{M}_n(\mathbb{K})$ définie par $\mathrm{I}d_n = \left[\ \delta_i^j\ \right]$ (symboles de Kronecker).

$$\mathrm{I}d_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

On a la propriété fondamentale

$$\forall A \in \mathcal{M}_n(\mathbb{K})\ ,\ \mathrm{I}d_n A = A\mathrm{I}d_n = A$$

Si A, non nécessairement carrée, est telle que $A\mathrm{I}d_n$ soit définie alors $A\mathrm{I}d_n = A$, de même $\mathrm{I}d_n A = A$ si le produit existe.

En résumé, $\mathcal{M}_n(\mathbb{K})$ est muni de deux lois internes

- Une addition notée $+$ qui est une loi de groupe commutatif.
- Un produit associatif, distributif à gauche et à droite par rapport à $+$ et admettant un élément neutre.
- D'une loi externe avec les propriétés déjà citées.

Muni de ces opérations $\mathcal{M}_n(\mathbb{K})$ est appelée *Algèbre des matrices d'ordre n à coefficients dans* $\mathbb{K}$.

1.3 Puissances d'une matrice carrée

$\forall A \in \mathcal{M}_n(\mathbb{K})$ on définit

$$A^0 = \mathrm{I}d_n,\ \ A^1 = A\ \ , A^2 = AA,\ \ A^m = AA^{m-1}\ .$$

Remarque : $(AB)^2 = ABAB$

On dira que deux matrices *commutent* si $AB = BA$, dans ce cas A et B sont nécessairement carrées de même ordre. Pour de telles matrices on a la formule du binôme :

$$(A+B)^k = \sum_{i=1}^{k} C_k^i A^i B^{k-i}$$

Cette formule est utilisée souvent dans le cas particulier :

$$(\mathrm{I}d_n + A)^k = \sum_{i=1}^{k} C_k^i A^i$$

1.4 Matrices inversibles

Définition 2.1
Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite inversible ou régulière s'il existe une matrice B telle que $AB = BA = \mathrm{Id}_n$.

Remarques :

- Seule une matrice carrée peut, éventuellement, être inversible.
- Si B existe alors $B \in \mathcal{M}_n(\mathbb{K})$.
- Pour A donnée, B est *unique en cas d'existence*, en effet, si B' vérifie $AB' = B'A = \mathrm{Id}_n$, alors $B' = BAB' = B$. Dans ce cas B sera notée A^{-1} et appelée matrice inverse de A.
- Si A est inversible alors $(A^{-1})^{-1} = A$.
- Le produit de deux matrices inversibles est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.
- **Important** [Inverse à gauche et à droite] En fait on peut montrer que si A est une matrice carrée et s'il existe une matrice « inverse à gauche » B telle que $BA = \mathrm{Id}_n$, alors A est inversible et $B = A^{-1}$. La démonstration de ce résultat nécessite des outils qu'on ne verra qu'un peu plus tard. Il est cependant bien pratique car il permet, lorsqu'on veut calculer l'inverse d'une matrice, de ne calculer que l'inverse à gauche (par l'algorithme de Gauss-Jordan qu'on verra prochainement) sans vérifier que $AB = \mathrm{Id}_n$.

L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est appelé *groupe linéaire d'ordre* n, il est noté $\mathrm{Gl}(n, \mathbb{K})$.

Définition 2.2
On appelle **transposée d'une matrice** $A \in \mathcal{M}_{n,p}(\mathbb{K})$*, la matrice de* $\mathcal{M}_{p,n}(\mathbb{K})$ *que l'on note* $A^t = \left[\ \alpha_{i,j}\ \right]$ *dont les coefficients sont définis par* $\alpha_{i,j} = a_{j,i}$*. Autrement dit la transposée de* A *est obtenue en permutant les lignes et les colonnes de* A*.*

Voici quelques propriétés de la transposition :

1. La transposition est involutive : $((A)^t)^t = A$.
2. Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$, alors $(A+B)^t \in \mathcal{M}_{p,n}(\mathbb{K})$ et $(A+B)^t = A^t + B^t$.
3. Soient $A \in \mathcal{M}_{n,p}(\mathbb{K}), B \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors $(AB)^t \in \mathcal{M}_{q,n}(\mathbb{K})$ et $(AB)^t = B^t A^t$.
4. Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors A^t est inversible et $(A^t)^{-1} = (A^{-1})^t$.

Les points 1. et 2. sont faciles. Pour le point 3., on a :

$$(AB)^t_{i,j} = (AB)_{j,i} = \sum_{k=1}^{p} a_{j,k} b_{k,i} = \sum_{k=1}^{p} A^t_{k,j} B^t_{i,k} = (B^t A^t)_{i,j}.$$

Enfin, si A est inversible alors $AA^{-1} = \mathrm{I}d$. Donc $(AA^{-1})^t = (\mathrm{I}d)^t = \mathrm{I}d$. Par le point précédent, $(AA^{-1})^t = (A^{-1})^t A^t$. Donc A^t est inversible et $(A^t)^{-1} = (A^{-1})^t$.

Définition 2.3
Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$, on dit que A est ***symétrique*** *si $A^t = A$ et* ***antisymétrique*** *si $A^t = -A$.*

Définition 2.4
Soit $P \in \mathcal{M}_n(\mathbb{R})$. On dit que P est une matrice de permutation si P a exactement un coefficient égal à 1 dans chaque ligne et chaque colonne, et que tous ses autres coefficients sont nuls. Une matrice de permutation a donc les mêmes lignes que la matrice identité mais dans un ordre qui peut être différent.

Quelques propriétés des matrices de permutation :

1. le produit de matrices de permutation est une matrice de permutation
2. toute matrice de permutation P est inversible et $P^{-1} = P^t$

2. Les matrices comme opérateurs linéaires

Désormais, et comme nous avons procédé au chapître précédent, les éléments $x = (x_1, \ldots, x_p)$ de $\mathbb{K}^p$ seront notés matriciellement en colonnes :

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}.$$

Le résultat du produit $Y = AX$ où A est une matrice à p colonnes et n lignes est une matrice–colonne à n lignes, donc est identifiée à un élément de $\mathbb{K}^n$.
Une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ définit donc une application :

$$\begin{aligned} A : \mathbb{K}^p &\to \mathbb{K}^n \\ X &\mapsto AX \end{aligned}.$$

Les propriétés de distributivité et d'associativité du produit matriciel, ont les conséquences suivantes :

$$\begin{gathered} \forall X, Y \in \mathbb{K}^p \,,\ \forall \lambda,\ \mu \in \mathbb{K} \\ A(\lambda X + \mu Y) = \lambda AX + \mu AY \ . \end{gathered}$$

On dit alors que A est *une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$*.

De plus le produit BA de deux matrices correspond à la composée des applications définies par les matrices A et B . Si A est une matrice inversible alors l'application qu'elle définit est une bijection dont A^{-1} définit la bijection réciproque. Nous verrons plus loin que la réciproque est vraie : si A définit une bijection alors A est inversible.

Réciproquement, il est aisé de constater que toute application $f : \mathbb{K}^p \to \mathbb{K}^n$ linéaire, c'est-à-dire vérifiant $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ ($\forall x, y \in \mathbb{K}^p$ et $\forall \lambda, \mu \in \mathbb{K}$), est un opérateur matriciel $\mathbb{K}^p \ni X \to AX$, en effet, en utilisant la notation matricielle des éléments de $\mathbb{K}^p$, on peut écrire :

$$f(X) = f(\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}) = \sum_i^p x_i f(E_i) \quad \text{, où } E_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{ème} \text{ ligne}$$

Les $f(E_i)$ sont des colonnes de $\mathbb{K}^n$ donc de la forme $f(E_i) = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix}$.

Ainsi :

$$f(X) = AX \quad \text{où } A = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ & \ddots & \\ a_{n1} & \dots & a_{np} \end{bmatrix} .$$

Exemples :

Dans $\mathbb{R}^2$, une rotation autour de l'origine d'angle θ , est caractérisée par la matrice :

$$A_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} .$$

On vérifie aisément que $A_\theta A_{\theta'} = A_{\theta+\theta'}$.

Dans $\mathbb{R}^2$, une homothétie de rapport λ , est caractérisée par la matrice :

$$H_\lambda = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} .$$

On obtient une matrice de similitude, en effectuant le produit $A_\theta H_\lambda$.

3. Systèmes linéaires

3.1 Écriture matricielle d'un système linéaire

Considérons par exemple le système d'équations

$$\left\{\begin{array}{l} 3x+2y-z=1 \\ 5x+3z=-8 \\ x-y+z=0 \end{array}\right. \tag{1}$$

Il s'écrit matriciellement :

$$\begin{bmatrix} 3 & 2 & -1 \\ 5 & 0 & 3 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -8 \\ 0 \end{bmatrix}$$

En toute généralité un système linéaire de n équations à p inconnues s'écrit

$$\left\{\begin{array}{l} a_{11}x_1+\ldots+a_{1p}x_p=b_1 \\ \vdots \\ a_{n1}x_1+\ldots+a_{np}x_p=b_n \end{array}\right. \Leftrightarrow \underbrace{\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ & & \ddots & \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}}_{=\boldsymbol{x}} = \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}}_{=\boldsymbol{b}}$$

Ce système est donc équivalent à l'écriture $A\boldsymbol{x}=\boldsymbol{b}$.

3.2 Elimination par les matrices

Remarque fondamentale :

On considère le système linéaire précédent. Soit P une matrice inversible telle que le produit PA soit bien défini. Dans ce cas :

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \text{ est solution de } A\boldsymbol{x}=\boldsymbol{b} \Leftrightarrow \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \text{ est solution de } PA\boldsymbol{x}=P\boldsymbol{b}.$$

En effet, si $A\boldsymbol{x}=\boldsymbol{b}$ alors $PA\boldsymbol{x}=P\boldsymbol{b}$. Réciproquement, si $PA\boldsymbol{x}=P\boldsymbol{b}$ alors $P^{-1}PA\boldsymbol{x}=P^{-1}P\boldsymbol{b}$ qui s'écrit donc $A\boldsymbol{x}=\boldsymbol{b}$.

C'est sur cette remarque que se fondent les méthodes d'élimination par le pivot de Gauss.

Cette méthode est basée sur la manipulation des lignes d'une matrice. Les manipulations autorisées sont :

- permuter deux lignes, codage : $\ell_i \leftrightarrow \ell_j$.
- ajouter à une ligne une combinaison linéaire des autres lignes. En pratique on se contente d'ajouter à une ligne un multiple d'une autre. Codage : $\ell_i \leftarrow \ell_i + \alpha\ell_j \ (j \neq i)$.

- multiplier une ligne par un scalaire non nul, codage : $\ell_i \leftarrow \lambda \ell_i$.

Chacune de ces manipulations correspond à la multiplication par une matrice inversible, comme on le verra d'abord dans l'exemple suivant.

3.3 Echelonnement d'une matrice 3×3. Gauss et opérations matricielles

Commençons par un exemple.

On considère le système $A\boldsymbol{x} = \boldsymbol{b}$, avec $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ et :

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ -1 & 1 & -2 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}.$$

On écrit la **matrice augmentée**, constituée de la matrice A et du second membre $\boldsymbol{b}$.

$$\tilde{A} = \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & 2 & -1 & 1 \\ -1 & 1 & -2 & -2 \end{bmatrix}.$$

La première ligne a un 1 en première position (en gras dans la matrice), on dit que c'est un **pivot**. On va pouvoir diviser toute la première ligne par ce nombre pour en soustraire un multiple à toutes les lignes suivantes, dans le but de faire apparatre des 0 dans tout le bas de la première colonne. La deuxième équation a déjà un 0 dessous, donc on n'a rien besoin de faire. On veut ensuite annuler le premier coefficient de la troisième ligne. On retranche donc (-1) fois la première ligne à la troisième[4] :

$$\begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & 2 & -1 & 1 \\ -1 & 1 & -2 & -2 \end{bmatrix} \overset{\ell_3 \to \ell_3 + \ell_1}{\longrightarrow} \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{2} & -1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

Ceci revient à multiplier $\tilde{A}$ à gauche par la matrice $T_{31}(1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

La deuxième ligne a un terme non nul en deuxième position (2) : c'est un pivot. On va maintenant annuler le deuxième terme de la troisième ligne ; pour cela, on retranche 1/2 fois la ligne 2 à la ligne 3 :

[4] Bien sûr, ceci revient à ajouter la première ligne ! il est cependant préférable de parler systématiquement de retrancher car c'est ce qu'on fait conceptuellement : pour l'élimination on enléve un multiple de la ligne du pivot à la ligne courante.

$$\begin{bmatrix} \mathbf{1} & 1 & 1 & 2 \\ 0 & \mathbf{2} & -1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \xrightarrow{\ell_3 \to \ell_3 - 1/2\ell_2} \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{2} & -1 & 1 \\ 0 & 0 & -\mathbf{\frac{1}{2}} & -\frac{1}{2} \end{bmatrix}$$

Ceci revient à multiplier la matrice précédente à gauche par la matrice

$$T_{32}(-\frac{1}{2}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}.$$

On a ici obtenu une matrice sous forme triangulaire supérieure à trois pivots : on peut donc faire la remontée pour obtenir la solution du système, et on obtient : $x_3 = 1$ puis $x_2 = 1$ et enfin $x_1 = 1$.

On a ainsi résolu le système linéaire.

Le fait de travailler sur la matrice augmentée est extrêmement pratique car il permet de travailler simultanément sur les coefficients du système linéaire et sur le second membre.

Finalement, au moyen des opérations décrites ci-dessus, on a transformé le système linéaire

$$A\boldsymbol{x} = \boldsymbol{b} \text{ en } U\boldsymbol{x} = \boldsymbol{c}, \text{ où } \boldsymbol{c} = T_{32}(-\frac{1}{2})T_{31}(1)\boldsymbol{b}\,, \quad U = T_{32}(-\frac{1}{2})T_{31}(1)A$$

est une matrice triangulaire supérieure.

3.4 Matrices élémentaires de $\mathcal{M}_n(\mathbb{K})$

Définition 2.5

Soit $i \in \{1, \dots, n\}$, on définit pour $a \in \mathbb{K}, a \neq 0$ la matrice $D_i(a) \in \mathcal{M}_n(\mathbb{K})$ qui est diagonale et dont tous les coefficients diagonaux sont égaux à 1 excepté le $i^{\text{ème}}$ qui est égal à a :

$$D_i(a) = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & a & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ - i^{\text{ème}} \textit{ ligne} \\ \\ \\ \\ \end{matrix}$$

On dit que la matrice $D_i(a)$ est une matrice de dilatation.

Définition 2.6
Soient i, j deux entiers distincts de $\{1, \dots, n\}$ ($i \neq j$), on définit la matrice $E_{i,j}$ comme la matrice dont tous les coefficients sont nuls sauf le coefficient i, j qui est egal à 1 .

Définition 2.7
Soient i, j deux entiers distincts de $\{1, \dots, n\}$ ($i \neq j$), pour tout $\lambda \in \mathbb{K}$ la matrice $T_{i,j}(\lambda) = \mathrm{Id}_n + \lambda E_{i,j}$:

$$T_{i,j}(\lambda) = \begin{array}{cl} \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \lambda & 0 & 1 & 0 & & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} & \begin{matrix} \\ \\ \\ - i^{\text{ème}} \textit{ ligne} \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} | \\ j^{\text{ème}} \textit{ colonne} \end{matrix} & \end{array}$$

On dit que la matrice $T_{i,j}(\lambda)$ est une matrice de transvection.

Les matrices de dilatation et de transvection sont appelées **matrices élémentaires**.

Proposition 2.8 *On a*

1. $(D_i(a))^t = D_i(a)$, $(T_{i,j}(\lambda))^t = T_{j,i}(\lambda)$.
2. *Si* $a \neq 0$, $(D_i(a))^{-1} = D_i\left(\frac{1}{a}\right)$, $(T_{i,j}(\lambda))^{-1} = T_{i,j}(-\lambda)$.

Démonstration. Le premier point est immédiat. Montrons le deuxième point. Pour les dilatations, on remarque que le produit de deux matrices diagonales D_1 et D_2 est une matrice diagonale dont les coefficients diagonaux sont les produits des coefficients diagonaux de D_1 et D_2. Effectuons le produit des matrices $T_{i,j}(\lambda)T_{i,j}(-\lambda)$. On a

$$T_{i,j}(\lambda)T_{i,j}(-\lambda) = (\mathrm{Id}_n + \lambda E_{i,j})(\mathrm{Id}_n - \lambda E_{i,j}) = \mathrm{Id}_n - \lambda^2 E_{i,j}E_{i,j} = \mathrm{Id}_n$$

car lorsque $i \neq j$, $E_{i,j}E_{i,j} = 0$. □

Lemme 2.9 *Si $E \in \mathcal{M}_n(\mathbb{K})$ est le produit de matrices élémentaires, alors E est inversible et son inverse est encore un produit de matrices élémentaires.*

Démonstration. On montre le résultat par récurrence sur le nombre de facteurs du produit, en utilisant la relation sur l'inverse d'un produit de matrices inversibles, et le fait que l'inverse d'une matrice élémentaire est une matrice élémentaire. □

Théorème 2.10

[Opérations élémentaires sur les lignes d'une matrice]

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

1. *La matrice* $D_i(a)A$ *est la matrice obtenue à partir de* A *en multipliant la* $i^{ème}$ *ligne de* A *par* a .
2. *La matrice* $T_{i,j}(\lambda)A$ *est la matrice obtenue à partir de* A *en ajoutant à la* $i^{ème}$ *ligne de* A, λ *fois la* $j^{ème}$ *ligne.*

3.5 Matrices échelonnées et pivot de Gauss

Définition 2.11

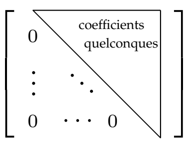
[Matrice échelonnée] *Soit* A *une matrice de* $\mathcal{M}_{n,p}(\mathbb{K})$. *On dit que la matrice* A *est* **échelonnée** *si elle vérifie les deux conditions suivantes :*

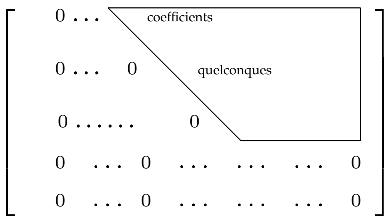
1. *Si une ligne est nulle, toutes les suivantes le sont.*
2. *Si la ligne* i *a son premier coefficient non nul sur la colonne* j , *alors le premier coefficient non nul de la ligne* $i+1$ *se trouve sur une colonne* $k > j$.

Autrement dit, une matrice $n \times p$ est sous forme échelonnée si les deux conditions suivantes sont vérifiées :

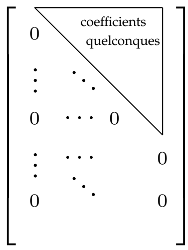
1. s'il existe des lignes de zéros, alors elles sont toutes en bas de la matrice ;
2. si deux lignes successives sont non nulles, alors la seconde a plus de zéros à gauche que la première.

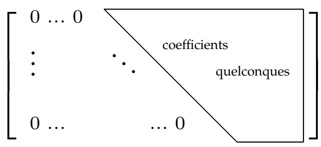
Voici à quoi ressemblent des matrices échelonnées :





ou encore :





Exemples de matrices échelonnées

Dans toutes les matrices ci dessous $\star$ désigne n'importe quel scalaire.

- Dans $\mathcal{M}_2(\mathbb{K})$

$$\begin{bmatrix} 2 & \star \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & \star \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

- Dans $\mathcal{M}_3(\mathbb{K})$

$$\begin{bmatrix} 1 & \star & \star \\ 0 & 1 & \star \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 5 & \star & \star \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 & \star \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 & \star \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Dans $\mathcal{M}_4(\mathbb{K})$

$$\begin{bmatrix} 1 & \star & \star & \star \\ 0 & 3 & \star & \star \\ 0 & 0 & 4 & \star \\ 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 8 & \star & \star \\ 0 & 0 & 1 & \star \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 & \star & \star \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 2 & \star \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- En revanche, la matrice $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$ n'est pas échelonnée.

Dans le cas des matrices carrées, une matrice échelonnée est triangulaire supérieure :

Proposition 2.12 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ échelonnée est triangulaire supérieure.*

Démonstration. Soit donc $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée. On montre par récurrence sur $2 \leq i \leq n$ que $a_{i,j} = 0$ si $j < i$.

Pour $i = 2$, si la ligne 2 est nulle, la propriété est trivialement vraie. Si la ligne 2 n'est pas nulle, le point 2. de la définition des matrices échelonnées montre que sur la ligne 2 le premier coefficient non nul se trouve sur une colonne $k > 1$, autrement dit $a_{21} = 0$. La propriété est donc vraie pour $i = 2$.

Supposons la propriété vraie pour $2 \leq i \leq n-1$ et montrons-la pour $i+1$. Si la ligne $i+1$ est nulle, la propriété est trivialement vraie. Si la ligne $i+1$ n'est pas nulle, la ligne i n'est pas nulle (par le point 1. de la définition des matrices échelonnées). Alors, comme $a_{i,j} = 0$ si $j < i$ par hypothèse de récurrence, le premier terme non nul de la ligne i se trouve sur une colonne $j_0 \geq i$. D'après le point 2., le premier terme non nul de la ligne $i+1$ se trouve donc sur une colonne $j \geq j_0 + 1 \geq i + 1$. Ainsi, $a_{i+1\,j} = 0$ pour tout $j < i + 1$.

La propriété est donc vraie pour tout i : la matrice A est donc triangulaire supérieure.
□

3.6 Existence de la forme échelonnée, algorithme d'échelonnement

Théorème 2.13 Echelonnement d'une matrice *Soit* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *. Il existe une matrice* $E \in \mathcal{M}_n(\mathbb{K})$ *produit de matrices élémentaires telle que la matrice* EA *est échelonnée.*

Pour démontrer le théorème, on décrit un algorithme qui donne explicitement la matrice E . Nous ne le ferons pas dans le le cas général, nous allons décrire l'algorithme sur un exemple.

On veut échelonner la matrice

$$A := \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 1 & 0 & -8 & 0 & 0 & 1 \\ 1 & 1 & -5 & 0 & 0 & 1 \end{bmatrix}.$$

L'idée est de procéder colonne par colonne. A l'étape i de l'algorithme, la matrice formée des i premières colonnes est échelonnée. Lorsqu'on arrive à l'étape 6 ($=$ le nombre de colonnes de A), la matrice obtenue est échelonnée.

On commence par l'étape $1.$ Il s'agit d'échelonner le premier vecteur colonne :

$$C_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Comme il est non nul, cela signifie qu'on veut, par des manipulations de lignes (c'est-à-dire en multipliant A à gauche par des matrices élémentaires), se ramener au vecteur :

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(Si le premier vecteur colonne était nul, en particulier il serait échelonné, on ne le modifierait pas et on passerait au vecteur colonne suivant). La première ligne de C_1 est un 1 donc on ne la change pas. Ensuite on fait apparaître des 0 sur les lignes suivantes de C_1. On commence par retrancher à la troisième ligne 1 fois la première. Cela correspond à effectuer $T_{31}(-1)A$. On obtient la matrice $A_1 = T_{31}(-1)A,$ c'est-à-dire :

$$A_1 = \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 0 & -2 & -6 & -4 & -1 & 0 \\ 1 & 1 & -5 & 0 & 0 & 1 \end{bmatrix}.$$

Pour obtenir A_1 à partir de A, on a effectivement ajouté à la troisième ligne -1 fois la première. On a obtenu ce qu'on voulait : un 0 sur la troisième ligne de la première colonne. Pour obtenir un zéro sur la quatrième ligne de la première colonne, on ajoute à la quatrième ligne -1 fois la première, c'est-à-dire on multiplie la matrice A_1 par $T_{41}(-1)$. On obtient la matrice $A_2 = T_{41}(-1)A_1 = T_{41}(-1)T_{31}(-1)A$:

$$A_2 = \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 0 & -2 & -6 & -4 & -1 & 0 \\ 0 & -1 & -3 & -4 & -1 & 0 \end{bmatrix}.$$

La première colonne de A_2 est de la forme

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

et donc on en a fini avec la première colonne, qui est échelonnée. De plus, on note qu'elle a 3 lignes nulles.

La deuxième étape consiste à échelonner la matrice formée des deux premières colonnes de A_2. Pour cela, on va échelonner le deuxième vecteur colonne de A_2 *à partir de la deuxième ligne* (noter que la deuxième ligne est la première ligne nulle de la première colonne). On verra que ça ne modifie pas la première colonne. La deuxième colonne de A_2 est

$$\begin{bmatrix} 2 \\ 1 \\ -2 \\ -1 \end{bmatrix}.$$

Echelonner ce vecteur colonne à partir de la deuxième ligne c'est se ramener par des combinaisons linéaires de lignes à une deuxième colonne du type

$$\begin{bmatrix} * \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

où $*$ représente un nombre quelconque. Cette opération revient à multiplier A_2 à gauche par des matrices élémentaires.

Il y a déjà un 1 sur la deuxième ligne. Pour faire apparaître un 0 sur la troisième ligne, on retranche à la troisième ligne -2 fois la seconde. On obtient

$$A_3 = T_{32}(2)A_2 = T_{32}(2)T_{41}(-1)T_{31}(-1)A = \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 0 & 0 & 0 & -12 & 3 & 2 \\ 0 & -1 & -3 & -4 & -1 & 0 \end{bmatrix}.$$

Noter qu'on n'a pas modifié la première colonne. Ceci est dû au fait qu'il y a un 0 sur la deuxième ligne de la première colonne.

On poursuit le travail sur la deuxième colonne, en gagnant un 0 sur la quatrième ligne. Pour cela, on retranche à la quatrième ligne -1 fois la deuxième, autrement dit, on multiplie A_3 par $T_{42}(1)$ pour obtenir

$$A_4 := T_{42}(1)T_{32}(2)T_{41}(-1)T_{31}(-1)A = \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 0 & 0 & 0 & -12 & 3 & 2 \\ 0 & 0 & 0 & -8 & 1 & 1 \end{bmatrix}.$$

La deuxième colonne a bien la forme qu'on attendait. La première colonne n'a pas été modifiée. La matrice formée des deux premières colonnes de A_4 est échelonnée, et on note qu'elle a deux lignes nulles.

Passons à la troisième étape : on travaille sur la troisième colonne. Il s'agit de l'échelonner *à partir de la troisième ligne* (qui est la première ligne nulle de la matrice formée des deux premières colonnes de A_4). La matrice formée des trois premières colonnes de A_4 sera ainsi échelonnée. Comme le vecteur colonne constitué des 2 dernières lignes de la troisième colonne est nul, on n'a rien à faire sur la troisième colonne : elle est déjà échelonnée à partir de la troisième ligne. On remarque que la matrice formée des trois premières colonnes de A_4 a encore deux lignes nulles.

On passe à la quatrième colonne, qu'on veut échelonner *à partir de la troisième ligne* car c'est la première ligne nulle de la matrice formée des trois premières colonnes de A_4. Ainsi, on veut ramener le quatrième vecteur colonne à un vecteur colonne de la forme :

$$\begin{bmatrix} * \\ * \\ a \\ 0 \end{bmatrix}.$$

avec a non nul. On voit que pour cela, il suffit de retrancher à la quatrième ligne $2/3$ fois la troisième, autrement dit de multiplier A_4 par $T_{43}(-2/3)$. On obtient

$$A_5 := T_{43}(2)T_{42}(1)T_{32}(2)T_{41}(-1)T_{31}(-1)A = \begin{bmatrix} 1 & 2 & -2 & 4 & 1 & 1 \\ 0 & 1 & 3 & -4 & 2 & 1 \\ 0 & 0 & 0 & -12 & 3 & 2 \\ 0 & 0 & 0 & 0 & -1 & -1/3 \end{bmatrix}.$$

Ceci achève le travail sur la quatrième colonne : la matrice formée des 4 premières colonnes est échelonnée. De plus, elle a 1 ligne nulle. Pour la cinquième étape : le cinquième vecteur colonne est échelonné à partir de la quatrième ligne. Donc on n'y touche pas. La matrice formée des 5 premières colonnes est échelonnée et n'a pas de ligne nulle. L'algorithme s'arrête donc ici : la matrice A_5 est échelonnée.

De plus, on a $A_5 = EA$ avec

$$E := T_{43}(-2/3)T_{42}(1)T_{32}(2)T_{41}(-1)T_{31}(-1).$$

Remarque :

Lorsque l'on rencontre un zéro en position pivotale (et non pas un « pivot nul », car, par définition, un pivot n'est jamais nul) durant l'échelonnement de la matrice on peut, au lieu de transformer la ligne en question, permuter la ligne où se trouve ce zéro avec une des lignes qui se trouve en dessous de celle où se trouve ce zéro.

Dans la suite, il sera utile de repérer quelles sont les colonnes pivotales d'une matrice échelonnée, dont on a déjà parlé lors de l'élimination de Gauss. En voici une définition pour une matrice échelonnée rectangulaire $n \times p$.

Définition 2.14

Pivots, colonnes pivotales, rang *On note r le nombre de lignes non nulles d'une matrice $n \times p$ échelonnée. On appelle **pivot** le premier terme non nul de chaque ligne non nulle de la matrice échelonnée. On appelle colonnes pivotales d'une matrice échelonnée les colonnes dans lesquelles aparaissent les pivots des lignes non nulles (attention ce ne sont pas forcément les r premières colonnes de la matrice). On note $k_1, \dots, k_r$ les indices de ces colonnes. On appelle colonnes non pivotales les autres colonnes.*

Les colonnes pivotales sont donc les colonnes $c_j(A)$ telles que $j \in J$, où $J = \{k_1, \dots, k_r\}$ est l'ensemble des indices des colonnes dans lesquelles apparaissent les pivots.

On définit le rang de la matrice comme le nombre de lignes non nulles (ou de colonnes pivotales).

La preuve du théorème d'échelonnement se fait en propageant l'échelonnement à partir de celui d'une colonne. Nous détaillons l'algorithme pour une colonne dans le paragraphe qui suit.

3.7 Algorithme d'échelonnement d'un vecteur colonne

Lemme 2.15 *Soit $A = (a_i)_{i=1,\dots,n} \in \mathcal{M}_{n,1}(\mathbb{K})$ un vecteur non nul. Alors il existe une matrice $E \in \mathcal{M}_n(\mathbb{K})$ produit de matrices élémentaires et $a \in \mathbb{K}$, $a \neq 0$, tels que*

$$EA = \begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- Si $a_1 \neq 0$ alors on utilise a_1 pour éliminer les coefficients qui sont en-dessous, à l'aide des matrices élémentaires. On commence par retrancher à la deuxième ligne a_2/a_1 fois la première :

$$T_{2,1}\left(-a_2/a_1\right) A = \begin{bmatrix} a_1 \\ 0 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}.$$

  On procède de même avec les lignes suivantes :

$$T_{n,1}\left(-a_n/a_1\right) \cdots T_{2,1}\left(-a_2/a_1\right) A = \begin{bmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- Si $a_1 = 0$, il existe $i \geq 2$ tel que $a_i \neq 0$ car on a supposé $A \neq 0$. Auquel cas, si on ajoute à la première ligne 1 fois la $i^{ième}$ ligne, on obtient

$$T_{1,i}\left(1\right) A = \begin{bmatrix} a_i \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

  On est alors ramené au cas précédent.

$$T_{n,1}(-a_n/a_i) \cdots T_{2,1}(-a_2/a_i) T_{1,i}\left(1\right) A = \begin{bmatrix} a_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

3.8 Calcul de l'inverse d'une matrice, échelonnement total

Nous allons prolonger l'algorithme de Gauss, ce sera *l'algorithme de Gauss-Jordan des matrices*. Nous l'illustrons d'abord par un exemple d'inversion d'une matrice.

Etant donnée $A \in \mathcal{M}_n(\mathbb{K})$. Il s'agira encore d'effectuer un produit P de matrice élémentaires, donc des manipulations de pivot, en sorte que $PA = \mathrm{Id}_n$. On aura donc $A^{-1} = P$ en cas de succès.

Au lieu d'introduire une matrice augmentée avec la matrice de départ et un vecteur second membre, on introduit une matrice augmentée avec la matrice de départ et la matrice unité ; ceci revient à considérer la matrice augmentée $n \times 2n$ constituée de la matrice A et la matrice $\mathrm{I}d_n$:

$$\tilde{A} = \begin{bmatrix} A & \mathrm{I}d_n \end{bmatrix}.$$

Si la matrice A est inversible, en appliquant l'algorithme de Gauss-Jordan décrit si dessous, on obtient alors l'inverse de la matrice dans la partie droite (où se trouvait auparavant la matrice identité) de la forme « totalement échelonné » de la matrice augmentée.

Si A est inversible, sa forme totalement échelonnée R est $\mathrm{I}d_n$, et l'algorithme de Gauss-Jordan sur $\tilde{A} = \begin{bmatrix} A & \mathrm{I}d_n \end{bmatrix}$ donne :

$$\begin{bmatrix} \mathrm{I}d_n & P \end{bmatrix} = \begin{bmatrix} \mathrm{I}d_n & A^{-1} \end{bmatrix}.$$

Voyons ceci d'abord sur un exemple.

3.9 Inversion d'une matrice 2 × 2

Prenons une matrice 2×2 : $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$. On écrit la matrice augmentée : $\tilde{A} = \begin{bmatrix} A & \mathrm{I}d_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{bmatrix}$, et on applique Gauss-Jordan. On commence par Gauss :

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{bmatrix} \xrightarrow{\ell_2 \to \ell_2 - 3\ell_1} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{bmatrix}$$

et on effectue ensuite la partie Jordan, qui consiste à rendre la matrice diagonale :

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{bmatrix} \xrightarrow{\ell_1 \to \ell_1 - 2\ell_2} \begin{bmatrix} 1 & 0 & 7 & -2 \\ 0 & 1 & -3 & 1 \end{bmatrix}$$

Gauss Jordan donne donc $A^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$. On vérifie qu'on a bien $AA^{-1} = \mathrm{I}d_2$

3.10 Inversion d'une matrice 3 × 3

Considérons la matrice $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix}$.

On écrit la matrice augmentée $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$.

On commence par appliquer l'algorithme de Gauss, c'est-à-dire se ramener à une matrice échelonnée :

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \overset{\ell_3 \to \ell_3 - \ell_1}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{bmatrix}$$

$$\overset{\ell_3 \to \ell_3 - (1/2)\ell_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1/2 & -1 & -1/2 & 1 \end{bmatrix}.$$

On a obtenu une matrice échelonnée. On reprend alors l'algorithme à la fin de Gauss, et on écrit maintenant la partie « Jordan »[5] de l'algorithme dit de Gauss-Jordan.

Pour cela, on fait apparaître des 0 au-dessus des pivots en commençant par les colonnes les plus à droite, et en progressant vers la gauche.

On multiplie la dernière ligne par -2 pour obtenir le pivot 1 :

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1/2 & -1 & -1/2 & 1 \end{bmatrix} \overset{\ell_3 \to -2\ell_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix}$$

Puis on met à zéro le troisième coefficient de la deuxième ligne (à l'aide de la troisième) :

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix} \overset{\ell_2 \to \ell_2 + \ell_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix}$$

On annule ensuite le troisième coefficient de la première ligne (toujours à l'aide de la troisième) :

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix} \overset{\ell_1 \to \ell_1 - \ell_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & 2 \\ 0 & 2 & 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix}$$

On multiplie enfin la deuxième ligne par $1/2$ pour obtenir le pivot 1 :

$$\begin{bmatrix} 1 & 0 & 0 & -1 & -1 & 2 \\ 0 & 2 & 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix} \overset{\ell_2 \to 1/2\ell_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 2 & 1 & -2 \end{bmatrix}$$

Les trois premières colonnes forment la matrice identité. Elle constitue la forme totalement échelonnée de la matrice A : c'est le cas pour toute matrice inversible. On verra plus loin

[5] Marie Ennemond Camille Jordan, né le 5 janvier 1838 à Lyon (Rhéne) et mort le 22 janvier 1922, est un mathématicien français, connu à la fois pour son travail fondamental dans la théorie des groupes et pour son cours d'analyse.

que la réciproque est également vraie : si la forme totalement échelonnée de la matrice A est l'identité, A est inversible.

Si au cours du calcul apparait une colonne de zéros, ou une demi-ligne de zéros dans la partie gauche, le processus s'arrête : la matrice n'est pa inversible

3.11 Echelonnement total

Définition 2.16
Matrice totalement échelonnée
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On dit que A est totalement échelonnée si :

1. *A est échelonnée,*
2. *le premier coefficient non nul de chaque ligne non nulle est égal à 1,*
3. *tous les éléments qui sont dans la colonne au dessus du premier coefficient non nul (et égal à 1) d'une ligne sont nuls.*

Notons que dans une forme totalement échelonnée, les pivots sont toujours égaux à 1, ce qu'on n'a pas demandé pour les matrices échelonnées.

Matrices totalement échelonnées ou non

- La matrice identité $\mathrm{I}d_n$ est totalement échelonnée et a n colonnes pivotales : $r = n$.
- La matrice nulle 0_n est totalement échelonnée et n'a aucune colonne pivotale : $r = 0$.
- La matrice suivante est totalement échelonnée :

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

  Elle a deux colonnes pivotales : la première, $k_1 = 1$, et la troisième : $k_2 = 3$, et une colonne non pivotale, la seconde : $k_3 = 2$.

- La matrice 4×6 :

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  est sous forme totalement échelonnée, et a trois colonnes pivotales $k_1 = 2$, $k_2 = 4, k_3 = 5$, et trois colonnes non pivotales : $k_4 = 1$, $k_5 = 3, k_6 = 6$.

- La matrice suivante est échelonnée, mais pas totalement échelonnée :

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

- La matrice suivante n'est ni totalement échelonnée ni même échelonnée :

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

On admettra le résultat suivant :

Théorème 2.17 Echelonnement total

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. *Il existe une matrice* E *produit de matrices élémentaires telle que la matrice* EA *est totalement échelonnée.*

Chapitre 3

Espaces Vectoriels

1. DŐfinitions et premiers exemples

Définition 3.1
On appelle structure de $\mathbb{K}$ *-espace vectoriel sur un ensemble* E *, la donnée :*
1– d'une loi de groupe commutatif notée additivement. L'élément neutre étant noté 0_E *.*
2– d'une loi externe, application de $\mathbb{K} \times E$ *dans* E *, notée* $(\lambda, x) \to \lambda x$ *et vérifiant, pour tous scalaires* λ, μ *et pour tous* $(x, y) \in E$ *:*

$$\begin{cases} (\lambda\mu)x = \lambda(\mu x) \\ \lambda(x+y) = \lambda x + \lambda y \\ (\lambda+\mu)x = \lambda x + \mu x \\ (1_{\mathbb{K}})x = x \end{cases}$$

Les éléments d'un espace vectoriel seront appelés *vecteurs*.

Exemples :

1. Pour tout entier n , $\mathbb{K}^n$ est un $\mathbb{K}$ -espace vectoriel pour les lois :

$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ \ldots, y_n) &:= (x_1 + y_1, \ldots, x_n + y_n) \\ \lambda(x_1, \ldots, x_n) &:= (\lambda x_1, \ldots, \lambda x_n) \end{aligned}$$

   En particulier pour $n = 1$. $\mathbb{K}$ est un $\mathbb{K}$ -espace vectoriel.

   Graphiquement, $\mathbb{R}^2$ possde une reprsentation trs simple : il s'agit du plan repr. De mme, l'espace repr est une bonne reprsentation de $\mathbb{R}^3$.

2. $E = \mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices $n \times p$ muni de la somme des matrices et du produit par un scalaire.

3. $\mathbb{R}^2$ est un $\mathbb{R}$-ev,
4. $\mathbb{C}^2$ est un $\mathbb{C}$-ev et un $\mathbb{R}$-ev,
5. $\mathcal{M}_2(\mathbb{K})$, $\mathcal{M}_{2,1}(\mathbb{K})$, $\mathcal{M}_{1,2}(\mathbb{K})$, $\mathcal{M}_{n,p}(\mathbb{K})$ munis de l'addition des matrices et de la multiplication par un scalaire (de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) sont des $\mathbb{K}$-ev,
6. L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$,
7. L'ensemble des fonctions continues sur un intervalle $[a, b]$ est un espace vectoriel sur $\mathbb{R}$,
8. L'ensemble des fonctions $u : \mathbb{R} \to \mathbb{R}$ solutions de l'équation $u'' + u' + u = 0$.
9. L'ensemble des suites relles est un $\mathbb{R}$-espace vectoriel.

Ensembles qui ne sont pas des espaces vectoriels :

- $\mathbb{Z}$ n'est pas un $\mathbb{R}$-ev,
- $(\mathbb{R}_+)^2$ n'est pas un $\mathbb{R}$-ev,
- L'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ qui valent 1 en 0,
- L'ensemble des fonctions $u : \mathbb{R} \to \mathbb{R}$ solutions de l'équation $u'' + u' + u = 1$.

1.1 Règles de calcul dans un espace vectoriel

Dans la suite de ce chapitre, et sauf indication contraire, la lettre E désignera un $\mathbb{K}$-espace vectoriel.

- Pour tout $x \in E$ on a : $(0_{\mathbb{K}})x = 0_E$.

  Il n'y a en général pas, de ce fait, d'équivoque à noter 0 aussi bien l'élément neutre de l'addition dans $\mathbb{K}$ que celui de l'addition dans E.
- $(-1)x = -x$ ($-x$ désigne le symétrique de x).

  En effet : $0_E = 0x = (1-1)x = x + (-1)x$.
- On pourra donc écrire sans ambiguïté : $(\lambda - \mu)x = \lambda x - \mu x$

1.2 Combinaisons linéaires

Soit une partie finie $A = \{a_1, \ldots, a_n\} \subset E$

Définition 3.2
On appelle combinaison linéaire de A, tout vecteur x de la forme

$$x = \sum_{i=1}^{n} \lambda_i a_i = \lambda_1 a_1 + \ldots + \lambda_n a_n, \text{ où } \forall i = 1 \ldots n, \ \ \lambda_i \in \mathbb{K}$$

Exemples :

- Dans $\mathbb{K}^3$ tout vecteur est combinaison linéaire de $\{(1,0,0);(0,1,0);(0,0,1)\}$.
- Dans $\mathbb{K}[X]$, tout polynôme de degré n est combinaison linéaire de $\{1, X, \ldots, X^n\}$.

Il nous arrivera d'utiliser l'abrviation CL pour « combinaison linaire ».

Remarques :

- 0_E est combinaison linéaire de toute partie finie.
- Toute combinaison linéaire de combinaisons linéaires d'une famille est encore une combinaison linéaire de cette famille.

Généralisation : Si $A \subset E$ est une partie quelconque, non nécessairement finie, on appellera combinaison linéaire de vecteurs de A toute combinaison linéaire de toute partie finie de A.

Exemple : dans $\mathbb{K}[X]$ tout polynôme est combinaison linéaire de $\{1, X, \ldots, X^n, \ldots\}$.

2. Sous-Espaces Vectoriels

Définition 3.3
On appelle sous-espace vectoriel de E, en abrg sev, toute partie F de E telle que :
– $F \neq \emptyset$
– $\forall x \in F, \forall y \in F, \forall (\lambda, \mu) \in \mathbb{K}^2 \quad \lambda x + \mu y \in F$.

Cela signifie que F est un sous-espace vectoriel si et seulement si F et stable pour l'addition de E et pour le produit d'un élément de F par un scalaire.

En particulier pour $x \in F$ et $0 \in \mathbb{K}$ on a $0x = 0_E \in F$, donc un sous-espace vectoriel **possde toujours le vecteur** 0_E.

Remarque : Un sous-espace vectoriel est lui-mme un espace vectoriel. Il suffira, pour montrer qu'un ensemble est muni d'une structure d'espace vectoriel, dans la grande majorit des cas, de vrifier qu'il est sous-espace vectoriel d'une strucure dj identifie comme espace vectoriel (cf. **2.2** ci-dessous).

2.1 Exemples

- E et $\{0_E\}$ sont des sev de E.
- $\mathbb{K}_n[X]$ l'ensemble des polynômes de degr infrieur ou gal n forment un sous-espace vectoriel de $\mathbb{K}[X]$.

La plupart des espaces vectoriels que l'on rencontre sont des sous-espaces d'espaces plus gnraux. Cela est le cas pour les espaces de fonctions : les fonctions polynmes relles sur un intervalle I forment un sous-espace des fonctions drivables sur I, qui elles mmes sont un sous-espace des fonctions continues, elles mmes sous-espace de $\mathbb{R}^I$.

Proposition 3.4 *Une partie F non vide de E est un sous-espace vectoriel si et seulement si F est stable par combinaison linéaire.*

Démonstration. Par récurrence évidente sur n ; pour $n = 1, 2$ c'est vrai par définition. En supposant la propriété vraie jusqu'à $n - 1$: alors $\sum_{i \leq n-1} \lambda_i a_i \in F$ et $a_n \in F$ donc, par stabilité de F pour l'addition, on a $\displaystyle\sum_{i \leq n-1} \lambda_i a_i + \lambda_n a_n = \sum_{i \leq n} \lambda_i a_i \in F$. □

Théorème 3.5 *Une intersection de deux sous-espaces vectoriels de E est un sous-espace vectoriel de E.*

Démonstration. Soit F et G deux sous-espaces vectoriels de E. $F \cap G$ n'est pas vide puisque 0_E appartient F et G. Par ailleurs pour tous x et $y \in F \cap G$ et pour tous λ et μ scalaires, on a $\lambda x + \mu y \in F$ et $\lambda x + \mu y \in G$ par stabilit de F et de G, ainsi $\lambda x + \mu y \in F \cap G$. □

Il en rsulte facilement qu'une intersection finie de sous-espaces vectoriels de E est un sous-espce vectoriel de E (c'est même le cas pour une intersection infinie).

2.2 Les ensembles d'applications Ĺ valeurs dans un espace vectoriel

La structure d'espace vectoriel se rencontre souvent dans le cas d'espaces d'applications. La plupart de ces espaces sont des sev d'un espace « englobant ». Prcisons cela.

Considrons un $\mathbb{K}$-espace vectoriel E. Soit X un ensemble quelconque et notons E^X l'ensemble de *toutes les applications* de X dans E. Ds lors, E^X hrite naturellement d'une structure d'espace vectoriel, comme suit : pour tous f et g dans E^X et pour tout $\lambda \in \mathbb{K}$, on posera $(\lambda f + g)(x) = \lambda f(x) + g(x)$. L'addition dans le membre de droite opre dans E, celle du membre de gauche sur des lments de E^X. La vrification que cela dfinit une structure de $\mathbb{K}$-espace vectoriel est facile et laisse en exercice (0_{E^X} tant l'application identiquement nulle).

Il faut retenir que c'est la structure vectorielle existant au but E des applications, et non pas leur source X, qui cre la structure de E^X. Il ne sera donc pas utile de vrifier la longue liste des proprits d'espace vectoriel chaque fois que l'on rencontrera ce cas de figure.

Ainsi toutes les fonctions dfinies sur $[a,b]$ valeurs relles forment un espace vectoriel. Les autres sous-ensembles couramment rencontrs, fonctions continues, drivables ou C^∞ sur $[a,b]$ etc. seront des espaces vectoriels car aisment reprs comme sous-espaces vectoriel de $\mathbb{R}^{[a,b]}$. De mme l'ensemble des suites relles $\mathbb{R}^\mathbb{N}$ est un espace vectoriel.

2.3 Sous-espaces engendrés

Étant donnée une famille de vecteurs $A = \{a_1, \dots, a_n\}$, l'ensemble des combinaisons linéaires de A est noté $\mathrm{Vect}(a_1, \dots, a_n)$.

> **Théorème 3.6**
> *1)* $\mathrm{Vect}(a_1, \dots, a_n)$ *est un sous-espace vectoriel.*
> *2) Pour tout sev* $F \subset E$ *;* $A \subset F \Rightarrow \mathrm{Vect}(a_1, \dots, a_n) \subset F$. *Ainsi* $\mathrm{Vect}(a_1, \dots, a_n)$ *est le plus petit (au sens de l'inclusion) sev contenant* A
> $\mathrm{Vect}(a_1, \dots, a_n)$ *est appelé sous-espace engendré par* A.

Démonstration. $\mathrm{Vect}(a_1, \dots, a_n)$ possde 0_E et n'est donc pas vide. Il est, par construction, stable par combinaison linaire, c'est bien un sous-espace. De plus

$$\mathrm{Vect}(a_1, \dots, a_n) \subset F$$

par stabilité ; c'est donc le plus petit sev contenant A. □

> **Définition 3.7**
> *Une famille* $(a_1, \dots, a_n)$ *est dite* génératrice *d'un sous-espace* F *si* $F = \mathrm{Vect}(a_1, \dots, a_n)$.

Il est donc équivalent de dire que tout vecteur de F est combinaison linéaire de $\{a_1, \dots, a_n\}$.

Remarque : $\mathrm{Vect}(a_1, \dots, a_n) = \mathrm{Vect}(a_1, \dots, a_{n-1})$ si et seulement si a_n est combinaison linéaire de $\{a_1, \dots, a_{n-1}\}$.

> **Définition 3.8**
> *Soit* $\boldsymbol{a} \in E$ *et* $\boldsymbol{a} \neq 0_E$*, le sous-espace* $Vect(\boldsymbol{a})$ *est appel* **droite vectorielle engendre par** $\boldsymbol{a}$*, on peut galement le noter* $[\boldsymbol{a}]$.

2.4 Familles libres – Familles liées

> **Définition 3.9**
> *Une famille* $\{a_1, \dots, a_n\}$ *est dite* libre*, si*
>
> $$\sum_{i=1}^{n} \lambda_i a_i = 0_E \Leftrightarrow \lambda_i = 0 \ \ \forall i \in \{1, \dots, n\}$$

Autrement dit une famille est libre si la combinaison linéaire nulle ne peut s'obtenir que par des coefficients nuls.

Définition 3.10
Une famille est dite liée *si elle n'est pas libre.*

Proposition 3.11 *Une famille $\{a_1, \ldots, a_n\}$ est liée si et seulement si un des vecteurs a_i est combinaison linéaire des autres.*

Démonstration. supposons $\{a_1, \ldots, a_n\}$ lie. Il existe une CL

$$\sum_{i=1}^{n} \lambda_i a_i = 0_E$$

avec les coefficients λ_i non tous nuls. On ne perd pas de gnralit supposer que $\lambda_1 \neq 0$; on peut alors crire

$$a_1 = \sum_{i=2}^{n} -\frac{\lambda_i}{\lambda_1} a_i$$

donc a_1 est bien CL des autres vecteurs.

Rciproquement, supposons que l'un des vecteurs soit CL des autres, l encore, on ne perd pas de gnralit supposer qu'il s'agit de $a_1 = \alpha_2 a_2 + \ldots + \alpha_n a_n$, ce qui entrane

$$-a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = 0_E$$

la famille est donc lie, puisqu'elle admet une CL nulle avec les coefficients non tous nuls : -1 est le coefficient de a_1 . □

Remarques :

- Une famille comprenant 0_E est toujours liée.
- Une famille réduite à un vecteur non nul est toujours libre.
- Deux vecteurs sont liés si et seulement si ils sont colinéaires.
- Toute famille contenant une famille liée est liée.
- Toute sous-famille d'une famille libre est libre.

2.5 Bases – Dimension

Proposition 3.12 *Si $\{a_1, \ldots, a_p\}$ est libre ; alors $\forall x \notin \mathrm{Vect}(a_1, \ldots, a_p)$ la famille $\{a_1, \ldots, a_p, x\}$ est libre.*

Démonstration. en effet la condition $\lambda x + \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_p a_p = 0_E$ entrane que $\lambda = 0$ sinon on pourrait exprimer $x = -\frac{\lambda_1}{\lambda} a_1 - \ldots - \frac{\lambda_p}{\lambda} a_p$ ce qui est contraire l'hypothse. Il reste alors $\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_p a_p = 0_E$ d'où la nullité de tous les coefficients par l'hypothèse que la famille $\{a_1, \ldots, a_p\}$ est libre. □

Lemme 3.13 [fondamental] *Si $\{a_1, \ldots, a_n\}$ engendre un sev F alors toute famille d'au moins $n+1$ vecteurs de F est nécessairement liée.*

Démonstration. Nous allons raisonner par récurrence sur n. Vérifions le pour $n = 1$. Soit $\{a_1\}$ engendrant F et $\{b_1, b_2\} \subset F$. On a par hypothèse de génération : $b_1 = \lambda_1 a_1$ et $b_2 = \lambda_2 a_1$ si les deux coefficients sont nuls la famille est liée. Sinon, supposons par exemple que $\lambda_1 \neq 0$, alors $a_1 = \dfrac{1}{\lambda_1} b_1$ en reportant dans $b_2 = \lambda_2 a_1 = \dfrac{\lambda_2}{\lambda_1} b_1$, ainsi $\{b_1, b_2\}$ sont liés car colinéaires.

Supposons la propriété vraie pour des sous-espaces engendrés par $n-1$ éléments. Démontrons la pour n. Soit $\{a_1, \ldots, a_n\}$ engendrant F et $\{b_1, \ldots, b_{n+1}\} \subset F$. Notons F' le sous-espace engendré par $\{a_1, \ldots, a_{n-1}\}$. On peut écrire

$$\begin{aligned}
b_1 &= b'_1 + \lambda_1 a_n \\
b_2 &= b'_2 + \lambda_2 a_n \\
&\vdots \\
b_{n+1} &= b'_{n+1} + \lambda_{n+1} a_n
\end{aligned}$$

où les $b'_i \in F'$. Si tous les λ_i sont nuls alors la famille considérée est dans F' et la question est réglée par hypothèse de récurrence. Supposons l'un des λ_i non nul ; on peut toujours supposer, au prix d'une renumérotation éventuelle, que $\lambda_{n+1} \neq 0$ donc $a_n = \dfrac{1}{\lambda_{n+1}}(b_{n+1} - b'_{n+1})$ en reportant cette valeur de a_n dans les n premières équations ci-dessus on a

$$\begin{aligned}
b_1 \quad &= b'_1 + \frac{\lambda_1}{\lambda_{n+1}}(b_{n+1} - b'_{n+1}) \\
b_2 \quad &= b'_2 + \frac{\lambda_2}{\lambda_{n+1}}(b_{n+1} - b'_{n+1}) \\
&\vdots \\
b_n \quad &= b'_n + \frac{\lambda_n}{\lambda_{n+1}}(b_{n+1} - b'_{n+1})
\end{aligned}$$

ce qui s'écrit également

$$\begin{aligned}
b_1 - \frac{\lambda_1}{\lambda_{n+1}} b_{n+1} \quad &= b'_1 - \frac{\lambda_1}{\lambda_{n+1}} b'_{n+1} \\
b_2 - \frac{\lambda_2}{\lambda_{n+1}} b_{n+1} \quad &= b'_2 - \frac{\lambda_2}{\lambda_{n+1}} b'_{n+1} \\
&\vdots \\
b_n - \frac{\lambda_n}{\lambda_{n+1}} b_{n+1} \quad &= b'_n - \frac{\lambda_n}{\lambda_{n+1}} b'_{n+1}
\end{aligned}$$

les $b_i - \dfrac{\lambda_i}{\lambda_{n+1}} b_{n+1} = b'_i - \dfrac{\lambda_i}{\lambda_{n+1}} b'_{n+1} \in F'$ sont liés par hypothèse de récurrence ; il existe donc une relation

$$\alpha_1(b_1 - \frac{\lambda_1}{\lambda_{n+1}} b_{n+1}) + \ldots \alpha_n(b_n - \frac{\lambda_n}{\lambda_{n+1}} b_{n+1}) = 0$$

avec les α_i non tous nuls. Après regroupement, cette relation s'écrit

$$\alpha_1 b_1 + \ldots \alpha_n b_n - \frac{1}{\lambda_{n+1}} [\sum_{i=1}^{n} (\alpha_i \lambda_i)] b_{n+1} = 0$$

qui établit que $\{b_1, \ldots, b_{n+1}\}$ est liée, car les α_i sont non tous nuls. $\square$

Définition 3.14
On appellera base d'un espace vectoriel E, toute famille libre et génératrice de E. On parlera aussi bien de base d'un sous-espace.

Exemples :

- L'espace $\mathbb{K}^n$ possde une **base canonique**. Elle est forme de la famille des n vecteurs :

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \ldots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

- L'espace $\mathbb{K}_n[X]$ admet la famille $\{1, X, X^2, \ldots, X^n\}$ comme base naturelle.
- Dans l'espace des applications d'une variable relle valeurs relles, la famille $\{\cos(x), \sin(x), \cos(2x), \sin($ est libre (le montrer titre d'exercice), elle forme donc une base de l'espace qu'elle engendre.
- Plus gnralement, toute famille libre dans un espace vectoriel est une base du sous-espace qu'elle engendre.
- Une droite vectorielle $[\boldsymbol{a}]$ admet $\{\boldsymbol{a}\}$ comme base.

Si $\{a_1, \ldots, a_n\}$ est une base de E tout vecteur x s'écrit de manière unique $x = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n$. Les scalaires x_i sont appelés *les coordonnées* de x dans la base.

Proposition 3.15 *Si E possède une base finie de n vecteurs alors toute les autres bases de E ont le même nombre de vecteurs. Ce nombre n est alors appelé* **dimension** *de E qui est alors dit de dimension finie. On écrira $dimE = n$.*

Démonstration.

Il existe, par hypothèse, une base de n vecteurs $A = \{e_1, \ldots, e_n\}$. Soit $B = \{e'_1, \ldots, e'_m\}$ une autre base de E, d'après le lemme fondamental on ne peut avoir $n < m$ (sinon B serait liée) ; de même on ne peut avoir $m < n$, donc $m = n$. □

Théorème 3.16 *Soit E tel que $\dim E = n$ alors*
1) Toute famille génératrice a au moins n éléments.
2) Toute famille libre a au plus n éléments.
3) Pour une famille $A = \{a_1, \ldots, a_n\}$ de n vecteurs on a les équivalences

$$A \text{ base } \Leftrightarrow A \text{ libre } \Leftrightarrow A \text{ génératrice}$$

Démonstration.

1) Si $\{a_1, \ldots, a_m\}$ est une famille génératrice de E, alors $n \leq m$. Sinon, d'après le lemme fondamental, la base serait liée, ce qui ne se peut pas.

2) Si $\{a_1, \ldots, a_m\}$ est libre alors $m \leq n$. Car elle serait nécessairement liée sinon.

3) Si A est une base alors A est libre et génératrice. Supposons $A = \{a_1, \ldots, a_n\}$ libre, alors $\forall x$, $\{a_1, \ldots, a_n, x\}$ est nécessairement liée d'après 2) ; donc $x \in \mathrm{Vect}(A)$. Donc A est génératrice donc une base.

Soit $A = \{a_1, \ldots, a_n\}$ génératrice, montrons qu'elle est libre. Si elle ne l'tait pas, l'un des vecteurs serait CL des autres, en l'enlevant on produirait une famille génératrice de $n - 1$ vecteurs (puisqu'une CL d'une CL est une CL) ce qui ne se peut pas. □

Conséquence de 1) : si $F \subset F'$, sont deux sev alors $\dim F \leq \dim F'$. En effet : la dimension d'un sous-espace est la taille maximale d'une famille libre, or toute base de F est libre dans F'. Donc tout sous-espace d'un espace de dimension finie est lui-mme de dimension finie.

On en dduit que $\mathbb{K}^n$ est de dimension n, alors que $\mathbb{K}_n[X]$ est de dimension $n + 1$.

Pour dterminer si une famille de trois vecteurs $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ est gnratrice de $\mathbb{R}^3$, donc est une base puisque $\mathbb{R}^3$ est de dimension trois, il suffira de vrifier qu'elle est libre, ce qui est plus facile en gnral.

Théorème 3.17 *Toute famille génératrice finie de E contient une base (dans ce cas E sera de dimension finie).*

Démonstration. On suppose E non nul. Par récurrence sur la longueur de la famille. Si $\{e_1\} \neq \{0\}$ engendre E alors $\{e_1\}$ est libre donc constitue une base. Supposons, par hypothèse de récurrence, que de toute famille génératrice de longueur $n - 1$ on peut extraire une base. Soit $\{a_1, \ldots, a_n\}$ une famille génératrice de E, si elle est libre c'est elle même une base, sinon un des vecteurs est combinaison linéaire des autres ; on peut toujours supposer qu'il s'agit de a_n, dès lors E est engendré par $\{a_1, \ldots, a_{n-1}\}$ qui contient une base par hypothèse de récurrence. □

Théorème 3.18 dit de la base incomplète . *Si E est de dimension finie, toute famille libre peut être complétée en une base.*

Démonstration. Supposons $\dim E = n$. Soit $A = \{a_1, \dots, a_m\}$ une famille libre dans E on a nécessairement $m \leq n$. Si $m = n$, A est alors une base et le théorème est prouvé. Sinon, il existe un vecteur $x_1 \in E$ qui ne soit pas combinaison linéaire de A , et d'après une règle précédente, $A_1 = \{a_1, \dots, a_m, x_1\}$ est libre. On renouvelle le procédé en choisissant $x_2 \notin \mathrm{Vect}(A_1)$ on obtient une famille libre $A_2 = \{a_1, \dots, a_m, x_1, x_2\}$, et ainsi de suite jusqu'à obtenir une famille libre de n éléments qui est donc nécessairement une base. □

2.6 Rang d'une famille de vecteurs.

Définition 3.19
On appelle rang d'un système (ou d'une famille) de vecteurs $\{a_1, \dots, a_m\}$ la dimension de $\mathrm{Vect}(a_1, \dots, a_m)$.

Théorème 3.20 *On ne change pas $\mathrm{Vect}(a_1, \dots, a_m)$ donc le rang d'un système en :*
- *multipliant l'un des vecteurs par un scalaire non nul.*
- *ajoutant à l'un des vecteurs une combinaison linéaire des autres.*

Démonstration. Vérifions la deuxième affirmation qui s'crit : $\mathrm{Vect}(a_1, \dots, a_n) = \mathrm{Vect}(a_1 + \sum_{1<i} \alpha_i a_i, a_2, \dots, a_n)$. En effet : il s'agit de vérifier que toute combinaison linéaire de $\{a_1, \dots, a_n\}$ est combinaison linéaire de $\{a_1 + \sum_{1<i} \alpha_i a_i, a_2, \dots, a_n\}$ et réciproquement. Il est clair que toute CL de la deuxième famille appartient à $\mathrm{Vect}(a_1, \dots, a_n)$. Réciproquement soit $x \in \mathrm{Vect}(a_1, \dots, a_n)$ alors x s'écrit :

$$\begin{aligned} x &= \lambda_1 a_1 + \ldots + \lambda_n a_n \\ &= \lambda_1 a_1 + \ldots + \lambda_n a_n + \lambda_1 \sum_{1<i} \alpha_i a_i - \lambda_1 \sum_{1<i} \alpha_i a_i \\ &= \lambda_1 (a_1 + \sum_{1<i} \alpha_i a_i) + (\lambda_2 - \lambda_1 \alpha_2) a_2 + \ldots + (\lambda_n - \lambda_1 \alpha_n) a_n \ . \end{aligned}$$

□

En particulier si une famille $\{a_1, \dots, a_m\}$ est libre, son rang est égal à sa longueur, la famille est alors une base du sev qu'elle engendre. Le théorème précédent prouve en particulier qu'une base le reste si on ajoute à un des vecteurs une CL des autres ou si on multiplie l'un des vecteurs par un scalaire non nul.

2.7 Espace des colonnes – Image d'une matrice

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, pour tout vecteur $x = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \in \mathbb{K}^p$, alors $Ax \in \mathbb{K}^n$. Ainsi A dfinit une application de $\mathbb{K}^p$ dans $\mathbb{K}^n$. Cette application vrifie $A(x+y) = Ax+Ay$ et $A(\lambda x) = \lambda Ax$, pour $x \in \mathbb{K}^p$, $y \in \mathbb{K}^p$ et $\lambda \in \mathbb{K}$ sont quelconques.

L'ensemble image de A , $A(\mathbb{K}^p)$, sera not $\mathrm{Im}(A)$.

Le vecteur Ax est combinaison linaire des colonnes $c_i(A)$, ainsi $\mathrm{Im}(A) \subset \mathrm{Vect}(c_1(A), \ldots, c_p(A))$.

Rciproquement tout lment $\lambda_1 C_1(A) + \ldots + \lambda_p C_p(A) \in \mathrm{Vect}(c_1(A), \ldots, c_p(A))$ s'crit $A \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix}$ donc appartient $\mathrm{Im}(A)$. On peut rsumer ainsi :

Définition 3.21
On appelle espace des colonnes *ou* Image *de* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *, le sous-espace vectoriel* $\mathrm{Vect}(c_1(A), \ldots, c_p(A)) \subset \mathbb{K}^n$ *. Il est not* $\mathrm{Im}(A)$.

Exemple

Déterminons l'image des matrices suivantes :

$$\mathrm{I}d_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}.$$

$\mathrm{Im}(\mathrm{I}d_2) = \mathbb{R}^2$, $\mathrm{Im}(A)$ est la droite passant par l'origine et de vecteur directeur $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

$$\mathrm{Im}(A) = \{\lambda \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \lambda \in \mathbb{R}\}.$$

$$\mathrm{Im}(B) = \{\lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \lambda, \mu \in \mathbb{R}\} = \mathbb{R}^2.$$

2.8 Bases et espace des colonnes

Nous allons voir comment l'chelonnement permet de dterminer une base de l'image d'une matrice.

Considrons une matrice carre $A \in \mathcal{M}_n(\mathbb{K})$, dont les colonnes sont notes $c_1, \ldots, c_n$. Rappelons que $\mathrm{Im}(A) = \mathrm{Vect}(c_1, \ldots, c_n)$. On a la proprit fondamentale :

Proposition 3.22 *La matrice carre A est inversible si et seulement si $\{c_1, \ldots, c_n\}$ est une famille libre. Donc si et seulement si les colonnes forment une base de $\mathbb{K}^n$.*

Démonstration. Supposons A inversible. Soit $x_1 c_1 + \ldots + x_n c_n = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ une combinaison linaire nulle des colonnes. On en dduit :

$$A^{-1} A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

donc $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, donc les colonnes forment une famille libre.

Rciproquement on suppose que $\{c_1, \ldots, c_n\}$ est une famille libre, donc est une base de $\mathbb{K}^n$ de dimension n. En particulier chaque vecteur $\boldsymbol{e}_j$ de la base canonique de $\mathbb{K}^n$ s'crit comme combinaison linaire des colonnes :

$$\boldsymbol{e}_j = \alpha_{1j} c_1 + \ldots + \alpha_{nj} c_n .$$

Par construction, la matrice $B = \begin{bmatrix} \alpha_{ij} \end{bmatrix}$ vrifie $AB = I_n$. Il est alors admis que $B = A^{-1}$ (cf. chapitre 2 paragraphe 1.4). $\square$

De plus :

> **Proposition 3.23** *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice échelonnée non nulle. Alors toute colonne non pivotale est :*
> *1) nulle si elle est situe gauche de la premire colonne pivotale,*
> *2) combinaison linaire des colonnes pivotales qui la prcdent.*
> *De plus les colonnes pivotales de A forment une famille libre.*

Démonstration. La proprit 1) est une consquence de la dfinition d'une matrice chelonne. La preuve de 2) est fastidieuse et nous l'omettons. Montrons l'indpendance linaire des colonnes pivotales. Soit $1 \le r \le n$ le nombre de lignes non nulles de A. Pour $1 \le i \le r$, on note j_i l'indice de la colonne qui contient le premier terme non nul de la ligne i. Par le point 2. de la définition d'une matrice échelonnée, $j_1 < \cdots < j_r$. Pour $1 \le i \le r$, la colonne d'indice j_i est de la forme $c_{j_i}(A) = \begin{bmatrix} a_{1,j_i} \\ \vdots \\ a_{i,j_i} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ avec $a_{i,j_i} \neq 0$. Soient $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$ tels que le vecteur colonne $X := \alpha_1 c_{j_1}(A) + \cdots + \alpha_r c_{j_r}(A)$ soit nul. La ligne r de X est $\alpha_r a_{r,j_r}$. Donc $\alpha_r = 0$. On montre de même (par récurrence) que $\alpha_{r-1} = \cdots = \alpha_1 = 0$. Donc la famille des colonnes pivotales est libre. $\square$

En combinant les deux dernires propositions on dduit les deux consquences suivantes :

Proposition 3.24 *Dans une matrice chelonne, les colonnes pivotales forment une base de l'espace des colonnes.*

Proposition 3.25 *Une matrice carre d'ordre n est inversible si et seulement si elle possde n pivots.*

Nous allons en dduire une rgle pour les matrices non ncessairement chelonnes. Rappelons que l'chelonnement d'une matrice A revient effectuer un produit PA o P est inversible. Considrons une famille $\{u_1,\ldots,u_k\}\subset\mathbb{K}^n$ et P une matrice inversible d'ordre n. Dans ces conditions :

$$\{u_1,\ldots,u_k\} \text{ est libre } \Leftrightarrow \quad \{Pu_1,\ldots,Pu_k\} \text{ libre.}$$
$$u=\sum_{i=1}^{k}\lambda_i u_i \quad \Leftrightarrow \quad Pu=\sum_{i=1}^{k}\lambda_i Pu_i$$

Ces quivalences sont videntes : il suffit d'appliquer P^{-1} aux galits concernes.

On peut alors noncer la rgle :

Les colonnes de A qui deviennent pivotales aprs chelonnement de A forment une base de l'image de A.

Cela prouve en passant que le nombre de pivot est gal la dimension de $\mathrm{Im}(A)$, en cohrence avec la dfinition du rang d'une famille de vecteurs. On peut aussi affirmer que le rang d'une matrice $n\times p$ est ncessairement infrieur au nombre de colonnes et la dimension de $\mathbb{K}^n$ donc $\mathrm{rg}(A)\leq\min(n,p)$.

2.9 Noyau d'une matrice – systŔmes linŐaires homogŔnes

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$.

Définition 3.26
On appelle **noyau** *de A, not $\ker(A)$, l'ensemble des vecteurs $\boldsymbol{x}\in\mathbb{K}^p$ tels que $A\boldsymbol{x}=0$.*

Proposition 3.27 *Le noyau de A est un sous-espace vectoriel de $\mathbb{K}^p$.*

Démonstration. Il est clair que $\ker(A)$ n'est pas vide, puisqu'au moins $0\in\ker(A)$. Supposons que $\boldsymbol{x}=(x_1,\ldots,x_p)$ et $\boldsymbol{y}=(y_1,\ldots,y_p)$ appartiennent $\ker(A)$. On a :

$$A(\boldsymbol{x}+\boldsymbol{y})=\sum_{i=1}^{p}(x_i+y_i)c_i(A)=\sum_{i=1}^{p}x_ic_i(A)+\sum_{i=1}^{p}y_ic_i(A)=A\boldsymbol{x}+A\boldsymbol{y}=0\,.$$

De plus, pour tout $\lambda\in\mathbb{K}$, $A(\lambda\boldsymbol{x})=\lambda A\boldsymbol{x}=0$. Ainsi $\ker(A)$ est stable par combinaison linaire. $\square$

Dterminer le noyau d'une matrice A revient dterminer l'ensemble des solutions $(x_1,\ldots,x_p)$

d'un systme linaire **homogne**, c'est--dire dont les seconds membres sont tous nuls :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1p}x_p = 0 \\ \vdots \\ a_{n1}x_1 + \ldots + a_{np}x_p = 0 \end{cases} \quad \Leftrightarrow \quad A\boldsymbol{x} = 0 \, .$$

On remarque que si P est inversible alors $A\boldsymbol{x} = 0$ est quivalent $PA\boldsymbol{x} = 0$. On se souvient que l'chelonnement revient multiplier gauche A par une matrice inversible, d'o ce rsultat important :

Le noyau d'une matrice est invariant aprs chelonnement.

Par ailleurs :

Proposition 3.28 *Les vecteurs colonnes d'une matrice* A *forment une famille libre si et seulement si* $\ker(A) = \{0\}$.

Démonstration. En effet, une combinaison linaire nulle des colonnes de A s'crit :

$$A \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix} = \sum_{k=1}^{p} \lambda_k c_k(A) = 0 \;\Leftrightarrow\; \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix} \in \ker(A) \, .$$

Ainsi le noyau est rduit zro si et seulement si les λ_i sont tous nuls. □

2.10 Échelonnement total et base du noyau

La pratique de la rsolution d'un systme linaire $A\boldsymbol{x} = \boldsymbol{b}$, par les mthodes d'limination, revient rsoudre un systme quivalent $PA\boldsymbol{x} = P\boldsymbol{b}$ o P est inversible, ainsi le noyau de PA reste gal celui de A. L'chelonnement total permet de dterminer en pratique une base de ce noyau.

Commençons par le cas d'une matrice A carrée ($n = p$).

Cas d'une matrice A carrée ($n = p$) inversible Si A est inversible, la seule solution du système $A\boldsymbol{x} = \boldsymbol{0}$ est $\boldsymbol{x} = \boldsymbol{0}$. On verra que la matrice A est inversible si et seulement si la forme totalement échelonnée de A est $R = \mathrm{Id}_n$. Dans ce cas, la matrice R a donc $r = n$ colonnes pivotales et aucune colonne non pivotale, et donc son rang est $r = n$.

Cas d'une matrice A carrée ($n = p$) non inversible Dans ce cas, la forme totalement échelonnée de A est $R \neq \mathrm{Id}_n$, et son rang (le nombre de colonnes pivotales) $r < n$: la matrice R possède au moins une ligne de 0 (en bas).

Exemple :

Considérons par exemple la matrice $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, et cherchons à déterminer le noyau de A. C'est l'ensemble des $\boldsymbol{x} = (x_1, x_2)$ solutions du système :

$$\begin{cases} x_1 + 2x_2 = 0 \\ 2x_1 + 4x_2 = 0 \end{cases}$$

Faisons l'élimination : $\ell_2 \to \ell_2 - 2\ell_1$. On obtient :

$$\begin{cases} x_1 + 2x_2 = 0 \\ 0 = 0 \end{cases}$$

La deuxième équation est toujours satisfaite, et donc il n'y a en fait qu'une seule équation. Le noyau de A est la droite d'équation $x_1 + 2x_2 = 0$.

Son quation paramtrique est un moyen simple de décrire cette droite de solutions. On choisit un point de la droite, donc une gnrateur du noyau. Les points de la droite sont tous des multiples de ce point (parce que la droite passe par $\mathbf{0}$). prenons par exemple $x_2 = 1$, dans ce cas on a $x_1 = -2$, et donc une solution particulière est $\boldsymbol{s} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Ainsi :

$$\mathrm{Ker}(A) = \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix}, t \in \mathbb{R} \right\}.$$

Cas d'une matrice rectangulaire quelconque

Exemple :

Le système linéaire à une équation et trois inconnues $x_1 + x_2 + x_3 = 0$, s'écrit aussi $A\boldsymbol{x} = \mathbf{0}$ où A est la matrice ligne : $A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, dont la forme totalement échelonnée est $R = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, qui n'a donc qu'un pivot ($r = 1$). Le noyau de A est le plan $\mathcal{P}$ de $\mathbb{R}^3$ d'équation $x_1 + x_2 + x_3 = 0$, dont on peut trouver deux solutions non colinéaires, qu'on dtermine en choisissant d'abord $x_2 = 1$ et $x_3 = 0$ puis $x_2 = 0$ et $x_3 = 1$:

$$\mathrm{s}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{et } \mathrm{s}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Notons que les variables x_2 et x_3 correspondant aux colonnes non pivotales sont arbitraires, ce sont des « variables libres » (au sens où on peut les choisir comme on veut) mais une fois qu'elles sont choisies, la première (et unique) ligne de la matrice totalement échelonnée R nous donne $x_1 = -x_2 - x_3$.

La première colonne de A contient le pivot, et donc la première composante n'est pas libre. On choisit des solutions gnratrices grâce aux variables libres correspondant aux colonnes non pivotales.

Remarquons enfin que les solutions $\mathbf{s_1}$ et $\mathbf{s_2}$ engendrent tout le plan $\mathcal{P}$ d'équation $x_1 + x_2 + x_3 = 0$. En effet, pour n'importe quel vecteur $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ du plan $\mathcal{P}$, on peut écrire

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = x_2 \boldsymbol{s}_1 + x_3 \boldsymbol{s}_2,$$

grâce au fait que $-x_2 - x_3 = x_1$.

On peut donc décrire $\mathrm{Ker}(A)$ (le plan $\mathcal{P}$) comme l'ensemble des combinaisons linéaires des solutions particulires qu'on vient de calculer :

$$\mathrm{Ker}(A) = \left\{ \lambda \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \mu \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \lambda \in \mathbb{R}, \mu \in \mathbb{R} \right\}.$$

Une base des solutions de $\ker(A)$ peut en fait se calculer très simplement à partir de la forme totalement échelonnée R de A : si on regarde comment sont faites les colonnes de R, on se rend compte qu'une colonne non pivotale peut s'écrire comme combinaison linéaire des colonnes pivotales précédentes. Et ce sont justement les coefficients de cette combinaison linéaire qui donnent une solution particulire de $A\boldsymbol{x} = \mathbf{0}$, puisque le produit $A\boldsymbol{x}$ est une combinaison linéaire des colonnes de A.

Exemple :

Prenons la matrice 3×4 sous forme totalement échelonnée

$$R = \begin{bmatrix} 1 & \mathbf{2} & 0 & \mathbf{3} \\ 0 & \mathbf{0} & 1 & \mathbf{-1} \\ 0 & \mathbf{0} & 0 & \mathbf{0} \end{bmatrix}$$

où on a écrit en gras les colonnes **non** pivotales. Notons $\boldsymbol{c}_j(R)$, $j = 1, \ldots, 4$ les 4 colonnes de R.

Les deux colonnes pivotales sont $\boldsymbol{c}_1(R)$ et $\boldsymbol{c}_3(R)$, et les deux non pivotales $\boldsymbol{c}_2(R)$ et $\boldsymbol{c}_4(R)$.

Chercher $\boldsymbol{x}$ tel que $R\boldsymbol{x} = 0$ revient à chercher $x_1, \ldots, x_4$ tels que $\sum_{i=1}^{4} x_i \boldsymbol{c}_i(R) = 0$. Or la lecture de la matrice totalement échelonnée R donne

$$\boldsymbol{c}_2(R) = 2\boldsymbol{c}_1(R) \text{ et } \boldsymbol{c}_4(R) = 3\boldsymbol{c}_1(R) - \boldsymbol{c}_3(R)$$

ce qui s'écrit aussi (pour bien faire apparaître le produit matrice vecteur) :

$$-2\,\boldsymbol{c}_1(R) + \boldsymbol{c}_2(R) + 0\,\boldsymbol{c}_3(R) + 0\,\boldsymbol{c}_4(R) = \mathbf{0} \text{ et } -3\,\boldsymbol{c}_1(R) + 0\,\boldsymbol{c}_2(R) + \boldsymbol{c}_3(R) + 1\,\boldsymbol{c}_4(R) = \mathbf{0}.$$

On a donc trouvé ainsi deux solutions qui vérifient $Ax = \mathbf{0}$, qui sont

$$\boldsymbol{s}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ et } \boldsymbol{s}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Ici encore, on peut remarquer que tout vecteur $\boldsymbol{x}$ qui vérifie $A\boldsymbol{x} = \mathbf{0}$ peut s'écrire comme combinaison linéaire de $\boldsymbol{s}_1$ et $\boldsymbol{s}_2$. En effet, pour n'importe quel vecteur $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ vérifiant $A\boldsymbol{x} = 0$, on a : $x_1 = -2x_2 - 3x_4$ et $x_3 = x_4$; on peut donc écrire

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 1 \end{bmatrix} = x_2\boldsymbol{s}_1 + x_4\boldsymbol{s}_2.$$

Les vecteurs $\boldsymbol{s}_1$ et $\boldsymbol{s}_2$ engendrent l'espace des solutions de $A\boldsymbol{x} = \mathbf{0}$.

Le noyau de A s'écrit alors :

$$\mathrm{Ker}A = \left\{ \lambda \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} -3 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \lambda \in \mathbb{R},\ \mu \in \mathbb{R} \right\}.$$

Rappelons que si R est la forme totalement échelonnée de A, les noyaux de A et de R sont donc identiques.

Il y a autant de solutions spéciales à $A\boldsymbol{x} = \mathbf{0}$ que de colonnes non pivotales, c.à.d. $p - r$. Dans l'exemple ci-dessus : $p = 4$, $r = 2$, donc il y a 2 solutions spéciales.

Cas d'une matrice $n \times p$ **avec** $p > n$ Remarquons que dans le dernier exemple ci-dessus, on a une matrice rectangulaire « couchée », c.à.d. avec plus d'inconnues que de lignes : $p > n$. Dans ce cas, le noyau de la matrice est forcément non réduit à $\mathbf{0}$. Le nombre r de pivots (ou de lignes non nulles) est forcément inférieur aux nombres de lignes, n. Comme $p > n$, il existe au moins une colonne non pivotale, et cette colonne est une combinaison linéaire d'autres colonnes de A. Il y a donc une infinité de solutions au système $A\boldsymbol{x} = \mathbf{0}$ (tous les multiples de s), comme dans l'exemple ci-dessus.

Unicité de la forme totalement échelonnée Question : la forme totalement échelonnée d'une matrice A est elle unique ? La réponse est oui, et on a même un résultat plus fort que cela : une matrice totalement échelonnée est entièrement déterminée par ses dimensions et son noyau.

Exemple :

Soit A une matrice 3×4 dont le noyau est l'ensemble $\{\boldsymbol{x}\in\mathbb{R}^2; x_1=x_4=0, x_2-2x_3=0\}$. Construisons sa forme totalement échelonnée. La première colonne de R est donc non nulle, sinon le vecteur $(1,0,0,0)$ serait dans $\mathrm{Ker}S$. On a donc forcément $\boldsymbol{c}_1(R)=\begin{bmatrix}1\\0\\0\end{bmatrix}$. Pour la même raison, la deuxième colonne est aussi non nulle. De plus, le vecteur $(1,-1,0,0)$ n'est pas dans le noyau. Donc forcément $\boldsymbol{c}_2(R)=\begin{bmatrix}0\\1\\0\end{bmatrix}$. La troisième colonne est entièrement déterminée par l'équation du noyau : $\boldsymbol{c}_3(R)=2\boldsymbol{c}_2(R)$. Enfin, la quatrième colonne ne dépend pas des trois premières puisque x_4 n'est pas lié aux autres variables dans les équations du noyau de A. On a donc un seul choix possible pour R :

$$R=\begin{bmatrix}1&0&0&0\\0&1&2&0\\0&0&0&1\end{bmatrix}$$

La construction qu'on vient de faire dans cet exemple peut se généraliser à n'importe quelle matrice.

2.11 Dimension sur $\mathbb{C}$ et dimension sur $\mathbb{R}$

On dmontrera titre d'exercice le fait suivant :

soit E un $\mathbb{C}$-espace vectoriel de dimension n, et $\{a_1,\dots,a_n\}$ une base de E. E peut tre considr comme espace vectoriel sur $\mathbb{R}$ (la rciproque est fausse en gnral), montrer qu'alors $\{a_1, ia_1,\dots,a_n, ia_n\}$ est une base de E comme $\mathbb{R}$-espace vectoriel.

2.12 Sommes – Sommes directes

Soient A et B deux parties d'un $\mathbb{K}$-ev E. On définit

$$A+B=\{x\in E\ /\ x=a+b,\ \ a\in A,\ \ b\in B\ \}$$

On a la proposition évidente suivante :

Proposition 3.29 *Si A et B sont des sev alors $A+B$ l'est aussi.*

Définition 3.30
On dira qu'un ev E est somme directe de deux sev F et G si

$$E=F+G \ \ et \ \ F\cap G=\{0_E\}$$

On dira alors que F et G sont supplémentaires et on notera $E=F\oplus G$.

On a une caractérisation d'une somme directe, donne par l'nonc suivant :

Proposition 3.31 *$E = F \oplus G$ si et seulement si tout vecteur $x \in E$ s'écrit de façon unique comme $x = u + v$ où $u \in F$ et $v \in G$.*

Démonstration. Si $E = F \oplus G$ alors tout vecteur de E s'crit $x = u + v$ avec $u \in F$ et $v \in G$, si $x = u + v = u' + v'$ avec $u' \in F$ et $v' \in G$ alors $u - u' = v' - v$ est lment dans $F \cap G = \{0_E\}$ donc $u = u'$ et $v = v'$. Supposons rciproquement que tout vecteur x de E s'crive de faon unique $x = u + v$ avec $u \in F$ et $v \in G$, on a donc $E = F + G$, il reste vrifier que $F \cap G = \{0_E\}$; soit $x \in F \cap G = \{0_E\}$, $x = 0_E + x = x + 0_E$, galit dans laquelle x et 0_E sont considrs chacun son tour comme lment de F puis de G, donc par hypothse d'unicit on a $x = 0_E$. □

En dimension finie on dispose d'un théorème relatif aux bases :

Théorème 3.32 *Soit $\{a_1, \ldots, a_p\}$ une base de F et $\{b_1, \ldots, b_q\}$ une base de G, alors :*

$$E = F \oplus G \Leftrightarrow \{a_1, \ldots, a_p, b_1, \ldots, b_q\} \text{ base de } E$$

Démonstration. si $E = F \oplus G$ alors $\{a_1, \ldots, a_p, b_1, \ldots, b_q\}$ est génératrice de $E = F + G$. Vrifions qu'elle est libre. Si

$$\lambda_1 a_1 + \ldots + \lambda_p a_p + \mu_1 b_1 + \ldots + \mu_q b_q = 0$$

alors

$$\lambda_1 a_1 + \ldots + \lambda_p a_p = -(\mu 1 b_1 + \ldots + \mu_q b_q) \in F \cap G = \{0\},$$

donc tous les coefficients λ_i et μ_j sont nuls par indpendance linaire des bases respectives de F et de G.

Réciproquement : si $\{a_1, \ldots, a_p, b_1, \ldots, b_q\}$ est une base de E il est clair que $E = F + G$, de plus si $x \in F \cap G$ alors $x = \sum_{i=1}^{p} \lambda_i a_i = \sum_{j=1}^{q} \mu_j b_j$ d'où $\sum_{i=1}^{p} \lambda_i a_i - \sum_{j=1}^{q} \mu_j b_j = 0_E$ donc tous les coefficients sont nuls puisque la famille est une base donc $x = 0_E$. □

Corollaire 3.33 *En dimension finie, tout sous-espace admet un supplémentaire.*

Démonstration. c'est une consquence directe du thorme de la base incomplte et du thorme prcdent : soit F un sous-espace de E (suppos de dimension finie). Toute base $\{a_1, \ldots, a_p\}$ de F peut tre complte par $\{b_1, \ldots, b_q\}$ en une base de E, et d'aprs le thorme prcdent on a $E = F \oplus \operatorname{Vect}(b_1, \ldots, b_q)$. □

Si F et G sont des sev de dimension finie on a alors la *formule de Grassman*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$$

Indication pour une vérification de cette relation : compléter une base $\{a_1, \ldots, a_p\}$ de

$F \cap G$ en une base $\{a_1, \ldots, a_p, f_1, \ldots, f_l\}$ de F et en une base $\{a_1, \ldots, a_p, g_1, \ldots, g_m\}$ de G puis vérifier que $\{a_1, \ldots, a_p, f_1, \ldots, f_l, g_1, \ldots, g_m\}$ est une base de $F + G$.

Chapitre 4

Applications linéaires

Les applications linéaires constituent un outil absolument essentiel. On rencontre de tels objets dans tous les recoins des mathématiques. Ramener un problème à une situation linéaire permet de le simplifier.

Dans ce chapitre E et F désigneront deux espaces vectoriels sur un même corps $\mathbb{K}$.

1. Définitions et propriétés

Définition 4.1

On dira qu'une application $f : E \to F$ *est* linéaire*, si pour tous* x , y *dans* E *et pour tous scalaires* λ *et* μ *on a :*

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \ .$$

On dira également qu'une telle application est un homomorphisme *d'espaces vectoriels.*

Une récurrence montre alors facilement qu'une application linéaire f respecte les combinaisons linéaires :

$$f(\sum_{i=1}^{n} \lambda_i x_i) = \sum_{i=1}^{n} \lambda_i f(x_i)$$

1.1 L'espace $\mathcal{L}(E, F)$

l'ensemble des applications linéaires de E vers F n'est jamais vide, car possédant au moins l'application nulle. On le notera $\mathcal{L}(E, F)$. Il est lui même un espace vectoriel car naturellement sous-espace vectoriel de F^E (cf. chap.3, **2.2**).

L'espace $\mathcal{L}(E, E)$ est noté $\mathcal{L}(E)$. Un élément de $\mathcal{L}(E)$ est appelé **endomorphisme de** E .

Remarques :

- Si f est linéaire alors $f(0_E) = 0_F$. En effet : $f(0_E) = f(0_E + 0_E) = f(0_E) + f(0_E)$ d'où $f(0_E) = 0_F$. Il s'en suit que, pour tout $x \in E$, $0_F = f(x+(-x)) = f(x)+f(-x)$ donc $f(-x) = -f(x)$.
- On montrera que f est linéaire si et seulement si, pour tous x , $y \in E$ et λ scalaire, on a $f(x + y) = f(x) + f(y)$ et $f(\lambda x) = \lambda f(x)$.

Remarquons que l'application identique $I_E : E \to E$ est linéaire ainsi que l'application nulle $\mathbf{0} : E \to F$.

Exemples :

1. $f : \mathbb{K}[X] \to \mathbb{K}[X]$, $f(P) = P'$.
2. $E = \mathcal{C}([a, b], \mathbb{R}) \ni \phi \to \displaystyle\int_a^b \phi(t)dt$.
3. Pour tout $a \in \mathbb{K}$ fixé, $val(a) : \mathbb{K}[X] \to \mathbb{K}$, $val(a)(P) = P(a)$.
4. Soit E l'espace des suite réelles convergentes. Alors l'application $\lim : E \to \mathbb{R}$ qui, à toute suite, associe sa limite, est linéaire.
5. $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par $f(x, y, z) = (2x - y + z, 5x - 3z)$
6. $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $f(x, y, z) = x$.
7. $(x, y, z) \to (x, 0, 0)$ est linéaire mais $(x, y, z) \to (x, 1, 1)$ n'est pas linéaire car $(0, 0, 0)$ n'a pas $(0, 0, 0)$ comme image.
8. Si $f : \mathbb{K} \to K$ est linéaire alors, pour tout $x \in \mathbb{K}$ on peut écrire $f(x) = f(x.1) = xf(1)$; en posant $a = f(1)$ on trouve que $f(x) = ax$. Réciproquement, si $a \in \mathbb{K}$ est fixé, il est immédiat que $x \to ax$ est linéaire de $\mathbb{K}$ dans lui-même. Ce résultat, caractérisant une application linéaire de $\mathbb{K}$ dans $\mathbb{K}$ par la donnée d'un scalaire, se généralisera en dimension finie en la caractérisation d'une application linéaire par une matrice.

Les vérifications de la linéarité des exemples ci-dessus sont faciles et laissées en exercice.

1.2 Premières propriétés

Notons deux propriétés évidentes :

- La composée d'applications linéaires est linéaire : si $f : E \to F$ et $g : F \to G$ sont linéaires alors $g \circ f$ est linéaire. Il suffit en effet de développer : $g(f(\lambda x + \mu y) = g(\lambda f(x) + \mu f(y)) = \lambda g(f(x)) + \mu g(f(y))$.
- La restriction d'une application linéaire $f : E \to F$ à tout sous-espace vectoriel de E est encore linéaire.

Plus généralement, l'effet des applications linéaires sur les sous-espaces est simple, la propriété d'être un sous-espace est conservée. On a précisément l'énoncé suivant :

> **Proposition 4.2** *Soit* $f : E \to F$ *linéaire.*
> *1) Pour tout sous-espace* H *de* E *, l'image* $f(H)$ *est un sous-espace de* F *.*
> *2) Pour tout sous-espace* G *de* F *, l'image réciproque* $f^{-1}(G)$ *est un sous-espace de* E *.*

Démonstration. Notons que $f(H)$ et $f^{-1}(G)$ ne sont pas vides ; l'égalité $f(0_E) = 0_F$ indique que le premier possède $f(0_E)$ et le second possède 0_E . Il reste à vérifier la stabilité de $f(H)$ et de $f^{-1}(G)$ pour les combinaisons linéaires. Pour tous scalaires λ et μ , x et y , quelconques de $f(H)$, sont de la forme $x = f(h)$ et $y = f(h')$ où h , $h' \in H$. On aura donc $\lambda x + \mu y = \lambda f(h) + \mu f(h') = f(\lambda h + \mu h') \in f(H)$. Pour tous x et $y \in f^{-1}(G)$, on a $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \in G$ donc $\lambda x + \mu y \in f^{-1}(G)$. □

Cette proposition a deux cas particuliers importants :

1) Le sous-espace $f(E)$ est noté $\mathrm{Im}(f)$ et est appelé *image de* f .

2) Le sous-espace $f^{-1}(\{0_F\})$, noté $\ker(f)$, est appelé *noyau* de f .

Important : deux points x et y de E auront même image par f : $f(x) = f(y)$ si et seulement si $f(x - y) = 0$ donc si et seulement si $x - y \in \ker(f)$.

> **Proposition 4.3**
> *1)* f *est surjective* $\Leftrightarrow$ $\mathrm{Im}(f) = F$ *.*
> *2)* f *est injective* $\Leftrightarrow$ $\ker(f) = \{0_E\}$ *.*

Démonstration. La première affirmation est une conséquence de la définition même d'une surjection. Pour ce qui est de la deuxième affirmation : $f(x) = f(y)$ étant équivalent à $x - y \in \ker(f)$, si le noyau est réduit à zéro, alors $x = y$ et f est injective. Réciproquement, tout $x \in \ker(f)$ vérifie $f(x) = f(0_E) = 0_F$, donc l'injectivité de f nécessite que $x = 0_E$. □

Isomorphismes

> **Proposition 4.4** *Si une application linéaire* f *est bijective, alors sa bijection réciproque* f^{-1} *est aussi linéaire.*

Démonstration. En effet, puisque f est bijective, l'égalité :

$$f^{-1}(\lambda x + \mu y) = \lambda f^{-1}(x) + \mu f^{-1}(y) ,$$

est équivalente, en appliquant f à chaque membre, à :

$$\lambda x + \mu y = f(\lambda f^{-1}(x) + \mu f^{-1}(y)) ,$$

mais f est linéaire, l'égalité devient donc $\lambda x+\mu y = \lambda x+\mu y$ qui est vraie, donc la première égalité est vraie : f^{-1} est bien linéaire. □

On dira alors, dans ce cas que f est un **isomorphisme**. Si $E = F$ on parle alors **d'automorphisme**

- Remarquons que la composée d'isomorphismes est encore un isomorphisme.

1.3 Applications linéaires et dimension

Il est essentiel de remarquer qu'une application linéaire est caractérisée par les images des vecteurs de bases.

On a vu qu'une application linéaire $f : \mathbb{K} \to E$ est connue dès lors que l'on fixe la valeur qu'elle prend pour $x = 1$. Cela est un cas particulier du fait suivant : une application linéaire est déterminée par les valeurs qu'elle prend sur une base. Si $x = \sum_{i=1}^{n} x_i e_i$ alors $f(x) = \sum_{i=1}^{n} x_i f(e_i)$, il suffit donc de connaitre les $f(e_i)$ pour calculer $f(x)$. Plus précisément :

Théorème 4.5 *Soit $\{e_1, \dots, e_n\}$ une base de E. Pour toute famille $\{a_1, \dots, a_n\}$ dans F, il existe une unique application $f \in \mathcal{L}(E, F)$ telle que, pour tout i, $f(e_i) = a_i$.*

Démonstration. L'unicité découle de ce que, pour tout $x = \sum_{i=1}^{n} x_i e_i$, on a nécessairement $f(x) = \sum_{i=1}^{n} x_i a_i$. L'unicité des coordonnées de tout $x \in E$ dans la base fixée, fait que la correspondance $x \mapsto \sum_{i=1}^{n} x_i a_i$ est une application bien définie. Il suffit alors de vérifier que cette correspondance est linéaire (vérification laissée en exercice). □

Conséquence : une application linéaire est *caractérisée* par ses valeurs prises sur une base. Avec les notations précédentes on a, pour $f, g \in \mathcal{L}(E, F)$, l'équivalence :

$$\forall i \, , \; f(e_i) = g(e_i) \quad \Leftrightarrow \quad f = g \, .$$

Soit $f \in \mathcal{L}(E, F)$, et $\{u_1, \dots, u_n\} \subset E$, on a alors :

Proposition 4.6

$$f(\mathit{Vect}(u_1, \dots, u_n)) = \mathit{Vect}(f(u_1), \dots, f(u_n)) \, .$$

Démonstration. Ce résultat est une conséquence immédiate de l'égalité :

$$f(\lambda_1 u_1 + \ldots + \lambda_n u_n) = \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) \; . \; \square$$

Il s'ensuit que si E est de dimension finie, $\mathrm{Im}(f)$ le sera aussi. On peut alors donner, dans ce cas, la définition suivante :

Définition 4.7
On appellera rang *de* f *, la dimension de* $\mathrm{Im}(f) = f(E)$ *, que l'on notera* $\mathrm{rg}(f)$ *.*

On démontrera à titre d'exercice les deux propositions suivantes :

Proposition 4.8 *Soit* $f \in \mathcal{L}(E, F)$ *,* f *est injective si et seulement si l'image par* f *de toute famille libre est libre.*

On suppose à présent que E est de dimension finie.

Corollaire 4.9 $f \in \mathcal{L}(E, F)$ *est un isomorphisme si et seulement si l'image par* f *d'une base de* E *est une base de* F *.*

Démonstration.

Ce dernier résultat, avec le théorème 4.5, aboutissent au résultat suivant :

Corollaire 4.10 *Deux espaces vectoriels* E *et* F *, de dimension finie, sont isomorphes si et seulement si ils ont même dimension.*

En particulier tout $\mathbb{K}$-espace de dimension n est isomorphe à $\mathbb{K}^n$.

Le fait que l'image par un isomorphisme d'une base soit une base, implique l'invariance du rang par isomorphisme :

Proposition 4.11 *Soit* f *une application linéaire et* g *un isomorphisme tel que* $f \circ g$ *(resp.* $g \circ f$ *) soit défini. Dans ces conditions :* $\mathrm{rg}(f \circ g) = \mathrm{rg}(f)$ *(resp.* $\mathrm{rg}(g \circ f) = \mathrm{rg}(f)$ *).*

Le théorème suivant, dit *Théorème Rang*, est fondamental pour ce qui concerne les rapports reliant dimension et applications linéaires.

Théorème 4.12 [du rang] *Soit* E *de dimension finie et* $f \in \mathcal{L}(E, F)$ *, alors :*
1) $\mathrm{Im}(f)$ *est isomorphe à tout supplémentaire de* $\ker(f)$ *.*
2) $\dim(E) = \mathrm{rg}(f) + \dim(\ker(f))$ *.*

Démonstration. Notons $\{k_1, \ldots, k_p\}$ une base de $\ker(f)$. Pour tout supplémentaire H de $\ker(f)$ et pour toute base $\{h_1, \ldots, h_m\}$ de H , alors $\{k_1, \ldots, k_p, h_1, \ldots, h_m\}$ sera une base de E (cf. 3.35), avec $\dim(E) = n = p + m$. Pour tout $x = \lambda_1 k_1 + \ldots + \lambda_p k_p + \mu_1 h_1 + \ldots + \mu_m h_m \in E$ alors $f(x) = \mu_1 f(h_1) + \ldots + \mu_m f(h_m)$, car $f(k_i) = 0$, ainsi $\{f(h_1), \ldots, f(h_m)\}$ est génératrice de $\mathrm{Im}(f)$. De plus, la restriction de f à H est injective car son noyau est égal à $H \cap \ker(f) = \{0\}$, donc $f_{|H}$ réalise un isomorphisme de H sur $\mathrm{Im}(f)$ dont la dimension est donc m . □

Théorème 4.13 *Si* $\dim(E) = \dim(F)$ *(finie) et* $f \in \mathcal{L}(E, F)$ *, on a les équivalences :*

$$f \text{ injective} \Leftrightarrow f \text{ surjective} \Leftrightarrow f \text{ bijective} .$$

Démonstration. En effet si $n = \dim(E) = \dim(F)$, l'égalité du rang donne $n = \mathrm{rg}(f) + \dim(\ker(f))$, ainsi $\ker(f)$ est réduit à zéro si et seulement si $\dim(F) =$

$\dim(f(E))$ donc si et seulement si f est surjective. $\square$

En particulier si $f \in \mathcal{L}(E)$ et $\dim(E) = n$, le théorème précédent s'applique. On est alors en présence d'un automorphisme de E.

• Le fait que E soit de dimension finie est essentiel. Un contre-exemple simple est le suivant : si $E = \mathbb{K}[X]$ alors $f(P) = P'$ est linéaire surjective de E dans E mais non injective (le montrer à titre d'exercice).

2. Matrices d'applications linéaires

Dans ce qui suit tous les espaces vectoriels seront supposés de dimension finie.

On a vu précédemment qu'une application linéaire $f :\to F$ est déterminée par les valeurs qu'elle prend sur une base de E. Ces valeurs vont elles-mêmes être repérées par leurs coordonnées dans une base de F, qui seront placées dans une matrice. Voici comment.

Attention : les bases sont désormais ordonnées.

Soit $\mathcal{B} = \{e_1, \ldots, e_p\}$ une base *fixée* de E. On représentera un vecteur $x = \sum_{i=1}^{p} x_i e_i$ par la matrice :

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

appelée *matrice de x dans la base $\mathcal{B}$*. Retenons cette première règle qui a été abondamment utilisée dans le chapitre 2 :

Matriciellement, les vecteurs sont représentés en colonnes.

Il est clair que si on permute deux vecteurs de $\mathcal{B}$, la matrice de x s'en trouvera changée, d'où la nécessité de considérer $\mathcal{B}$ avec son ordre. Une base $\mathcal{B} = \{e_1, e_2\}$ est différente de $\mathcal{B}' = \{e_2, e_1\}$.

Fixons donc $\mathcal{B} = \{e_1, \ldots, e_p\}$ comme base de E et $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ comme base de F. Soit $f \in \mathcal{L}(E, F)$. La matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est constituée par les vecteurs-colonnes $f(e_1), f(e_2), \ldots, f(e_p)$ représentés dans la base $\mathcal{B}'$. Donc s'écrivent : $f(e_1) = a_{11}e'_1 + a_{21}e'_2 + \ldots + a_{n1}e'_n$, $f(e_2) = a_{12}e'_1 + a_{22}e'_2 + \ldots + a_{n2}e'_n$ etc. on aura donc une matrice à p colonnes et n lignes :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & & \vdots & \\ a_{n1} & a_{12} & \dots & a_{np} \end{bmatrix}$$

On notera parfois cette matrice $M_{\mathcal{B},\mathcal{B}'}(f)$. Ce qu'il faut retenir :

La colonne j de $M_{\mathcal{B},\mathcal{B}'}(f)$ est formée des coordonnées de $f(e_j)$ dans $\mathcal{B}'$.

La donnée d'une telle matrice caractérise f **, à condition que les bases soient fixées**. On a précisément l'énoncé suivant :

Théorème 4.14 *Les bases $\mathcal{B}$ de E et $\mathcal{B}'$ de F étant données, pour toute matrice à n lignes et p colonnes*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & & \vdots & \\ a_{n1} & a_{12} & \dots & a_{np} \end{bmatrix},$$

il existe une unique application $f \in \mathcal{L}(E,F)$ telle que $A = M_{\mathcal{B},\mathcal{B}'}(f)$.

Démonstration. Cela découle du théorème 4.5 : la règle de construction de $M_{\mathcal{B},\mathcal{B}'}(f)$ donne $f(e_i) = \sum_{j=1}^{p} a_{ji}e'_j$ donc détermine les valeurs des $f(e_i)$ donc détermine une unique f . □

2.1 Utilisation de la matrice d'une application linéaire.

Nous conservons les mêmes notations. Soit $\boldsymbol{x} \in E$ et $y \in F$ s'écrivant respectivement $\boldsymbol{x} = x_1e_1 + \ldots + x_pe_p$ et $\boldsymbol{y} = y_1e'_1 + \ldots + y_ne'_n$. Notons X et Y leur représentation matricielle en colonne, respectivement dans $\mathcal{B}$ et $\mathcal{B}'$ et A la matrice de f dans ces mêmes bases. On a :

$$\begin{aligned} f(x) &= \sum_{i=1}^{p} x_i f(e_i) \\ &= \sum_{i=1}^{p} x_i f(\sum_{j=1}^{n} a_{ji}e'_j) \\ &= (\sum_{j=1}^{p} x_j a_{j1})e'_1 + \ldots + (\sum_{j=1}^{p} x_j a_{jn})e'_n \end{aligned}$$

Dans ce cas, l'égalité $y = f(x)$ se traduira, pour tout $k \in \{1,\ldots,n\}$, par $y_k = \sum_{j=1}^{p} x_j a_{jk}$, donc par un système linéaire d'équations dont l'écriture matricielle est $Y = AX$. On a finalement l'équivalence suivante :

$$\boxed{y = f(x) \quad \Leftrightarrow \quad Y = AX \ .}$$

Cette première relation entre le calcul matriciel et les applications linéaires est complétée par le théorème fondamental suivant :

Théorème 4.15 *Soit E, F deux espaces vectoriels de bases respectives $\mathcal{B}$, $\mathcal{B}'$. Soit $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(E,F)$. On a alors,pour tout scalaire λ, la relation :*

$$M_{\mathcal{B},\mathcal{B}'}(\lambda f + g) = \lambda M_{\mathcal{B},\mathcal{B}'}(f) + M_{\mathcal{B},\mathcal{B}'}(g)$$

Si de plus, G est un espace vectoriel de base $\mathcal{B}''$, pour toutes applications linéaires $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$. On a la relation :

$$M_{\mathcal{B}',\mathcal{B}''}(g) M_{\mathcal{B},\mathcal{B}'}(f) = M_{\mathcal{B},\mathcal{B}''}(g \circ f) \ .$$

Autrement dit : **la matrice de la composée est égale au produit des matrices**.

Démonstration. Nous vérifions uniquement la dernière relation. Notons $\mathcal{B} = \{e_1, \dots, e_p\}$, $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ et $\mathcal{B}'' = \{e''_1, \dots, e''_m\}$. Notons $A = M_{\mathcal{B},\mathcal{B}'}(f) = [a_{ij}]$ et $B = M_{\mathcal{B}',\mathcal{B}''}(g) = [b_{kl}]$. Evaluons les coordonnées $g(f(e_j))$ dans $\mathcal{B}''$:

$$\begin{aligned}
g(f(e_j)) &= g(a_{1j}e'_1 + \ldots + a_{nj}e'_n) \\
&= a_{1j}g(e'_1) + \ldots + a_{nj}g(e'_n) \\
&= a_{1j}(b_{11}e''_1 + b_{21}e''_2 + \ldots + b_{n1}e''_m) + \\
& \quad \vdots \\
& \quad a_{nj}(b_{1n}e''_1 + b_{2n}e''_2 + \ldots + b_{nm}e''_m) \\
&= (\sum_{k=1}^{n} b_{1k}a_{kj})e''_k + \ldots + (\sum_{k=1}^{n} b_{nk}a_{kj})e''_m
\end{aligned}$$

Les coordonnées sont bien les coefficients de la colonne j du produit BA $\square$

Important : Ce dernier résultat, en sus de l'équivalence $y = f(x) \Leftrightarrow Y = AX$ (**2.1** ci-dessus), permet, une fois les bases à la source et au but des applications linéaires fixées, d'identifier applications linéaires en dimension finie et matrices.

Cela permet, par exemple, de déduire l'associativité du produit matriciel à partir de l'associativité de la composition des applications. De même, toutes les propriétés des opérations matricielles se déduisent de celles opérant sur $\mathcal{L}(E,F)$, telles que la double distributivité du produit par rapport à la somme.

Notons l'équivalence entre isomorphismes et matrices inversibles, précisée par le théorème suivant. Soient $\mathcal{B} = \{e_1, \ldots, e_p\}$ et $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ deux bases respectivement de E et F, dans ces conditions :

Théorème 4.16

- $M_{\mathcal{B},\mathcal{B}}(Id_E) = Id_p$

- *$f \in \mathcal{L}(E, F)$ est un isomorphisme si et seulement si $M_{\mathcal{B},\mathcal{B}'}(f)$ est inversible, et dans ce cas $n = p$ et $M_{\mathcal{B}',\mathcal{B}}(f^{-1}) = M_{\mathcal{B},\mathcal{B}'}(f)^{-1}$.*

Démonstration. Le premier point est évident : la colonne i de Id_E dans la base $\mathcal{B}$ est donnée par $Id_E(e_i) = e_i$. Supposons que $f \in \mathcal{L}(E, F)$ est un isomorphisme. D'après le corollaire **4.10** cela impose $n = p$. Notons $g = f^{-1}$. On a $g \circ f = Id_E$ et $f \circ g = I_F$, ce qui se traduit matriciellement par

$$\left\{\begin{array}{l} M_{\mathcal{B},\mathcal{B}}(g \circ f) = M_{\mathcal{B},\mathcal{B}}(Id_E) = Id_n \\ M_{\mathcal{B}',\mathcal{B}'}(f \circ g) = M_{\mathcal{B}',\mathcal{B}'}(Id_F) = Id_n \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} M_{\mathcal{B}',\mathcal{B}}(g) M_{\mathcal{B},\mathcal{B}'}(f) = Id_n \\ M_{\mathcal{B},\mathcal{B}'}(f) M_{\mathcal{B}',\mathcal{B}}(g) = Id_n \end{array}\right.$$

d'où l'inversibilité et $M_{\mathcal{B}',\mathcal{B}}(f^{-1}) = M_{\mathcal{B},\mathcal{B}'}(f)^{-1}$.

Supposons, réciproquement, que $M_{\mathcal{B},\mathcal{B}'}(f)$ soit inversible. Il existe une unique $g \in \mathcal{L}(F, E)$ telle que $M_{\mathcal{B}',\mathcal{B}}(g) = M_{\mathcal{B},\mathcal{B}'}(f)^{-1}$, ainsi $M_{\mathcal{B},\mathcal{B}}(g \circ f) = M_{\mathcal{B},\mathcal{B}}(Id_E) = Id_n$, par la proposition **4.14** d'unicité, on a $g \circ f = Id_E$. De même $f \circ g = Id_F$, ainsi f est un isomorphisme. $\square$

Ce théorème démontre que l'opération d'échelonnement d'une matrice est la traduction, en termes de calcul matriciel, de la composition à gauche d'un isomorphisme avec une application linéaire.

L'équivalence entre matrices et applications linéaires permet d'établir le résultat admis au chapitre deux :

Proposition 4.17 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée.*

1. *On a l'équivalence :*

$$A \text{ inversible} \Leftrightarrow \ker(A) = \{0\} \Leftrightarrow \operatorname{Im}(A) = \mathbb{K}^n .$$

2. *A est inversible s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = \mathrm{Id}_n$ ou telle que $BA = \mathrm{Id}_n$. Dans ces conditions $B = A^{-1}$.*

Démonstration. Notons f l'unique endomorphisme de $\mathbb{K}^n$ dont A est la matrice dans la base canonique.

1. Ce point est la conséquence du théorème **4.12** et de l'équivalence

$$\boldsymbol{y} = f(\boldsymbol{x}) \Leftrightarrow Y = AX.$$

2. Notons g l'unique endomorphisme de $\mathbb{K}^n$ dont B est la matrice dans la base canonique. L'égalité $AB = \mathrm{Id}_n$ se traduit par $f \circ g = I_{\mathbb{K}^n}$ donc f est surjective. De la même faon si $BA = \mathrm{Id}_n$, c'est que $g \circ f = I_{\mathbb{K}^n}$, donc que f est injective. Dans les deux cas on conclut, d'après **4.12**, que f est un isomorphisme dont g est la réciproque, ainsi A est inversible avec $B = A^{-1}$ □

Remarque : notons enfin que la proposition **4.11** permet d'affirmer que le rang d'une matrice est invariant en la multipliant, à gauche ou à droite, par une matrice inversible, notamment lors d'un échelonnement.

3. Systèmes linéaires

Nous disposons à présent des outils permettant de discuter en toute généralité de la résolution d'un système d'équations linéaires.

Soit $f : E \to E'$ une application linéaire. Une *équation linéaire* s'écrit :

$$f(\boldsymbol{x}) = \boldsymbol{b} \quad \text{où} \quad \boldsymbol{x} \in E \ , \ \ \boldsymbol{b} \in E'.$$

L'équation *homogène* associée est définie par :

$$f(\boldsymbol{x}) = 0_{E'} \ .$$

Résoudre une telle équation linéaire revient à déterminer l'ensemble des solutions :

$$S = \{ \ \boldsymbol{x} \in E \ , \ f(\boldsymbol{x}) = \boldsymbol{b} \ \} \ .$$

Remarquons que :

$$S \neq \emptyset \quad \Leftrightarrow \quad \boldsymbol{b} \in \mathrm{Im}(f) \ .$$

L'équation homogène admet le sous-espace $\ker(f)$ comme ensemble de solutions, qui possède au moins 0_E, donc n'est jamais vide.

Rappel de notation : Soit $x \in E$ et $A \subset E$, on note

$$x + A = \{ \ y \in E \ / \ \exists a \in A \ , \ y = x + a \ \} \ ,$$

appelé *translaté* de A par x.

Supposons donc que $S \neq \emptyset$, dans ces conditions :

Proposition 4.18 *Soit $\boldsymbol{x}_0$ une solution particulière de $f(\boldsymbol{x}) = \boldsymbol{b}$. On a :*

$$S = \boldsymbol{x}_0 + \ker(f) \ .$$

Démonstration. On a en effet les équivalences :

$$\boldsymbol{x} \in S \quad \Leftrightarrow \quad f(\boldsymbol{x}) = f(\boldsymbol{x}_0) \quad \Leftrightarrow \quad \boldsymbol{x} - \boldsymbol{x}_0 \in \ker(f) \ .$$

□

Les solutions s'obtiennent donc à partir de celles de l'équation homogène, translatées par une solution particulière.

On peut donc résumer les trois seuls cas possibles pour l'équation linéaire :

1. L'équation n'admet pas de solution. C'est le cas où $\boldsymbol{b} \notin \mathrm{Im}(f)$.
2. L'équation admet une unique solution. C'est le cas où $\boldsymbol{b} \in \mathrm{Im}(f)$ et $\ker(f) = \{0_E\}$ (f injective).
3. L'équation admet une infinité de solutions. C'est le cas où $\boldsymbol{b} \in \mathrm{Im}(f)$ et $\ker(f) \neq \{0_E\}$ (f non injective).

Considérons à présent le cas général d'un système de n équations linéaires à p inconnues :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \ldots + a_{np}x_p = b_n \end{cases} \quad \text{(SL).}$$

Il existe une unique application linéaire $f : \mathbb{K}^p \to \mathbb{K}^n$, dont la matrice dans les bases canoniques est égale à la matrice $A = [a_{ij}]$ de l'écriture matricielle du système. En posant $\boldsymbol{b} = (b_1, \ldots, b_n)$ et $\boldsymbol{x} = (x_1, \ldots, x_p)$, le système s'écrit donc $f(\boldsymbol{x}) = \boldsymbol{b}$. Le rang du système sera par définition celui de la matrice A , lui-même égal au rang de f .

L'énumération des cas possible est alors :

1. Le système n'admet pas de solution. C'est le cas où $\boldsymbol{b} \notin \mathrm{Im}(A)$, $\boldsymbol{b}$ n'appartient pas à l'espace des colonnes de A .
2. Le système admet un unique n -uplet de solutions $\boldsymbol{x} = (x_1, \ldots, x_p)$. C'est le cas où $\boldsymbol{b} \in \mathrm{Im}(A)$ et $\ker(A) = \{0_E\}$.
3. Le système admet une infinité de n -uplets de solutions. C'est le cas où $\boldsymbol{b} \in \mathrm{Im}(A)$ et $\ker(A) \neq \{0\}$. Ce dernier cas sera **toujours vrai si** $p > n$ et si $\boldsymbol{b} \in \mathrm{Im}(A)$, cas où il y a plus d'inconnues que d'équations : en effet, comme $\mathrm{rg}(A) \leq n$, si $n < p$ on aura $\dim(\ker(A)) = p - \mathrm{rg}(A) \geq p - n > 0$, le noyau est de dimension ≥ 1 .
4. Comme sous-cas du cas précédent : un système homogène avec $p > n$ aura toujours infinité de n -uplets de solutions puisque la condition $0 \in \mathrm{Im}(A)$ est nécessairement réalisée.

Dans le cas d'une infinité de solutions, donc de la forme $\boldsymbol{x}_0 + \ker(A)$ avec $\dim(\ker(A)) \geq 1$, il est utile d'exhiber une base de $\ker(A)$ pour exprimer les solutions. On est ramené à la méthode d'échelonnement décrite dans le chapitre précédent.

4. Changement de base

Définition 4.19
Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, alors $M_{\mathcal{B}',\mathcal{B}}(Id_E)$ s'appelle la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ de $\mathcal{B}$ à $\mathcal{B}'$.

Les colonnes de la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ sont constitués des coordonnées des éléments de $\mathcal{B}'$ dans la base $\mathcal{B}$. Par le Corollaire, on a $P_{\mathcal{B},\mathcal{B}'} = (P_{\mathcal{B}',\mathcal{B}})^{-1}$.

Théorème 4.20 *Soit E un $\mathbb{K}$-espace vectoriel, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $\boldsymbol{x} \in E$. Alors si X et X' désignent les composantes de $\boldsymbol{x}$ dans respectivement la base $\mathcal{B}$ et la base $\mathcal{B}'$ et si $P_{\mathcal{B},\mathcal{B}'}$ désigne la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, alors $X = P_{\mathcal{B},\mathcal{B}'}X'$.*

Théorème 4.21 *Soit E un $\mathbb{K}$-espace vectoriel, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $T \in L(E)$, alors si $P_{\mathcal{B},\mathcal{B}'}$ désigne la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, on a*

$$\boxed{M_{\mathcal{B}'}(T) = P_{\mathcal{B},\mathcal{B}'}^{-1} M_{\mathcal{B}}(T) P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}} M_{\mathcal{B}}(T) P_{\mathcal{B},\mathcal{B}'}.}$$

On a aussi

$$M_{\mathcal{B},\mathcal{B}'}(T) = P_{\mathcal{B},\mathcal{B}'}^{-1} M_{\mathcal{B}}(T) = P_{\mathcal{B}',\mathcal{B}} M_{\mathcal{B}}(T) \text{ et } M_{\mathcal{B},\mathcal{B}'}(T) = M_{\mathcal{B}'}(T) P_{\mathcal{B}',\mathcal{B}}.$$

$$\begin{tikzcd}
(E,\mathcal{B}') \arrow[r, "T"] \arrow[d, "Id_E"'] & (E,\mathcal{B}') & \Leftrightarrow & (E,\mathcal{B}') \arrow[r, "M_{\mathcal{B}'}(T)"] \arrow[d, "P_{\mathcal{B},\mathcal{B}'}"'] & (E,\mathcal{B}') \\
(E,\mathcal{B}) \arrow[r, "T"'] & (E,\mathcal{B}) \arrow[u, "Id_E"'] & & (E,\mathcal{B}) \arrow[r, "M_{\mathcal{B}}(T)"'] & (E,\mathcal{B}) \arrow[u, "P_{\mathcal{B},\mathcal{B}'}^{-1}=P_{\mathcal{B}',\mathcal{B}}"']
\end{tikzcd}$$

$$\begin{tikzcd}
(E,\mathcal{B}) \arrow[r, "T"] \arrow[d, "T"'] & (E,\mathcal{B}') & \Leftrightarrow & (E,\mathcal{B}) \arrow[r, "M_{\mathcal{B},\mathcal{B}'}(T)"] \arrow[d, "M_{\mathcal{B}}(T)"'] & (E,\mathcal{B}') \\
(E,\mathcal{B}) \arrow[ur, "Id_E"'] & & & (E,\mathcal{B}) \arrow[ur, "P_{\mathcal{B}',\mathcal{B}}"'] &
\end{tikzcd}$$

5. Applications linéaires remarquables

5.1 Homothéties

Définition 4.22
Pour $\lambda \in \mathbb{K}$, on appelle **homothétie** *de rapport λ l'endomorphisme $T_\lambda \in L(E)$ qui à tout $x \in E$ associe $T_\lambda(x) = \lambda x$.*

Proposition 4.23 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Pour toute base $\mathcal{B}$, la matrice d'une homothétie T_λ de rapport λ est donnée par*

$$\mathcal{M}_{\mathcal{B}}(T_\lambda) = \lambda \mathrm{Id}_n.$$

Démonstration. Il suffit d'écrire que pour tout élément $\boldsymbol{e}$ de la base $\mathcal{B}$, $T_\lambda(\boldsymbol{e}) = \lambda \boldsymbol{e}$. □

5.2 Projecteur

Définition 4.24
Soit E un $\mathbb{K}$-espace vectoriel. On dit que $p \in L(E)$ est un projecteur si $p \circ p = p$.

Si E est un ensemble et F un sous-ensemble de E, et si T est une application de E dans un ensemble G, on appelle restriction de T au sous ensemble F l'application de F dans G qui à $u \in F$ associe $T(u) \in G$. On note $T_{|_F}$ cette restriction. C'est donc la même application que T, sauf qu'on la considère maintenant sur un espace de départ plus petit.

Proposition 4.25 *Soit E un $\mathbb{K}$-espace vectoriel. Si $p \in L(E)$ est un projecteur alors*

- $E = \mathrm{Im}(p) \oplus \mathrm{Ker}(p)$.
- $p_{|_{\mathrm{Im}(p)}} = \mathrm{Id}_{\mathrm{Im}(p)}$*, où* $\mathrm{Id}_{\mathrm{Im}(p)}$ *désigne l'endomorphisme identité sur l'espace vectoriel* $\mathrm{Im}(p)$.

On dit aussi que p est une projection sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$:

$$\forall x \in E,\ x = x_{\mathrm{Im}} + x_{\mathrm{Ker}},\ \textit{avec}\ x_{\mathrm{Im}} \in \mathrm{Im}(p),\ x_{\mathrm{Ker}} \in \mathrm{Ker}(p)\ \textit{et}\ p(x) = x_{\mathrm{Im}}.$$

Démonstration. Soit $x \in E$ on peut écrire $x = p(x) + x - p(x)$. Il est clair que $p(x) \in \mathrm{Im}(p)$ et $x - p(x) \in \mathrm{Ker}(p)$ car $p(x - p(x)) = p(x) - p^2(x) = 0$. On en déduit que $\mathrm{Im}(p) + \mathrm{Ker}(p) = E.$ Soit $x \in \mathrm{Ker}(p) \cap \mathrm{Im}(p)$, on a $x = p(y)$ et $p(x) = 0$ ou encore $p(p(y)) = p(y) = 0 = x$. On en déduit que $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont supplémentaires dans E. De plus, si $x \in \mathrm{Im}(p)$, $x = p(y)$ et $p(x) = p^2(y) = p(y) = x.$ Donc $p_{|_{\mathrm{Im}(p)}} = \mathrm{Id}_{\mathrm{Im}(p)}$. □

Exemples dans $\mathbb{R}^3$

- Projection p sur une droite $\mathcal{D}$ de vecteur directeur $\boldsymbol{f}_1$: $\dim(\mathrm{Im}(T)) = 1$. Il existe alors deux vecteurs $\boldsymbol{f}_2, \boldsymbol{f}_3$ linéairement indépendants de $\mathrm{Ker}(p)$ tel que $\mathcal{B}' = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3)$ soit une base de $\mathbb{R}^3$. De plus

$$M_{\mathcal{B}'}(p) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Exemple : si $\boldsymbol{f}_1 = (1, -1, 0)$, $\boldsymbol{f}_2 = (0, 1, 1)$ $\boldsymbol{f}_3 = (1, 0, -1)$, la matrice de cette projection dans la base canonique est donnée par

$$M_{\mathcal{B}}(p) = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

En effet, $M_{\mathcal{B}}(p) = P_{\mathcal{B},\mathcal{B}'} M_{\mathcal{B}'}(p) P_{\mathcal{B}',\mathcal{B}}$ avec

$$P_{\mathcal{B},\mathcal{B}'} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \text{ et } P_{\mathcal{B}',\mathcal{B}} = \left(P_{\mathcal{B},\mathcal{B}'}\right)^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

- Projection p sur un plan $\mathcal{P}$ parallèlement à une droite $\mathcal{D}$. On note $(\boldsymbol{f}_1, \boldsymbol{f}_2)$ une base de $\mathcal{P}$ et $\boldsymbol{f}_3$ un vecteur directeur de $\mathcal{D}$. La famille $\mathcal{B}' = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3)$ est alors une base de $\mathbb{R}^3$ et

$$M_{\mathcal{B}'}(p) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Exemple : si $\boldsymbol{f}_1 = (1, -1, 0)$, $\boldsymbol{f}_2 = (0, 1, 1)$ $\boldsymbol{f}_3 = (1, 0, -1)$, la matrice de cette projection dans la base canonique est donnée par

$$M_{\mathcal{B}}(p) = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

5.3 Symétrie

Définition 4.26
Soit E un $\mathbb{K}$-espace vectoriel. On dit que $S \in L(E)$ est une symétrie si $S \circ S = \mathrm{Id}$.

Exemples dans $\mathbb{R}^3$

- Symétrie S par rapport à un plan $\mathcal{P}$ parallèlement à une droite $\mathcal{D}$. On note $(\boldsymbol{f}_1, \boldsymbol{f}_2)$ une base de $\mathcal{P}$ et $\boldsymbol{f}_3$ un vecteur directeur de $\mathcal{D}$. La famille $\mathcal{B}' = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3)$ est alors une base de $\mathbb{R}^3$ et
$$M_{\mathcal{B}'}(S) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$
Exemple : si $\boldsymbol{f}_1 = (1, -1, 0)$, $\boldsymbol{f}_2 = (0, 1, 1)$ $\boldsymbol{f}_3 = (1, 0, -1)$, la matrice de cette projection dans la base canonique est donnée par
$$M_{\mathcal{B}}(S) = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

- On remarquera que $-S$ est alors la symétrie par rapport à la droite $\mathcal{D}$ parallèlement au plan $\mathcal{P}$.

5.4 Formes linéaires et sous-espaces

5.5 Formes linéaires

Définition 4.27
Soit E un $\mathbb{K}$-espace vectoriel. On appelle forme linéaire sur E, une application de $L(E, \mathbb{K})$.

Exemples

1. $\varepsilon_i : E = \mathbb{K}^n \ni x = (x_1, \ldots, x_n) \mapsto x_i$.
2. $E = \mathcal{C}([a, b], \mathbb{R}) \ni f \mapsto \displaystyle\int_a^b f(t)dt$.
3. $val(x_0) : \mathcal{C}([a, b], \mathbb{R}) \ni f \mapsto f(x_0)$, avec x_0 fixé dans $[a, b]$.
4. $\mathcal{C}^1([a, b], \mathbb{R}) \ni f \mapsto f'(x_0)$, avec x_0 fixé dans $[a, b]$.
5. Soit E de dimension finie n et $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base fixée. On définit ε_i qui, tout $x \in E$, associe sa i^{eme} coordonnée x_i dans la base $\mathcal{B}$. Vérifier que ε_i est une forme linéaire.

Définition 4.28
Soit E un espace de dimension n. On appelle **hyperplan** *de E tout sous-espace de dimension $n - 1$.*

Dans $\mathbb{R}^2$ les hyperplans sont de dimension un, donc sont les droites passant par l'origine. Les hyperplans dans $\mathbb{R}^3$ sont de dimension deux, donc sont les plans passant par l'origine.

Remarque : dans un espace E de dimension n, le théorème de la base incomplète **3.19** et la proposition **3.36** font que les hyperplans sont constitués des sous-espaces H admettant une droite vectorielle $[\boldsymbol{a}]$ comme supplémentaire : $E = H \oplus [\boldsymbol{a}]$.

Proposition 4.29 *Soit E un $\mathbb{K}$-espace vectoriel et T une forme linéaire sur E. Soit $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E, il existe $(\alpha_1, \dots \alpha_n) \in \mathbb{K}^n$ tel que pour tout $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$ on ait*

$$T(\boldsymbol{x}) = \alpha_1 x_1 + \dots + \alpha_n x_n.$$

Démonstration. On note $\alpha_i = T(\boldsymbol{e}_i)$. On a donc

$$T(x) = T(x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n) = x_1T(\boldsymbol{e}_1) + \dots + x_nT(\boldsymbol{e}_n) = x_1\alpha_1 + \dots + x_n\alpha_n.$$

□

Cette dernière proposition rend simplement compte de ce qu'une forme linéaire est matriciellement représentée par une ligne. En choisissant $\mathcal{B}$ dans E et 1 comme base de $\mathbb{K}$, la matrice de T est :

$$\begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix}.$$

Notons que si $\mathbb{K} = \mathbb{R}$, et $E = \mathbb{R}^n$ muni de sa base canonique, $T(\boldsymbol{x})$ est le produit scalaire dans $\mathbb{R}^n$ des deux vecteurs $(\alpha_1, \dots, \alpha_n)$ et $(x_1, \dots, x_n)$.

Proposition 4.30 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et T une forme linéaire non nulle sur E. On a $\dim(\mathrm{Im}(T)) = 1$ et $\dim(\mathrm{Ker}(T)) = n - 1$.*

Démonstration. On sait que $\mathrm{Im}(T)$ est un sous-espace vectoriel de $\mathbb{K}$ non réduit à $\{0\}$. Par conséquent $\mathrm{Im}(T) = \mathbb{K}$. On en déduit que $\dim(\mathrm{Im}(T)) = 1$ et par le théorème du rang que $\dim(\mathrm{Ker}(T)) = n - 1$ □

Ainsi, le noyau d'une forme linéaire non nulle sur E est un hyperplan de E.

Réciproquement, si H est un hyperplan de E, il existe alors une forme linéaire non nulle sur E dont le noyau est H. En effet, considérons un supplémentaire donné par la remarque ci-dessus : $E = H \oplus [\boldsymbol{a}]$. On définit une base de E en complétant une base de H par $\boldsymbol{a}$ (cf. **3.35**), il existe alors une forme linéaire φ sur E dont la matrice est donnée par la ligne : $\begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix}$. Par construction on a $H = \ker(\varphi)$.

Montrer, à titre d'exercice, qu'étant donné une forme linéaire non nulle φ sur E, alors pour tout $\boldsymbol{a} \in E$ tel que $\varphi(\boldsymbol{a}) \neq 0$, on a $E = \ker(\varphi) \oplus [\boldsymbol{a}]$.

On peut résumer ce qui précède par l'énoncé :

Proposition 4.31 *Dans E de dimension finie, un sous-espace H est un hyperplan si et seulement s'il existe une forme linéaire de noyau H.*

Il n'est pas très difficile de généraliser cet énoncé pour des espaces non nécessairement de dimension finie.

Exemple : [Forme linéaire de $\mathbb{R}^3$] La forme linéaire définie par $T(x, y, z) = x + y + z$ peut aussi s'écrire $A\boldsymbol{x} = 0$ avec $A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ et $\boldsymbol{x} = (x, y, z)$ et on a déjà vu que le noyau de A est l'hyperplan d'équation $x + y + z = 0$. La forme linéaire T a donc aussi pour noyau le plan $\mathcal{P}$ d'équation $x + y + z = 0$.

Considérons à présent deux formes linéaires définies sur $\mathbb{K}^n$, identifiées à leurs matrices dans les bases canoniques :

$$A = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix} \quad , \quad B = \begin{bmatrix} \beta_1 & \dots & \beta_n \end{bmatrix}$$

Le sous-espace $F = \ker(A) \cap \ker(B)$ est défini par l'ensemble des $x = (x_1, \dots, x_n)$ vérifiant :

$$\begin{cases} \alpha_1 x_1 + \ldots + \alpha_n x_n = 0 \\ \text{et} \\ \beta_1 x_1 + \ldots + \beta_n x_n = 0 \end{cases}$$

Il est également égal au noyau de la matrice $M = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \end{bmatrix}$, qui est de rang un ou deux, ainsi $\dim(F) = n - 2$ si $\operatorname{rg}(M) = 2$.

Cette remarque se généralise :

Pour tout sous-espace F de E de dimension $n-p$, il existe p formes linéaires $\varphi_1, \dots, \varphi_p$, linéairement indépendantes dans $\mathcal{L}(E, \mathbb{K})$, telles que :

$$F = \ker(\varphi_1) \cap \ldots \cap \ker(\varphi_p) ,$$

ou bien encore, F est formé des n-uplets vérifiants :

$$\begin{cases} \alpha_{11} x_1 + \ldots + \alpha_{1n} x_n = 0 \\ \vdots \\ \alpha_{p1} x_1 + \ldots + \alpha_{p1} x_n = 0 \end{cases} .$$

Cette écriture constitue *l'équation cartésienne* de F.

Exemple L'équation cartésienne d'une droite vectorielle dans $\mathbb{R}^3$ est donnée par un système linéaire :

$$\begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$$

dans lequel (a, b, c) et (a', b', c') ne sont pas colinéaires.

Chapitre 5

Déterminants

On suppose que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

1. Définition et propriétés

Définition 5.1
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ *, on appelle* **déterminant** *de la matrice* A *un élément de* $\mathbb{K}$ *, noté* $\det(A)$ *que l'on définit par récurrence de la façon suivante :*

- *Si* $n = 1$ *et* $A = (a)$ *, alors* $\det(A) = a$ *.*
- *Si* $n \geq 2$ *, alors*

$$\det(A) = a_{11}\Delta_{11} - a_{21}\Delta_{21} + \cdots + (-1)^{n+1}a_{n1}\Delta_{n1},$$

où $\Delta_{i,j}$ *est le déterminant de la matrice* $A_{i,j} \in \mathcal{M}_{n-1}(\mathbb{K})$ *obtenue à partir de la matrice* A *en rayant la* $i^{\text{ème}}$ *ligne et la* $j^{\text{ème}}$ *colonne.*

Remarque : On note généralement pour $A = [a_{i,j}] \in \mathcal{M}_n(\mathbb{K})$

$$\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

On ne définit le déterminant que pour des matrices carrées.

1.1 Propriétés élémentaires des déterminants

Proposition 5.2 *Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice dont une ligne ou une colonne est identiquement nulle alors son déterminant est nul.*

Démonstration. On montre le résultat par récurrence sur la taille de la matrice. Supposons que la ligne L_{i_0} est nulle. On vérifie que les matrices A_{i1} ont une ligne nulle si $i \neq i_0$. La récurrence assure que $\Delta_{i1} = 0$ pour $i \neq i_0$ et comme $a_{i_0 1} = 0$, on en déduit que $\det(A) = 0$.

Supposons maintenant que la colonne C_j de la matrice A soit identiquement nulle. Si $j = 1$ alors la définition du déterminant assure directement que $\det(A) = 0$. Si $j \neq 1$, alors toutes les matrices A_{i1} ont une colonne nulle et donc $\Delta_{i1} = 0$. □

Proposition 5.3 *Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice triangulaire supérieure ou inférieure, alors le déterminant de A est égal au produit des coefficients diagonaux de la matrice.*

Démonstration. On montre le résultat par récurrence sur la taille de la matrice. Lorsque $n = 2$ le résultat est immédiat. Supposons le résultat vrai pour toute matrice triangulaire de taille $n - 1$.

Si A est une matrice triangulaire supérieure de taille n. Comme $a_{i1} = 0$ pour tout $i \geq 2$ on a $\det(A) = a_{11}\Delta_{11}$ avec Δ_{11} déterminant d'une matrice triangulaire supérieure de taille $n-1$ dont les coefficients diagonaux sont les coefficients diagonaux de A, $a_{i,i}$ pour $i \geq 2$. D'où le résultat.

Si la matrice A est triangulaire inférieure, la matrice A_{11} est une matrice triangulaire inférieure, l'hypothèse de récurrence assure que $\Delta_{11} = \prod_{i=2}^{n} a_{i,i}$. Si $i > 1$, alors la matrice A_{i1} a une ligne identiquement nulle, par suite son déterminant Δ_{i1} est nul (Proposition 5.2). Le résultat suit immédiatement. □

Corollaire 5.4 *Le déterminant de la matrice $D_i(a)$ est égal à a et si $i \neq j$, le déterminant de $T_{i,j}(\lambda)$ est égal à 1.*

Proposition 5.5 *Soit $\lambda \in \mathbb{K}$. Si on multiplie par λ une ligne d'une matrice de $\mathcal{M}_n(\mathbb{K})$ alors son déterminant est multiplié par λ.*

Démonstration. On montre le résultat par récurrence sur la taille de la matrice. Pour $n = 2$ le résultat est immédiat. On suppose le résultat vrai pour toute matrice de taille $n - 1$. Soit A une matrice de taille n et $A_\lambda = D_i(\lambda)A$. On utilise la définition du déterminant pour calculer $\det(D_i(\lambda)A)$. On a

$$\begin{aligned}\det(D_i(\lambda)A) = a_{11}\Delta_{11}^{\lambda} - a_{21}\Delta_{21}^{\lambda} + \cdots + (-1)^i a_{i-11}\Delta_{i-11}^{\lambda} \\ + (-1)^{i+1}\lambda a_{i1}\Delta_{i1}^{\lambda} + (-1)^{i+2}a_{i+11}\Delta_{i+11}^{\lambda} + \cdots + (-1)^{n+1}a_{n1}\Delta_{n1}^{\lambda}.\end{aligned}$$

Remarquons que $\Delta_{i1}^{\lambda} = \Delta_{i1}$ et que $\Delta_{j1}^{\lambda} = \lambda\Delta_{j1}$ si $j \neq i$ en utilisant l'hypothèse de réccurence. On en déduit que

$$\det(D_i(\lambda)A) = \lambda \det(A). \quad \square$$

Proposition 5.6 *On considère trois matrices (écrites en juxtaposant leurs lignes) :*

$$A = \begin{bmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L'_i + L''_i \\ L_{i+1} \\ \vdots \\ L_n \end{bmatrix} \quad , \; A' = \begin{bmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L'_i \\ L_{i+1} \\ \vdots \\ L_n \end{bmatrix} \quad , \; A'' = \begin{bmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L''_i \\ L_{i+1} \\ \vdots \\ L_n \end{bmatrix} \quad .$$

On a alors :

$$\det(A) = \det(A') + \det(A'') \, .$$

Démonstration. On montre le résultat par récurrence sur la taille de la matrice. Le résultat est vrai pour $n = 2$. On a

$$\det(A) \quad = \quad a_{11}\Delta_{11}^{A} - a_{21}\Delta_{21}^{A} + \cdots + (-1)^{i_0+1}a_{i_0 1}\Delta_{i_0 1}^{A} + \cdots + (-1)^{n+1}a_{n1}\Delta_{n1}^{A}$$

Or $a_{i_0 1} = a'_{i_0 1} + a''_{i_0 1}$, $\Delta_{i_0 1}^{A} = \Delta_{i_0 1}^{A'} = \Delta_{i_0 1}^{A''}$. De plus si $i \neq i_0$, en utilisant la récurrence on a

$$\Delta_{i1}^{A} = \Delta_{i1}^{A'} + \Delta_{i1}^{A''}.$$

On en déduit immédiatement que

$$\det(A) = \det(A') + \det(A''). \quad \square$$

Proposition 5.7 *Soit $A \in M_n(\mathbb{K})$. On définit $A' \in M_n(\mathbb{K})$, obtenue en échangeant deux lignes de A . Alors*

$$\det(A) = -\det(A').$$

Démonstration. On montre le résultat par récurrence sur la dimension de la matrice.

$$\det(A') = a_{11}\Delta_{11}^{A'} - a_{21}\Delta_{21}^{A'} + \cdots + (-1)^{i_0+1}a_{j_0 1}\Delta_{i_0 1}^{A'} + \cdots \\ + (-1)^{j_0+1}a_{i_0 1}\Delta_{j_0 1}^{A'} + \cdots + (-1)^{n+1}a_{n1}\Delta_{n1}^{A}$$

Par récurrence, on a si $j \neq i_0, j_0$, $\Delta_{j1}^{A'} = -\Delta_{j1}^{A}$ et $\Delta_{i_0 1}^{A'}$ est le déterminant de la matrice extraite de A lorsque on enlève la première colonne et la j_0 ème ligne et dont la ligne i_0

est à la place de la ligne j_0. Pour ramener cette ligne à sa position il faut alors faire j_0-i_0+1 échange de lignes. Par conséquent, $\Delta_{i_0}^{A'} = (-1)^{j_0-i_0+1}\Delta_{j_01}^{A}$, $\Delta_{j_0}^{A'} = (-1)^{i_0-j_0+1}\Delta_{i_01}^{A}$. On en déduit le résultat. □

Corollaire 5.8 *On suppose que $n \geq 2$. Si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ a deux lignes identiques alors son déterminant est nul.*

Démonstration. Si une matrice A a deux lignes L_{i_0} et L_{j_0} identiques, alors la matrice A' obtenue en échangeant ces deux lignes est égale à la matrice A. Mais d'après la proposition précédente, on a $\det(A) = -\det(A') = -\det(A)$, par suite on en déduit que $\det(A) = 0$. □

Proposition 5.9 *On suppose que $n \geq 2$. Soit $\lambda \in \mathbb{K}$ et $A, A' \in \mathcal{M}_n(\mathbb{K})$. On suppose qu'il existe i_0 et j_0 tel que $i_0 \neq j_0$*

$$L_i = L'_i, \forall i \neq i_0 \text{ et } L'_{i_0} = L_{i_0} + \lambda L_{j_0}$$

alors

$$\det(A) = \det(A').$$

Autrement dit, le déterminant de A ne change pas si on ajoute à une ligne de A le produit par λ d'une autre ligne de A.

Démonstration. On considère A'' la matrice telle que pour tout $i \neq i_0$, $L''_i = L_i$ et $L''_{i_0} = L_{j_0}$. D'après le corollaire précédent, on a $\det(A'') = 0$. On multiplie maintenant cette ligne L''_{i_0} par λ, la matrice obtenue $A^{(3)}$ est toujours de déterminant nul (cf Proposition 5.8). On utilise maintenant la proposition 5.6 pour les matrices A, A' et $A^{(3)}$ et on obtient que $\det(A) = \det(A') + \det(A^{(3)}) = \det(A')$. □

Théorème 5.10 *Si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$ alors*

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

Démonstration.

• **Première étape.** Notons que ce résultat est démontré si A est une matrice de dilatation (Proposition 5.4) ou si A est une matrice de transvection (Proposition 5.4). Par conséquent le résultat est également démontré si A est un produit de matrices de dilatation ou de transvection.

• **Deuxième étape.** Notons que si A est une matrice inversible alors A est un produit de matrices élémentaires. Donc si A est inversible $\det(AB) = \det(A)\det(B)$.

• **Troisième étape.** Supposons maintenant que la matrice A soit non inversible. Il existe une matrice E inversible telle que EA soit échelonnée, en particulier triangulaire supérieure. Or EA n'est pas inversible. Donc la dernière ligne de la matrice EA est une ligne de zéros. On déduit de la Proposition 5.2 que $\det(EA) = 0$ et d'après l'étape précédente, comme E est inversible, $\det(A) = \det(E^{-1}EA) = \det(E^{-1})\det(A) = 0$. Remarquons alors que la dernière ligne de la matrice EAB est nulle de sorte (Proposition 5.2) que $\det(EAB) = 0$. A nouveau comme P est inversible on sait que $\det(E^{-1}EAB) = \det(E^{-1})\det(EAB)$, on en déduit que $\det(AB) = 0$ et donc que l'on a bien $\det(AB) = \det(A)\det(B) = 0$. □

Théorème 5.11 *Si A est une matrice de $\mathcal{M}_n(\mathbb{K})$ inversible alors*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Démonstration. On utilise le résultat précédent. On a

$$\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(\mathrm{I}d) = 1.$$

Par suite

$$\det(A^{-1}) = \frac{1}{\det(A)}. \quad \square$$

Théorème 5.12 *Si A est une matrice de $\mathcal{M}_n(\mathbb{K})$ alors $\det(A^t) = \det(A)$.*

Démonstration. On remarque que le résultat est vrai pour les matrices élémentaires et pour les matrices triangulaires. Pour une matrice A quelconque on utilise une forme échelonnée $A' = PA$. On a $\det(A') = \det((A')^t)$ car A' est triangulaire supérieure et $\det(P^{-1}) = \det((P^{(-1)})^t)$ car P est un produit de matrices élémentaires. On en déduit que

$$\det(A^t) = \det((A')^t(P^{-1})^t) = \\ \det((A')^t)\det((P^{-1})^t) = \det(A')\det(P^{-1}) = \det(P^{-1}A') = \det(A). \quad \square$$

Corollaire 5.13 *Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, on a pour tout $j = 1, \dots, n$*

$$\det(A) = (-1)^{j+1}\left(a_{1j}\Delta_{1j} - a_{2j}\Delta_{2j} + \cdots + (-1)^{n+1}a_{nj}\Delta_{nj}\right). \tag{5.17}$$

et pour tout $i = 1, \dots, n$

$$\det(A) = (-1)^{i+1}\left(a_{i1}\Delta_{i1} - a_{i2}\Delta_{i2} + \cdots + (-1)^{n+1}a_{in}\Delta_{in}\right), \tag{5.18}$$

Autrement dit, on peut développer le déterminant par rapport à n'importe quelle ligne ou colonne.

Démonstration. La première équation découle de la Proposition 5.6, tandis que la deuxième équation est une conséquence du théorème 5.12. □

Définition 5.14
Soit $A \in \mathcal{M}_n(\mathbb{K})$ *. Le déterminant* $(-1)^{i+j}\Delta_{i,j}$ *est appelé* **cofacteur** *de* $a_{i,j}$ *. La* **comatrice de** A *, notée* $\mathrm{Com}(A)$ *est la matrice de* $\mathcal{M}_n(\mathbb{K})$ *dont le coefficient* i, j *est le cofacteur* $(-1)^{i+j}\Delta_{i,j}$ *.*

On a en fait pour tout $i = 1, \dots, n$

$$\det(A) = (-1)^{i+1}\Delta_{i1}\, a_{i1} + (-1)^{i+2}\Delta_{i2}\, a_{i2} + \cdots + (-1)^{i+n}\Delta_{in}\, a_{in},$$

et pour tout $j = 1, \dots, n$

$$\det(A) = (-1)^{j+1}\Delta_{1j}\, a_{1j} + (-1)^{j+2}\Delta_{2j}\, a_{2j} + \cdots + (-1)^{j+n}\Delta_{nj}\, a_{nj}.$$

2. Usages

2.1 Application à l'inversibilité

Proposition 5.15 *Soit* $n \geq 2$ *et* $A \in \mathcal{M}_n(\mathbb{K})$ *. Alors on a*

$$(Com(A))^t A = A(Com(A))^t = \det(A)\mathrm{Id}_n.$$

La conséquence en est que la matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si $\det(A) \neq 0$ et, dans ce cas, on a :

$$A^{-1} = \frac{1}{\det(A)}(\mathrm{Com}(A))^t.$$

Démonstration. On a

$$((\mathrm{Com}(A))^t A)_{i,j} = (-1)^{i+1}\Delta_{1i}\, a_{1j} + (-1)^{i+2}\Delta_{2i}\, a_{2j} + \cdots + (-1)^{i+n}\Delta_{ni}\, a_{nj}$$

En utilisant la formule 5.18, on déduit que $((\mathrm{Com}(A))^t A)_{i,i} = \det(A)$. De plus si $i \neq j$ $((\mathrm{Com}(A))^t A)_{i,j}$ est le déterminant de la matrice obtenue à partir de la matrice A en remplaçant la colonne i par la colonne j. La matrice ainsi obtenue a deux colonnes identiques, son déterminant est donc nul. Par conséquent si $i \neq j$ $((\mathrm{Com}(A))^t A)_{i,j} = 0$. On a donc bien montré que $(\mathrm{Com}(A))^t A = A(\mathrm{Com}(A))^t = \det(A)\mathrm{Id}_n$. $\square$

Proposition 5.16 *Soient* E *et* F *deux* $\mathbb{K}$*-espaces vectoriels de dimension* n*. On se donne* $\mathcal{B} = \{bsbe_1, \dots, \boldsymbol{e}_n\}$ *une base de* E *et* $\mathcal{B}' = \{\varepsilon_1, \dots, \varepsilon_n\}$ *une base de* F*. Soit* $f \in \mathcal{L}(E, F)$*. Alors* f *est un isomorphisme de* E *sur* F *si et seulement si*

$$\det(\mathcal{M}_{\mathcal{B},\mathcal{B}'}(f)) \neq 0.$$

On a vu dans le Corollaire coro 411 que f est un isomorphisme si et seulement si la matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}'}(f)$ est inversible. On utilise ensuite la Proposition 5.15.

2.2 Résolution de systèmes linéaires et formules de Cramer

Soit $n \geq 2$ et $A \in \mathcal{M}_n(\mathbb{K})$ **inversible** et $b \in \mathcal{M}_{n,1}(\mathbb{K})$. La solution $X = (x_i) \in \mathcal{M}_{n,1}(\mathbb{K})$ du système linéaire $AX = b$ est donnée par $X = A^{-1}b$ ou encore

$$X = \frac{1}{\det(A)}(\mathrm{Com}(A))^t b,$$

ou composantes par composantes

$$\begin{aligned} x_i &= \frac{1}{\det(A)}\big((-1)^{i+1}\Delta_{1i}\, b_1 + (-1)^{i+2}\Delta_{2i}\, b_2 + \cdots + (-1)^{i+n}\Delta_{ni}\, b_n\big) \\ &= \frac{1}{\det(A)}\begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn} \end{vmatrix} \end{aligned}$$

C'est ce que l'on appelle les **formules de Cramer**.

• On utilise ces formules pour $n = 2$ ou $n = 3$. Pour $n \geq 4$, il est plus rapide d'utiliser l'algorithme du pivot de Gauss.

2.3 Quelques calculs de déterminants

Le polynôme caratéristique d'une matrice

Proposition 5.17 *On définit pour une matrice* $A \in \mathcal{M}_n(\mathbb{R})$ *,*

$$P_A(X) = \det(A - X\mathrm{Id}_n).$$

C'est un polynôme de degré n *appelé polynôme caractéristique.*

Il suffit de remarquer par réccurence que le déterminant d'une matrice de taille n est une somme de produit de n termes distincts de la matrice. Par conséquent c'est bien un polynôme en X. On montre alors par réccurence que son terme dominant est $(-1)^n X^n$.

Le déterminant de la matrice de Frobenius

Proposition 5.18 *Le polynôme*

$$P(X) = (-1)^{n+1}\left(X^{n+1} - a_n X^n - \cdots - a_1 - a_0\right)$$

est le polynôme caractéristique de la matrice de Frobenius de taille $n+1$

$$A = \begin{bmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & \ddots & & & \vdots & a_2 \\ 0 & 1 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & & 0 & 1 & 0 & a_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & a_n \end{bmatrix}$$

Démonstration. On doit calculer

$$\Delta = \begin{vmatrix} -X & 0 & \cdots & \cdots & \cdots & 0 & a_0 \\ 1 & -X & \ddots & & & \vdots & a_1 \\ 0 & 1 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & & 0 & 1 & -X & a_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & a_n - X \end{vmatrix}$$

Pour cela on va ajouter à la première ligne une combinaison linéaire des précédentes :

$$L_1 \to L_1 + XL_2 + X^2L_3 + \cdots + X^{n-1}L_{n-1} + X^nL_n$$

On obtient

$$\Delta = \begin{vmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (a_n - X)X^n \\ 1 & -X & \ddots & & & \vdots & a_1 \\ 0 & 1 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & & 0 & 1 & -X & a_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & a_n - X \end{vmatrix}$$

En développant le déterminant obtenu par rapport à la première ligne, on a

$$\Delta = (-1)^{n+2}(a_0 + a_1 X + \cdots + a_{n-1}X^{n-1} + (a_n - X)X^n) \begin{vmatrix} 1 & -X & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 & -X \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{vmatrix}$$

Par suite on a bien $\Delta = P(X)$. $\square$

Exemple 5.19

Déterminant de Van der Monde

On s'intéresse pour quatre réels distincts x_1, x_2, x_3, x_4 au déterminant suivant :

$$\Delta(x_1, x_2, x_3, x_4) = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & (x_4)^2 \\ (x_1)^3 & (x_2)^3 & (x_3)^3 & (x_4)^3 \end{vmatrix}$$

En developpant ce déterminant par rapport à la première colonne on remarque que $\Delta(x_1, x_2, x_3, x_4)$ est un polynôme de degré 3 en la variable x_1 . De plus si $x_1 = x_2$ ou $x_1 = x_3$ ou $x_1 = x_4$, c'est le déterminant d'une matrice qui a deux colonnes identiques et par conséquent $\Delta(x_i, x_2, x_3, x_4) = 0$ pour $i = 2, 3, 4$. On en déduit que x_2, x_3, x_4 sont les racines de ce polynôme de degré 3 et qu'il existe une fonction $\delta(x_2, x_3, x_4)$ telle que

$$\Delta(x_1, x_2, x_3, x_4) = \delta(x_2, x_3, x_4)(x_1 - x_2)(x_1 - x_3)(x_1 - x_4).$$

Si on développe par rapport à la première colonne ce déterminant on vérifie que

$$\delta(x_2, x_3, x_4) = - \begin{vmatrix} 1 & 1 & 1 \\ x_2 & x_3 & x_4 \\ (x_2)^2 & (x_3)^2 & (x_4)^2. \end{vmatrix}$$

On recommence le même raisonnement

$$\delta(x_2, x_3, x_4) = - \begin{vmatrix} 1 & 1 \\ x_3 & x_4 \end{vmatrix} (x_2 - x_3)(x_2 - x_4) = (x_3 - x_4)(x_2 - x_3)(x_2 - x_4).$$

Autrement dit

$$\boxed{\Delta(x_1, x_2, x_3, x_4) = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & (x_4)^2 \\ (x_1)^3 & (x_2)^3 & (x_3)^3 & (x_4)^3 \end{vmatrix} = \prod_{1 \leq i \leq 4 j > i} (x_i - x_j).}$$

Ce résultat se généralise par récurrence à la dimension n.

Exemple 5.20

Déterminant d'une matrice tridiagonale On s'intéresse au déterminant de la matrice tridiagonale suivante

$$\begin{bmatrix} a & b & 0 & 0 & 0 \\ b & a & b & 0 & 0 \\ 0 & b & a & b & 0 \\ 0 & 0 & b & a & b \\ 0 & 0 & 0 & b & a \end{bmatrix}$$

On a en développant par rapport à la première colonne

$$\begin{aligned} \Delta_{a,b}^5 = \begin{vmatrix} a & b & 0 & 0 & 0 \\ b & a & b & 0 & 0 \\ 0 & b & a & b & 0 \\ 0 & 0 & b & a & b \\ 0 & 0 & 0 & b & a \end{vmatrix} &= a \begin{vmatrix} a & b & 0 & 0 \\ b & a & b & 0 \\ 0 & b & a & b \\ 0 & 0 & b & a \end{vmatrix} - b \begin{vmatrix} b & 0 & 0 & 0 \\ b & a & b & 0 \\ 0 & b & a & b \\ 0 & 0 & b & a \end{vmatrix} \\ &= a \begin{vmatrix} a & b & 0 & 0 \\ b & a & b & 0 \\ 0 & b & a & b \\ 0 & 0 & b & a \end{vmatrix} - b^2 \begin{vmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{vmatrix} \\ &= a\Delta_{a,b}^4 - b^2\Delta_{a,b}^3 \end{aligned}$$

De même

$$\begin{aligned} \Delta_{a,b}^4 &= a\Delta_{a,b}^3 - b^2\Delta_{a,b}^2 \\ \Delta_{a,b}^3 &= a\Delta_{a,b}^2 - b^2\Delta_{a,b}^1 \\ \Delta_{a,b}^2 &= a^2 - b^2 \\ \Delta_{a,b}^1 &= a \end{aligned}$$

On en déduit que

$$\begin{aligned} \Delta_{a,b}^3 &= a(a^2 - b^2) - b^2 a = a^3 - 2b^2 a \\ \Delta_{a,b}^4 &= a(a^3 - 2b^2 a) - b^2(a^2 - b^2) = a^4 - 3b^2a^2 + b^4 \\ \Delta_{a,b}^5 &= a(a^4 - 3b^2a^2 + b^4) - b^2(a^3 - 2b^2 a) = a^5 - 4b^2a^3 + 3ab^4 \end{aligned}$$

Le cas classique est $a = -2b$, on trouve alors

$$\begin{aligned} \Delta_{a,b}^3 &= = -4b^3 \\ \Delta_{a,b}^4 &= a(a^3 - 2b^2 a) - b^2(a^2 - b^2) = 5b^4 \\ \Delta_{a,b}^5 &= a(a^4 - 3b^2a^2 + b^4) - b^2(a^3 - 2b^2 a) = -6b^5 \end{aligned}$$