Algèbre linéaire 1

Christophe Reutenauer
Laboratoire de combinatoire et d'informatique mathématique,
Université du Québec à Montréal

December 3, 2017

Contents

Remerciements: Benjamin Blanchette, Christophe Hohlweg.

1 Introduction

Nous notons $\mathbb{R}$ l'ensemble des nombres réels.

1.1 Système d'équations linéaires: méthode d'élimination des variables, ou de substitution

Cette méthode est peut-être la plus naturelle. On choisit une des variables du système, on l'exprime en fonction des autres en utilisant une des équations; puis on remplace dans toutes les autres équations cette variable par l'expression obtenue. Cela donne un nouveau système, avec moins de variables, et moins d'équations (il se peut que des équations soient identiques, auquel cas on supprime les équations redondantes). On continue jusqu'à obtenir des variables sujettes à aucune équation: elles sont *libres.* En remontant le calcul à l'envers, on exprime toutes les variables non libres en fonctions des variables libres; celle-ci peuvent prendre des valeurs arbitraires, et ça donne pour chaque choix une solution du système. Il y a lors une infinité de solutions, qui sont paramétrisées par les variables libres.

Il peut arriver qu'on ne trouve aucune variable libre: il y aura alors une solution unique.

Il peut arriver aussi qu'on arrive à une équation contradictoire (comme 0=1), et dans ce cas le système n'a aucune solution.

Ces trois alternatives, qui sont les seules possibles, sont illustrées dans les trois exemples très simples suivants.

Exemple 1.1.

$$a + 2b - c + 3d = 0,\ 2a - b - c - 2d = 0,\ 3a + b - 2c + d = 0.$$

Exprimons c en fonction des autres variables, en utilisant la première équation: $c = a + 2b + 3d$. Remplaçons c par l'expression $a + 2b + 3d$ dans les autres équations: $2a-b-a-2b-3d-2d = 0$, $3a+b-2a-4b-6d+d = 0$. Ces deux équations se simplifient toutes les deux en la même équation $a - 3b - 5d = 0$. Celle-ci permet d'exprimer a: $a = 3b + 5d$. Il n'y a plus d'autres équations; les variables libres sont b et d.

On remonte: $c = a + 2b + 3d = 3b + 5d + 2b + 3d = 5b + 8d$. On obtient donc

$$a = 3b + 5d, c = 5b + 8d,$$

avec de valeurs arbitraires pour b et d. Il y a une infinité de solutions.

Exemple 1.2.

$$x + 3y = 8, 3x + y = 0.$$

Exprimons y en fonction de x avec la deuxième équation: $y = -3x$. Reportons y dans la première: $x + 3(-3x) = 8$, c'est-à-dire $-8x = 8$. Donc $x = -1$ et en remontant: $y = -3(-1) = 3$. Solution unique:

$$x = -1, y = 3.$$

Exemple 1.3.

$$u - v = 1, u + v + w = 2, 2u + w = 0.$$

Avec la première équation $u = v + 1$. Reportons u dans les autres: $v + 1 + v + w = 2$, $2(v + 1) + w = 0$. Ça se réécrit $2v + w = -1$, $2v + w = -2$. On paut déjà voir qu'il n'y aura pas de solutions. Mais soyons plus systématique en exprimant w avec la deuxième équation: $w = -2v - 1$. Reportons dans la dernière équation: $2v + (-2v - 1) = -2$, ce qui se réécrit en $-1 = -2$. Pas de solution.

Exercice 1.1. *Résoudre:* $x + 2y = 4, 3x + 7y = 2$.

Exercice 1.2. *Résoudre:* $2x + 3y - 2z = 7, x - y - z = 1, 3x + 2y - 3z = 8$.

Exercice 1.3. *Résoudre:* $2a + 3b + 4c = 4, a - b + c = 1, 3a + 2b + 5c = 8$.

Exercice 1.4. *Résoudre:* $i + j + k + l = 0, i + j + k - l = 4, i + j - k + l = -4, i - j + k + l = 2$.

2 Matrices

2.1 Définitions (rappels)

Nous notons $M_{np}(\mathbb{R})$ l'ensemble des matrices de taille $n \times p$ sur $\mathbb{R}$. Une telle matrice a donc n lignes et p colonnes. On note une telle matrice $[a_{ij}]_{1 \leq i \leq n, 1 \leq j \leq p}$, où a_{ij} désigne le coefficient (ou élément) en position i, j (ligne i, colonne j) de la matrice. On note la matrice aussi plus simplement $[a_{ij}]$ si aucune confusion n'est à craindre.

Une matrice-ligne (resp. matrice-colonne) est un élément de $M_{1p}(\mathbb{R})$ (resp. $M_{n1}(\mathbb{R})$). Nous écrirons aussi $\mathbb{R}^p$ pour $M_{1p}(\mathbb{R})$. Les lignes d'une matrice sont des matrices-lignes. Les colonnes d'une matrice sont des matrices-colonnes.

Deux matrices sont égales si elles ont la même taille et si pour tous i, j leur élément en position i, j sont égaux.

La matrice nulle de taille $n \times p$ est la matrice dont tous les coefficients sont nuls. On la note 0_{np} et plus souvent simplement 0.

Une matrice carrée d'ordre n est un élément de $M_n(\mathbb{R}) = M_{nn}(\mathbb{R})$. Les éléments diagonaux d'une matrice carrée $[a_{ij}]$ sont les éléments a_{ii}. On appelle diagonale d'une matrice les positions des éléments diagonaux. Une matrice carrée $[aiij$ est dite trianagulaire supérieure (resp inférieur) si pour tous i, j, $i > j \Rightarrow a_{ij} = 0$ (resp. $i < j \Rightarrow a_{ij} = 0$). Une matrice diagonale est une matrice qui est triangulaire à la fois supérieure et inférieurs; autrement dit, les éléments d'une matrice diagonale, qui sont en dehors de la diagonale, sont nuls.

Ce vocabulaire est très intuitif: voici des exemples de matrice triangulaire supérieure, triangulaire inférieure, diagonale (de gauche à droite):

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}, \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}, \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

On note I_n (ou simplement I) la matrice carrée d'ordre n, qui est diagonale, avec des des 1 comme éléments diagonaux. On l'appelle la matrice identité d'ordre n.

La *transposée* d'une matrice $A = [a_{ij}]$, de taille $n \times p$, est la matrice, notée A^t, de taille $p \times n$, telle que son élément en ligne i, colonne j, est l'élément de A en colonne j, ligne i. Ecrivant $A^t = [b_{ij}]$, ceci s'exprime par

$$b_{ij} = a_{ji}.$$

2.2 Matrices: somme et produit externe

La somme de deux matrices $[a_{ij}]$ et $[b_{ij}$ de même taille est la matrice $[c_{ij}]$ dont les coefficients sont définis par $c_{ij} = a_{ij} + b_{ij}$. Le *produit externe* d'un réel r par la matrice $[a_{ij}]$ est la matrice $[a_{ij}]$ dont les coefficients sont définis par $a_{ij} = ra_{ij}$. Par exemple

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} a+p & b+q \\ c+r & d+s \end{bmatrix}, \ r \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}.$$

Ces deux opérations jouissent des propriétés suivantes: quelles que soient les matrices de même taille X, Y, Z et les réels a, b:

1. $X + Y = Y + X$ (commutativité de la somme);
2. $(X + Y) + Z = X + (Y + Z)$ (associativité de la somme);
3. $X + 0 = 0$ (0 est élément neutre pour l'addition);
4. Il existe une matrice X' telle que $X + X' = 0$ (existence de la matrice opposée);
5. $1x = x$ (le produit externe de $1 \in R$ avec X est égal à X);
6. $(a + b)X = aX + bX$ (distributivité à gauche du produit externe sur l'addition dans $\mathbb{R}$);
7. $a(X + Y) = aX + aY$ (distributivité à droite du produit externe sur l'addition des matrices);

8. $(ab)X = a(bX)$ (associativité).

L'associativité de la somme permet d'omettre les parenthèses: on écrit $A + B + C$ au lieu de $(A + B) + C$. Plus généralement, $A_1 + \ldots + A_k$ désigne la somme des ces k matrices (de même taille!) avec n'importe quel parenthésage.

Soient $A_1, \ldots, A_k \in M_{np}(\mathbb{R})$ et $a_1, \ldots, a_k \in R$. On appelle $a_1 A_1 + \cdots + a_k A_k$ une *combinaison linéaire* des matrices $A_1, \ldots A_k$. Les a_i s'appellent les *coefficients* de cette combinaison linéaire; plus précisément a_i est le coefficient de A_i.

Exercice 2.1. *Calculer la combinaison linéaire dans $M_2(\mathbb{R})$:*

$$x \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{y+z}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{z-y}{2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

2.3 Produit de matrices

Soient $A = [a_i j], B = [b_{ij}]$ des matrices de taille respectives $n \times p$ et $p \times q$ (le nombre de colonnes de A est égal au nombre de lignes de B). Le produit AB de ces deux matrices est la matrice $C = [c_{ij}]$, de taille $n \times q$ (le nombre de lignes de C est égal au nombre de lignes de A et son nombre de colonnes à celui de B), est défini par:

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

Notez que cela s'exprime aussi comme suit:

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ip} b_{pj}.$$

Le produit des matrices jouit des propriétés suivantes: soient $A, A' \in M_{np}(\mathbb{R})$, $B \in M_{pq}(\mathbb{R})$, $C \in M_{qr}(\mathbb{R})$, on a:

1. $(AB)C = A(BC)$ (associativité du produit)
2. $(A + A')B = AB + A'C$ (distributivité à gauche de la somme sur le produit);
3. $A(B + B') = AB + AB'$ (distributivité à droite de la somme sur le produit);
4. $I_n A = A$, $BI_p = B$ (élément neutre pour le produit);

5. $a(AB) = (aA)B = A(aB)$;

Comme pour l'addition, l'associativité permet d'omettre les parenthèses: on écrit simplement ABC au lieu de $(AB)C$. Plus généralement $A_1 \cdots A_k$ désigne le produit de ces k matrices, avec 'importe quel parentésage.

Pour un entier naturel $k > 0$, on définit la puissance k-ème d'une matrice carrée $A \in M_n(\mathbb{R})$: $A^n = AA \cdots A$, avec n facteurs A.

Proposition 2.1. *La transposée d'un produit est égal au produit des transposées dans l'autre sens.*

La preuve est laissée en exercice.

Exercice 2.2. *Trouver des matrices A et B telles que $AB \neq BA$.*

Exercice 2.3. *Calculer le cube de la matrice* $\begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$.

Exercice 2.4. *Calculer la puissance n-ème de la matrice* $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Exercice 2.5. *Calculer la puissance n-ème de la matrice* $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. *Indication: essayer avec $n = 1, 2, 3$, deviner une formule générale et la prouver par récurrence sur n.*

Exercice 2.6. *Soit* $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Montrer que $M^2 - (a+d)M + (ad - bc)I_2 = 0$.*

Exercice 2.7. *Soit $L = (a_1, \ldots, a_n)$ une matrice ligne et M une matrice de taille $n \times p$ dont les lignes sont $L_1, \ldots, L_n$. Montrer que LM est la matrice-ligne égale à la combinaison linéaire $a_1L_1 + \cdots + a_nL_n$.*

Exercice 2.8. *Enoncer un analogue du résultat de l'exercice précédent avec une matrice-colonne et avec les colonnes d'une matrice.*

Exercice 2.9. *La* trace *d'une matrice carrée M, notée* $\mathrm{Tr}(M)$, *est la somme de ses éléments diagonaux. a) Montrer que si $A \in M_{np}(\mathbb{R})$ et $B \in M_{pn}(\mathbb{R})$, alors AB et BA sont des matrices carrées et $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. b) Montrer qu'il n'existe pas de matrices $A, B \in M_n(K)$ telles que $AB - BA = I_n$ ($n \geq 1$).*

Exercice 2.10. *Montrer que si M est une matrice, alors MM^t est une matrice carrée, dont la trace est égale à la somme des carrés de tous les coefficients de M.*

Exercice 2.11. *Soit A une matrice carrée et k un entier naturel > 0. Montrer par récurrence sur k que $(I-A)(I+A+A^2+\cdots+A^k) = I-A^{k+1}$. En déduire que si A est* nilpotente *(c'est-à-dire: il existe une puissance de A qui est nulle), alors $I-A$ est inversible.*

Exercice 2.12. *Soient A, B des matrices triangulaires supérieures d'ordre n, dont les éléments diagonaux sont repectivement $a_1, \ldots, a_n$ et $b_1, \ldots, b_n$. Montrer que les éléments diagonaux de la matrice AB sont $a_1b_1, \ldots, a_nb_n$.*

Exercice 2.13. *Soient A, B des matrices carrées de même taille. Montrer que la formule du binôme $(A+B)^n = \sum_{i=1}^{i=n} \binom{n}{i} A^i B^{n-i}$ est vraie si $AB = BA$. Montrer par un contre-exemple qu'elle n'est pas vraie en général ($n = 2$ suffira).*

2.4 Matrices inversibles

Définition 2.1. *Une matrice carrée A d'ordre n est dite* inversible à gauche *(resp.* inversible à droite*) s'il existe une matrice de même taille B (resp. C) telle que $BA = I_n$ (resp. $AC = I_n$).*

Elle est dite inversible *s'il existe B telle que $AB = BA = I_n$.*

Proposition 2.2. *Si une matrice a un inverse à gauche et un inverse à droite, alors ces inverses sont égaux et la matrice est inversible.*

Proposition 2.3. *Soient $A_1, \ldots, A_k$ des matrices carrées de même taille, inversibles. Alors leur produit est inversible et $(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}$.*

Autrement dit, l'inverse d'un produit de matrices inversibles est leur produit, mais dans l'autre sens.

Proposition 2.4. *Si une matrice est inversible, sa transposée l'est aussi et l'inverse de la transposée et la transposée de l'inverse.*

La preuve est laissée en exercice.

Exercice 2.14. *Soient A, B des matrices carrées de même taille. Montrer que si AB est inversible à droite, alors A est inversible à droite.*

Exercice 2.15. *Montrer que si A est une matrice inversible à droite et B une matrice telle que $BA = 0$, alors $B = 0$.*

Exercice 2.16. *Soit*

$$R(t) = \begin{bmatrix} e^t & 2te^t & (t^2-4)e^t & 2+2(t-1)e^t \\ 0 & e^t & te^t & e^t-1 \\ 0 & 0 & e^t & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

a) Montrer que $R(t+s) = R(t)R(s)$ pour tous nombres réels t et s. b) Quelle est l'inverse de $R(t)$?.

Exercice 2.17. *Montrer que si deux matrices carrées A, B commutent (c'est-à-dire $AB = BA$), et si elles sont inversibles, alors A et B^{-1} commutent, ainsi que A^{-1} et B^{-1}.*

Exercice 2.18. *Soient $A \in M_n(\mathbb{R})$, On suppose que A et $I_n + A$ sont inversibles. a) Montrer que I_n+A^{-1} a pour inverse $A(I_n+A)^{-1}$. b) Montrer que $(I_n + A^{-1})^{-1} + (I_n + A)^{-1} = I_n$.*

Exercice 2.19. *Montrer que si une matrice carrée A satisfait $A^2 = A$, alors $I - A$ n'est pas inversible, sauf si $A = 0$.*

Exercice 2.20. *Soit J la matrice $\frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Calculer J^2 et $(I-J)J)$. En désuire que les matrices J et $I - J$ ne sont pas inversibles.*

Exercice 2.21. *Montrer que si A, P sont des matrices carrées de même taille, et si P est inversible, alors $\mathrm{Tr}(P^{-1}AP) = \mathrm{Tr}(A)$. Utiliser l'exercice 2.9.*

Exercice 2.22. *On suppose que A est une matrice carrée non nulle telle que $A^2 = I$. Pour $r, s \in R$, montrer que $I+rA)(I+sA)$ est une combinaison linéaire de I et A. En déduire que $I + rA$ est inversible sauf si $r = -1$.*

2.5 Système d'équations linéaires de Cramer

Définition 2.2. *Le système d'équations linéaires $AX = B$, où A est une matrice inversible d'ordre n, où $X = (x_1, \ldots, x_n)^t$ est une matrice-colonnes de variables et où $B = (b_1, \ldots, b_n)^t$ est une matrice colonne, est appelé un* système de Cramer.

Théorème 2.1. *Un système de Cramer a une unique solution, qui est $X = A^{-1}B$.*

Proof. Si $AX = B$, on obtient par multiplication par A^{-1} à gauche: $X = IX = A^{-1}AX = A^{-1}B$. Réciproquement, si $X = A^{-1}B$, on obtient par multiplication à gauche par A: $AX = AA^{-1}B = B$. □

2.6 Opérations de lignes

Définition 2.3. *Les* opérations de lignes *sont de trois sortes, où A, B sont des matrices de même taille et où $A \longrightarrow B$ signifie qu'on transforme A en B par l'opération en question:*

- $A \stackrel{l_i+al_j}{\longrightarrow} B$, *qui signifie qu'on a ajouté à la ligne i de A la ligne j de A préalablement multipliée par a ($i \neq j$, $a \in \mathbb{R}$, $a \neq 0$).*
- $A \stackrel{(ij)}{\longrightarrow} B$, *qui signifie qu'on a échangé les lignes i et j dans A ($i \neq j$).*
- $A \stackrel{al_i}{\longrightarrow} B$, *qui signifie qu'on a multiplié par a la ligne i de A ($a \in \mathbb{R}$, $a \neq 0$).*

Définition 2.4. *Une* matrice élémentaire *est une matrice obtenue à partir d'une matrice identité par une opération de ligne.*

Proposition 2.5. *Si $A \longrightarrow B$ par une opération de ligne, alors $B = PA$, où P est la matrice élémentaire correspondant à cette opération.*

Corollaire 2.1. *Toute matrice élémentaire a un inverse qui est aussi une matrice élémentaire.*

Corollaire 2.2. *Si on transforme A en B par des opérations de lignes, alors il existe une matrice inversible P telle que $B = PA$. En fait, P est produit de matrices élémentaires.*

Définition 2.5. *Une matrice est dite* échelonnée *si elle a les propriétés suivantes:*

- *si une ligne est nulle, touts les nulles plus basses sont nulles;*
- *si une ligne est non nulle, le premier coefficient non nul dans cette ligne est* 1 *(on appellera* pivot *un tel élément);*
- *si une ligne est non nulle, avec le premier coefficient non nul en colonne j, alors la ligne suivante a son premier coefficient non nul en colonne $> j$.*

Une matrice échelonnée a l'allure suivante:

$$\begin{bmatrix}
0 & \dots & 0 & 1 & \times & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \times \\
0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 & \times & \dots & \dots & \dots & \dots & \dots & \times \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & \times & \dots & \times \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0
\end{bmatrix}$$

Théorème 2.2. *On peut par une suite d'opérations de lignes transformer toute matrice en une matrice échelonnée.*

Preuve et algorithme. On va raisonner de manière un peu littéraire.

Etape 1. Si la matrice a sa première colonne nulle, on travaille sur la matrice obtenue en la supprimant, et on la rétablit à la fin. On obtient bien une matrice échelonnée, car si on rajoute une colonne nulle à gauche d'une matrice échelonnée, on obtient une matrice échelonnée.

Etape 2. On peut donc supposer que la première colonne comporte un élément non nul. On en choisit un (de préférence le plus petit possible, 1 si c'est possible). Par une opération $(1i)$, on le ramène en haut à gauche. Par une opération al_1, on le change en 1. Puis par des opérations $l_i + al_1$, on annule tous les autres éléments de la première colonne (donc avec $i \geq 2$).

Etape 3: On a maintenant une matrice de la forme.

$$\begin{bmatrix}
1 & \times & \dots & \times \\
0 & \times & \dots & \times \\
\vdots & \vdots & \vdots & \vdots \\
0 & \times & \dots & \times
\end{bmatrix}$$

On retourne à l'étape 1 pour la matrice obtenue en supprimant la première ligne et la première colonne. On les rétablira à la fin, et on obtiendra une matrice échelonnée. $\square$

Dans la pratique, on peut ne pas supprimer des lignes ou des colonnes, mais on les garde tout au cours du calcul. C'est peu plus lourd, mais ça évite de devoir les garder en mémoire.

Proposition 2.6. *Si une matrice échelonnée a plus de lignes que de colonnes, alors sa dernière ligne est nulle.*

Proof. C'est évident. $\square$

Corollaire 2.3. *Si une matrice a plus de lignes que de colonnes, ses lignes sont linéairement dépendantes.*

Autrement dit: soit $M \in M_{np}(K]$ avec $n > p$; soient $l_1, \ldots, l_n$ ses lignes (chacune d'elles est une élément de $M_{1p}(\mathbb{R})$); alors il existe des réels $a_a, \ldots, a_n$ tels que $a_1 l_1 + \ldots + a)n l_n = 0$ et que (condition essentielle!) $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$.

Par exemple, pour

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

on a $(1,2) - 2(3,4) + (5,6) = (0,0)$ (autrement dit les réels a_1, a_2, a_3 sont ici $1, -2, 1$).

Preuve. Il existe par les propositions 2.2 et 2.2 une matrice inversible P est une matrice échelonnée N telle que $N = PM$. Celle-ci a sa dernière ligne nulle par la proposition 2.6. On a donc, avec $v = (0, \ldots, 0, 1)$, qui est un vecteur-ligne de longueur n: $vN = 0$. Posons $u = (a_1, \ldots, a_n) = vP$; alors $uM = 0$, ce qui exprime exactement que $a_1 l_1 + \ldots + a_n l_n = 0$. Mais u n'est pas nul: sinon $v = uP^{-1}$ serait nul, ce qui n'est pas. □

Exercice 2.23. *Soit $A \in M_n(R)$. Montrer que si on transforme par des opérations de lignes la matrice $[A, I_n]$ (c'est une matrice de taille $n \times 2n$ obtenue en plaçant I_n à la droite de A) en la matrice $[B, P]$, alors $B = PA$ (indication: montrer d'abord que $Q[A, B] = [QA, QB]$).*

Exercice 2.24. *Pour des matrices A, B de même taille, on écrit $A \sim B$ si on peut obtenir B à partir de A par une suite d'opérations de ligne. Montrer que $\sim$ est* relation d'équivalence, *c'est-à-dire qu'on a les trois propiétés (A, B, C sont quelconques, mais de même taille): (i) $A \sim A$; (ii) $A \sim B$ implique $B \sim A$; (iii) $A \sim B$ et $B \sim C$ implique $A \sim C$.*

Exercice 2.25. *Montre que certaines opérations de lignes commutent. Par exemple que si $A \xrightarrow{l_i + a l_j} B \xrightarrow{l_i + b l_k} C$ et $A \xrightarrow{l_i + b l_k} B' \xrightarrow{l_i + a l_j} C'$, alors $C = C'$. En déduire que les matrices élémentaires correspondant aux opérations $l_i + a l_j$ et $l_i + b l_k$ commutent. Montrer que ce n'est pas vrai pour les opérations $l_1 + l_2$ et $l_2 + l_1$, ni pour les matrices élémentaires correspondantes.*

Exercice 2.26. ** Pour n, p fixé, quel est le nombre de formes possibles de matrice échelonnées?*

2.7 Système d'équations linéaires: méthode de Gauss

Définition 2.6. *Une matrice est dite* échelonnée réduite *si elle est échelonnée et si de plus pour chaque pivot, sa colonne ne comporte à part lui que des 0.*

Proposition 2.7. *On peut par des transformations de lignes transformer toute matrice en un matrice échelonnée réduite.*

Preuve et algorithme. On peut partir d'une matrice échelonnée. Il suffit d'appliquer pour chaque pivot, en position i, j, des opérations de lignes $l_k + al_i$, $k \neq i$, de manière a annuler l'élément en position k, j. $\square$

Définition 2.7. *La* matrice *du système d'équations linéaires* homogène *(c'est-à-dire, dont les seconds membres sont nuls), à p inconnues $x_1, \ldots x_p$ et n équations*

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = 0,$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = 0,$$

$$\ldots,$$

$$a_{n1}x_2 + a_{n2}x_2 \cdots + a_{np}x_p = 0,$$

est la matrice $[a_{ij}]_{1 \leq i \leq n, 1 \leq j \leq p}$.

Pour résoudre ce système, on transforme la matrice en une matrice échelonnée réduite. Le système correspondant à cette nouvelle matrice est *équivalent* au précédent (c'est-à-dire, a les mêmes solutions). Alors les variables correspondant aux pivots (x_j avec un pivot en colonne j) sont appelées *liées* et les autres sont *libres* le nouveau système permet d,exprimer les variables liées en fonction des variables libres, qui peuvent prendre des valeurs arbitraires.

Pour un système général, où les seconds membres sont $b_1, \ldots, b_n$ (en place de 0 ci-dessus), la matrice du système est la même que ci-dessus sauf qu'on y rajoute à droite la colonne $(b_1, \ldots, b_n)^t$. On transforme cette matrice en une matrice échelonnée réduite. Le nouveau système est équivalent au précédent. Les lignes nulles donnent des équation du type $0x_1 + \cdots + 0x_p = c$; s'il y en a une avec $c \neq 0$, le sytème n'a pas de solution. Sinon, on obtient comme ci-dessus des expressions pour les variables liées en fonction des variables libres. S'il n'y a pas de variable libre, il y a une solution unique. Sinon, il y a une infinité de solutions obtenues en donnant des valeurs arbitraires aux variables libres.

Un examen attentif de la méthode de substitution et de la méthode ci-dessus révèle qu'elles sont essentiellement équivalentes: à condition, pour la méthode de substitution, de se donner un ordre sur les variables, qu'on éliminera dans cet ordre.

Une autre conséquence de l'existence de la matrice échelonnée réduite est la caractérisation suivante des matrices inversibles.

Corollaire 2.4. *Une matrice carrée est inversible si et seulement si elle est produit de matrices élémentaires.*

Proof. Un produit de matrices élémentaires est inversible (proposition 2.3 et corollaire 2.1). Réciproquement, soit A est une matrice inversible. Il existe donc, par la proposition 2.7 et le corollaire 2.2, une matrice P qui est produit de matrice élémentaires et une matrice B échelonnée réduite telle que $B = PA$. Alors B est inversible; donc B doit être la matrice identité (sinon elle a une ligne nulle, et ne peut être inversible). Donc l'inverse de A est P, et $A = P^{-1}$ est un produit de matrices élémentaires par la proposition 2.3 et le corollaire 2.1. $\square$

Exercice 2.27. ** Retravailler la preuve du corollaire 2.4 pour montrer que si A est carrée et inversible à droite, alors elle est produit de matrices élémentaires (indication: B est inversible à droite, carrée et réduite échelonnée; donc sa dernière ligne n'est pas nulle); donc inversible. En déduire, par transposition, qu'une matrice est inversible à droite si et seulement si elle est inversible à gauche.*

3 Espaces vectoriels et applications linéaires

3.1 Les huit axiomes d'un espace vectoriel

Définition 3.1. *Un* espace vectoriel *est un ensemble E qui a deux opérations. La première, appelé* addition*, ou* somme*, et la seconde est appelée* produit externe*. L'addition associe à deux éléments quelconques x, y de E un élément noté $x + y$. Le produit externe associe à un nombre réel A et à un élément x de E un élément de E noté ax. Ces deux opérations jouissent des propriétés suivantes (appelées* axiomes des espaces vectoriels*): quels que soient les éléments x, y, z de E et les réels a, b, on a:*

1. *$x + y = y + x$ (commutativité);*
2. *$(x + y) + z = (x + y) + z$ (associativité;*

3. *Il existe un élément* 0 *de* E *tel que* $x + 0 = x$ *(existence de l'élément neutre);*

4. *Il existe un élément* x' *de* E *tel que* $x + x' = 0$ *(existence de l'opposé);*

5. $1x = x$ *(le produit externe de* $1 \in R$ *avec* x *est égal à* x*);*

6. $(a + b)x = ax + bx$ *(distributivité à gauche du produit externe sur l'addition dans* $\mathbb{R}$*);*

7. $a(x + y) = ax + ay$ *(distributivité à droite du produit externe sur l'addition dans* E*);*

8. $(ab)x = a(bx)$ *(associativité).*

Pour un exemple concret de ces propriétés, regardez l'exemple 3.2 ci-dessous: x, y, z sont des matrices de même taille et ax désigne (comme déjà vu) le produit du réel a par la matrice x (ax s'obtient de x en y multipliant tous les coefficients par a)

On appelle souvent *scalaire* un élément de $\mathbb{R}$; ceci, par opposition aux éléments des éléments des espaces vectoriels, qui sont souvent appelés *vecteurs*.

On a $0e = 0$: en effet $0e + 0e = (0 + 0)e = 0e$, donc en ajoutant de chaque côté l'opposé de $0e$, on trouve $0e = 0$.

De plus l'opposé de e est $(-1)e$ (qu'on note $-e$): en effet, $e + (-1)e = 1e + (-1)e = (1 + (-1))e = 0e = 0$.

Exemple 3.1. *L'*espace vectoriel nul *est l'ensemble à un élément* $\{0\}$*. C'est un espace vectoriel sur* $\mathbb{R}$*. On a évidemment* $0 + 0 = 0$ *et* $a0 = 0$ *pour tout scalaire* a*.*

Exemple 3.2. *Fixons des entiers naturels* n *et* p*. Alors l'ensemble* $M_{np}(\mathbb{R})$ *est un espace vectoriel; car on peut additionner deux matrices, et on peut multiplier toute matrice par un scalaire, et tous les huit axiomes sont satisfaits.*

Exemple 3.3. *Fixons un intervalle* I *de* $\mathbb{R}$ *et soit* F *l'ensemble des fonctions de* I *dans* $\mathbb{R}$*. Alors* F *est un espace vectoriel. Car on peut additionner deux fonctions dans* F *et multiplier une fonction dans* F *par un scalaire; de plus les 9 axiomes sont satisfaits.*

Définition 3.2. *Soit* E *un espace vectoriel. Une* combinaison linéaire *dans* E *est une expression de la forme* $a_1e_1 + \ldots + a_ne_n$ *où* $a_1, \ldots, a_n$ *sont dans* $\mathbb{R}$ *et* $e_1, \ldots, e_n$ *sont dans* E*. On appelle* coefficients *de la combinaison linéaire les réels* $a_1, \ldots, a_n$*.*

Une telle expression s'écrit aussi $\sum_{i=1}^{i=n} a_i e_i$, ou aussi $\sum_{1 \leq i \leq n} a_i e_i$. On appelle *longueur* de la combinaison linéaire l'entier n. Le cas particulier $n = 0$ est aussi considéré; il est utile pour faire des raisonnements par récurrence.

La combinaison linéaire représente un vecteur dans E, défini par récurrence sur n: si $n = 0$, c'est le vecteur nul; si n=1, c'est $a_1 e_1$; et si $n \geq 2$, c'est $(a_1 e_1 + \ldots + a_{n-1} e_{n-1}) + a_n e_n$. le fait qu'on omet les parenthèses provient de ce que l'addition dans E est associative.

La combinaison linéaire ci-dessous est dite *triviale* si tous ses coefficients sont nuls.

Exercice 3.1. *On note E l'ensemble des matrices infinies à coefficients réels, dont les lignes et les colonnes sont indexées par les entiers naturels positifs 1,2,3,4.... Définir de manière naturelle une addition dans E. De même, un produit d'un scalaire par une telle matrice. Montrer que E est alors un espace vectoriel sur les réels (vérifier les 8 axiomes).*

Exercice 3.2. *Montrer que l'ensemble des polynômes de degré au plus 4, à coefficients réels, est un espace vectoriel.*

3.2 Applications linéaires

Une application linéaire est une fonction d'un espace vectoriel vers un autre qui préserve les opérations de ces espaces. Plus précisément:

Définition 3.3. *Soient E, F des espaces vectoriels. Une fonction $f : E \to F$ (c'est-à-dire une fonction de E vers F) est appelée une* application linéaire *si pour tous vecteurs x, y dans E et tout scalaire a on $f(x+y) = f(x)+f(y)$ et $f(ax) = af(x)$.*

Un exemple est la fonction *trace*, qui est une application linéaire de $M_n(\mathbb{R})$ dans $\mathbb{R}$: elle envoie toute matrice carrée sur la somme de ses éléments diagonaux. Un autre exemple est la fonction qui à toute fonction $I \to \mathbb{R}$ (où I est un intervalle fixé) associe $f(a)$ ($a \in I$ fixé). Un autre exemple est la fonction transposition $M_{np}(\mathbb{R}) \to M_{pn}(\mathbb{R})$.

3.3 Produit d'espaces vectoriels

Rappelons que le *produit cartésien* de deux ensembles E, F est l'ensemble des couples (x, y), $x \in E$, $y \in F$.

Si E, F sont des espaces vectoriels, on définit deux opérations sur leur produit cartésien: la somme et le produit externe. On les définit pour tous $x, x' \in E$, $y, y' \in F$, $a \in \mathbb{R}$ par les formules

$$(x, y) + (x', y') = (x + x', y + y'),\ a(x, y) = (ax, ay).$$

Proposition 3.1. *Soient E, F deux espaces vectoriels. Le produit cartésien $E \times F$, avec les deux opérations ci-dessus, est un espace vectoriel. Le vecteur nul de $E \times F$ est $(0_E, O_F)$.*

Proof. Vérifions par exemple l'associativité de l'addition: $((x, y) + (x', y')) + (x'', y'') = (x + x, , y + y') + (x'', y'') = ((x + x') + x'', (y + y') + y'') = (x + (x' + x''), y + (y' + y'')) = (x, y) + (x' + x'', y' + y'') = (x, y) + ((x', y') + (x'', y''))$. On a utilisé tour à tour la définition de la somme dans $E \times F$, ainsi que l'associativité de l'addition dans E et F.

Les 7 autres axiomes sont vérifiés sans problème. □

Plus généralement, si $E_1, \ldots, E_n$ sont des espaces vectoriels, $E_1 \times \cdots \times E_n$ est aussi un espace vectoriel. La somme et le produit externe sont définis par

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n).$$

Le cas particulier où $E_i = \mathbb{R}$ pour tout i est particulièrement important: $\mathbb{R}^n$ est un espace vectoriel. En fait on l'a déjà vu (exemple 3.2), puisque $\mathbb{R}^n = M_{1n}(\mathbb{R})$.

4 Sous-espaces vectoriels

4.1 Définition et caractérisation

Définition 4.1. *Un* sous-espace vectoriel *d'un espace vectoriel V est un sous-ensemble non vide de V qui est fermé sous les deux opérations de V.*

Notez que puisque le sous-espace est non vide, il contient 0 (obtenu en multipliant par le scalaire 0 n'importe quel vecteur du sous-espace).

Exemple 4.1. *Soir E l'espace vectoriel des matrices carrées d'ordre n. Considérons le sous-ensemble F des matrices de trace nulle. La matrice nulle est dans F; la somme de deux matrices de trace nulle est de trace nulle; et le produit externe d'une matrice de trace nulle par un réell est encore de trace nulle. Donc F est un sous-epace (attention: on considère ici le produit externe par un réel, pas le produit des matrices). Soyons encore plus concret: par exemple $n = 2$, $E = M_2(\mathbb{R}) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix}\}$ et $F = \{\begin{bmatrix} a & b \\ c & -a \end{bmatrix}\}$.*

Exemple 4.2. *Soit I un intervalle de $\mathbb{R}$, F l'espace vectoriel des fonctions de I dans $\mathbb{R}$, et C l'ensemble des fonctions continues de I dans $\mathbb{R}$. Alors C est un sous-espace vectoriel de F. En effet, C est bien un sous-ensemble de F. De plus, la fonction nulle est dans C (elle est continue). De plus si $f, g \in C$, alors $f + g \in C$ (la somme de deux fonctions continues sur I est continue). Endfin, si $f \in C$ et $a \in \mathbb{R}$, alors la fonction af est continue.*

Lorsque F est un sous-espace de l'espace vectoriel de E, F devient un espace vectoriel avec les *opérations induites* de E à F; c'est-à-dire, la somme dans E et le produit externe de E définissent sur F une somme et un produit externe car F est fermé sous ces opérations.

Proposition 4.1. *Un sous-espace vectoriel d'un espace vectoriel est un espace vectoriel avec les opération induites.*

Proof. Soit F un sous-espace de E. Le 0 est dans F, comme nous l'avons vu. De plus si $x \in F$, alors son opposé est $(-1)x$: il est aussi dans F. Donc les axiomes 3 et 4 sont satisfaits. Les 6 autres axiomes sont évidemment satisfaits. $\square$

Pour montrer qu'un sous-ensemble donné d'un espace vectoriel en est un sous-espace, on applique la recette suivante.

Proposition 4.2. *Soit F un sous-ensemble de l'espace vectoriel de E. Alors F est un sous-espace vectoriel de E si et seulement si on a les 3 propriétés suivantes:*

1. $0 \in F$;
2. *pour tous x, y dans F, on a $x + y \in F$;*
3. *pour tous $x \in E$ et pour tout $a \in \mathbb{R}$, on a $ax \in F$.*

Proof. C'est évident. $\square$

Exercice 4.1. *Montrer que si G est sous-espace de F et si F est sous-espace de E, alors G est sous-espace de E.*

Exercice 4.2. *Montrer que l'espace des fonctions dérivables de I dans $\mathbb{R}$ est un sous-espace de l'ensemble des fonctions continues de I dans $\mathbb{R}$, lui-même un sous-espace des fonctions de I dans R.*

Exercice 4.3. *Montrer que l'ensemble des fonctions deux fois dérivables $f : \mathbb{R} \to \mathbb{R}$ telles que $f'' + f = 0$ est un espace vectoriel. Indication: montrer que c'est un sous-espace d'un espace judicieusement choisi.*

Exercice 4.4. *Montrer que l'ensemble des* matrices magiques *dans* $M_{np}(\mathbb{R})$ *(c'est-à-dire l'ensemble des matrices dont la somme de chaque ligne et de chaque colonne est nulle) est un sous-espace de* $M_{np}(\mathbb{R})$*.*

Exercice 4.5. *Est-ce que l'ensemble des matrices de trace égale à 2 est un sous-espace de* $M_2(\mathbb{R})$*?*

Exercice 4.6. *Est-ce que l'ensemble des matrices de déterminant non nul est un sous-espace de* $M_2(\mathbb{R})$*?*

4.2 Intersection de sous-espaces et systèmes d'équations linéaires

Rappelons que si F, G sont ensembles, leur *intersection* est l'ensemble des éléments de F qui sont aussi dans G:

Proposition 4.3. *L'intersection de deux sous-espaces d'un espace vectoriel est un sous-espace.*

Proof. Si $x, y \in F \cap G$, alors $x \in Fety \in F$; donc $x + y \in F$. De même, $x + y \in G$. Donc $x + y \in F \cap G$.

Si $x \in F \cap G$ et $a \in \mathbb{R}$, alors $x \in F$, donc $ax \in F$. De même $ax \in G$. Donc $ax \in F \cap G$. □

Corollaire 4.1. *L'intersection d'un nombre fini de sous-espaces d'un espace vectoriel est un sous-espace.*

Proof. Récurrence sur le nombre de sous-espaces. □

Un exemple important de sous-espace vectoriel est le sous-espace de $\mathbb{R}^p$ associé à un système d'équations linéaires homogène. Considérons le système écrit dans la définition 2.7. L'ensemble des $(x_1, \ldots, x_p \in \mathbb{R}^p$ qui satisfont de ssytème est un sous-espace de $\mathbb{R}^p$. Pour le voir, on remarque que cet ensemble est l'intersection des n sous-ensembles $\{(x_1, \ldots, x_p \in \mathbb{R}^p, a_{i1}x_1 + \cdots + a_{ip}x_p = 0\}$, $i = 1, \ldots, n$.

On est donc ramené (par le corollaire 4.1) au cas d'une seule équation linéaire: $a_1x_1 + \ldots + a_px_p = 0$. On remarque que l'ensemble de p-uplets qui satisfont une équation linéaire satisfait les 3 conditions de la proposition 4.2 (une preuve plus moderne consiste à vérifier que cet ensemble est le noyau de l'application linéaire $(x_1, \ldots, x_p) \in \mathbb{R}^p \mapsto a_1x_1 + \ldots + a_px_p, \mathbb{R}^p \to \mathbb{R}$, voir le corollaire 6.1)

Qu'en est-il des systèmes d'équations linéaires généraux non homogènes? Ceci est traité dans l' exercice 4.8.

Exercice 4.7. *Montrer que l'intersection d'une famille quelconque de sous-espaces d'un espace vectoriel est un sous-espace.*

Exercice 4.8. *Un* sous-espace affine *d'un espace vectoriel E est une partie de E qui est soit vide, soit de la forme $e + F = \{e + f, f \in F\}$, où F est un sous-espace vectoriel de E et $e \in E$. Montrer que si $e' \in e + F$, alors $e + F = e' + F$. Montrer que l'intersection de deux sous-espaces affines est un sous-espace affine; de même pour un nombre fini quelconque de sous-espaces affines. Montrer que l'ensemble des solutions d'un système d'équations linéaires en p variables est un sous- espace affine de $\mathbb{R}^p$.*

4.3 Sommes de deux sous-espaces

Définition 4.2. *On dit que E est* somme *de ses deux sous-espaces F et G si si tout vecteur dans E est somme d'un vecteur dans F est d'un vecteur dans G. Notation: $E = F + G$.*

Autrement dit, on a:
(i) $\forall x \in E$, $\exists y \in F$ et $\exists z \in G$ tels que $x = y + z$.

Définition 4.3. *On dit que F, G, sous-espaces de E, sont* supplémentaires, *si tout vecteur dans E est de manière unique somme d'un vecteur dans F est d'un vecteur dans G. On dit aussi que E est* somme directe *de F et G.*

Autrement dit on a la propriété (i) ci-dessus (existence) ainsi que la propriété (ii) ci-dessous:
(ii) (unicité) $\forall x, x' \in F$ et $\forall y, y' \in G$, $x + y = x' + y'$ implique $x = x'$ et $y = y'$.

Proposition 4.4. *Les deux sous-espaces F, G de E sont supplémentaires si et seulement si $F + G = E$ et $F \cap G = \{0\}$.*

Proof. Supposons que F, G soient supplémentaires. Alors □

Exercice 4.9. *Soient E, F deux espaces vectoriels. Montrer que leur sous-espace $G_1 = \{(0, f), f \in F\}$ et $G_2 = \{(e, o), e \in E\}$ sont supplémentaires.*

Exercice 4.10. *Montrer que l'ensemble des matrices symétriques et l'ensemble des matrices anti-symétriques dans $M_n(\mathbb{R})$ en sont des sous-espaces supplémentaires.*

Exercice 4.11. *Montrer que l'ensemble des fonctions paires et l'ensemble des fonctions impaires sont des sous-espaces supplémentaires de l'espace vectoriel des fonctions $\mathbb{R} \to \mathbb{R}$.*

5 Bases et dimension

5.1 Dépendance linéaire

Définition 5.1. *On dit que des vecteurs* $v_1, \ldots, v_n$ *d'un espace vectoriel* V *sont* linéairement dépendants *s'il existe des scalaires* $a_1, \ldots, a_n$ *tels que* $(a_1, \ldots, a_n) \neq 0$ *et que* $a_1 v_1 + \cdots + a_n v_n = 0$.

Pour exprimer que $(a_1, \ldots, a_n) \neq 0$, on dit que $a_1, \ldots, a_n$ sont *non tous nuls*. Cette condition est essentielle, car si on l'omet, la définition précédente n'a pas de sens (puisque pour tous vecteurs $v_1, \ldots, v_n$ on a $0v_1 + \ldots + 0v_n = 0$).

Définition 5.2. *On dit que des vecteurs* $v_1, \ldots, v_n$ *d'un espace vectoriel* V *sont* linéairement indépendants *s'ils ne sont pas linéairement dépendants.*

Cela signifie que quels que soient les scalaires $a_1, \ldots, a_n$,

$$a_1 v_1 + \cdots + a_n v_n = 0 \Rightarrow a_1 = \ldots = a_n = 0.$$

Ceci donne la recette suivante: si vous voulez prouver que des vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants, vous prenez des scalaires quelconques $a_1, \ldots, a_n$ (qui seront donc des variables) et vous prouvez l'implication ci-dessus.

Exemple 5.1. *Prouvons que les vecteurs* $(1, 2)$ *et* $(2, 3)$ *sont linéairement indépendants. Soient* a, b *des réels quelconques; écrivons qu'on a le côté gauche de l'implication ci-dessus:* $a(1, 2) + b(2, 3) = 0$. *On obtient* $(a + 2b, 2a + 3b) = (0, 0)$, *donc* $a + 2b = 0$ *et* $2a + 3b = 0$. *On résout ce système d'équations linéaires, et on trouve* $a = b = 0$*; cqfd.*

Définition 5.3. *On dit qu'un vecteur* v dépend linéairement de $v_1, \ldots, v_n$ *si* v *est égal à une combinaison linéaire de* $v_1, \ldots, v_n$.

Proposition 5.1. *Soit* $p < n$. *Si* $v_1, \ldots, v_p$ *sont linéairement indépendants et si* $v_1, \ldots, v_n$ *sont linéairement dépendants, alors il existe* $j \in \{p+1, \ldots, n\}$ *tel que* v_j *dépend linéairement des autres vecteurs* v_k, $k = 1, \ldots, n$, $k \neq j$.

Preuve. Il existe des scalaires $a_1, \ldots, a_n$, non tous nuls, tels que $a_1 v_1 + \cdots + a_n v_n = 0$. On ne peut avoire $a_{p+1} = \ldots = a_n = 0$: sinon en effet les $a_1, \ldots, a_p$ sont non tous nuls et $a_1, \ldots, a_p$ sont linéairement dépendants.

Il existe donc $j \in \{p+1, \ldots, n\}$ tel que $a_j \neq 0$. Alors $a_j v_j = -\sum_{i \neq j} a_i v_i$ et par suite $v_j = -\sum_{i \neq j} \frac{a_i}{a_j} v_i$. □

Proposition 5.2. *Si v dépend linéairement de $v_1, \ldots, v_n$ et si v_n dépend linéairement de $v_1, \ldots, v_{n-1}$, alors v dépend linéairement de $v_1, \ldots, v_{n-1}$.*

Preuve. On a en effet $v = a_1 v_1 + \ldots + a_n v_n$ et $v_n = a_1 v_1 + \ldots + a_n v_{n-1}$ etc... □

Proposition 5.3. *Si des vecteurs $v_1, \ldots, v_n$ d'un espace vectoriel dépendent linéairement de vecteurs $x_1, \ldots, x_p$ et si $n > p$, alors $v_1, \ldots, v_n$ sont linéairement dépendants.*

Preuve. Ecrivons que chaque v_i est combinaison linéaire des x_j: $v_i = \sum_{j=1}^{j=p} a_{ij} x_j$, pour tout $i = 1, \ldots, n$, avec des scalaires a_{ij}. Ceci définit une matrice $A \in M_{np}(\mathbb{R})$. Cette matrice a plus de lignes que de colonnes. Donc par le corollaire 2.3, ses lignes sont linéairement dépendantes: il existe des scalaires $\lambda_1, \ldots, \lambda_n$, non tous nuls, tels que $\lambda_1 l_1 + \ldots + \lambda_n l_n = 0$. Ceci signifie que pour tout $j = 1, \ldots, p$, la j-ème composante de ce vecteur (qui est dans $\mathbb{R}^p$) est nul, donc $\sum_{i=1}^{i=n} \lambda_i a_{ij} = 0$.

On en déduit que $\sum_{i=1}^{i=n} \lambda_i v_i = \sum_{i=1}^{i=n} \lambda_i \sum_{j=1}^{j=p} a_{ij} x_j = \sum_{i=1}^{i=n} \sum_{j=1}^{j=p} \lambda_i a_{ij} x_j = \sum_{j=1}^{j=p} \sum_{i=1}^{i=n} \lambda_i a_{ij} x_j$ (par interversion des sommations) $= \sum_{j=1}^{j=p} (\sum_{i=1}^{i=n} \lambda_i a_{ij}) x_j = \sum_{j=1}^{j=p} 0 x_j = 0$.

Comme les λ_i sont non tous nuls, les v_i sont linéairement dépendants. □

Le lecteur qui a du mal à comprendre les formules avec les $\sum$ peut faire le calcul de la preuve précédente en prenant $n = 3, p = 2$.

Exercice 5.1. *Montrer que si x, y sont linéairement indépendants, alors aussi $x, x + y$. Généraliser: si $x_1, \ldots, x_n$ sont linéairement dépendants, alors aussi $x_1, \ldots, x_{n-1}, x_n + a_1 x_1 + \cdots + a_{n-1} x_{n-1}$.*

Exercice 5.2. *Dans $\mathbb{R}^n$: on suppose que $a_1 + \cdots + a_n = 0$. Montrer que $(a_1, \ldots, a_n)$ est combinaison linéaire de $(1, -1, 0, \ldots, 0), (0, 1, -1, 0, \ldots, 0), (0, \ldots, 1, -1)$.*

5.2 Bases: existence et unicité de la dimension

Définition 5.4. *On dit que l'espace vectoriel V est* finiment engendré *s'il existe des vecteurs $v_1, \ldots, v_n$ dans V, en nombre fini, tels que tout vecteur dans V dépend linéairement de $v_1, \ldots, v_n$. Dans ce cas, on dit que $v_1, \ldots, v_n$* engendrent *V.*

Définition 5.5. *Soit V un espace vectoriel finiment engendré. Une* base *de V est une suite de vecteurs $v_1, \ldots, v_n$ de V qui l'engendrent et qui sont linéairement indépendants.*

Un exemple typique est l'espace vectoriel $M_{np}(\mathbb{R})$ et sa *base canonique*. Celle-ci consiste en les np matrices E_{ij}: cette matrice a tous ses coefficients nuls, sauf celui en position i, j, qui vaut 1.

Théorème 5.1. *Soit $p \leq q$. Soient $v_1, \ldots, v_q$ des vecteurs d'un espace vectoriel V, qui engendrent V, et tels que $v_1, \ldots, v_p$ soient linéairement indépendants. Il existe alors parmi les vecteurs $v_{p+1}, \ldots, v_q$ des vecteurs $u_1, \ldots, u_r$ tels que $v_1, \ldots, v_p, u_1, \ldots, u_r$ forment une base de V.*

Preuve. (récurrence sur $q-p$) Si $q-p = 0$, il n'y a rien à démontrer: en effet, dans ce cas, $v_1, \ldots, v_p$ est déjà une base de V. Supposons que $q - p > 0$, c'est-à-dire $p < q$. Si les vecteurs $v_1, \ldots, v_q$ sont linéairement indépendants, alors ils forment une base de V et on a fini. Sinon, ils sont linéairement dépendants. Alors par la proposition 5.1, il existe $j > p$ tel que v_j soit linéairement dépendant des autres vecteurs v_k, $k \in \{1, \ldots q\} \backslash \{j\}$. Sans perte de généralité, on peut supposer que c'est v_q qui est linéairement dépendant de $v_1, \ldots, v_{q-1}$. Alors ces $q - 1$ vecteurs engendrent V (proposition 5.3). On a $p \leq q - 1$ et $q - 1 - p < p$. Par hypothèse de récurrence, on obtient donc qu'il existe des vecteurs $u_1, \ldots, u_r$, pris parmi $v_{p+1}, \ldots, v_{q-1}$ tels que $v_1, \ldots, v_p, u_1, \ldots, u_r$ est une base de V. □

Corollaire 5.1. *Tout espace vectoriel finiment engendré possède une base finie.*

Preuve. Il existe des vecteurs $v_1, \ldots, v_q$ qui engendrent V. On applique le théorème avec $p = 0$. □

Corollaire 5.2. *(théorème dit de la"base incomplète") Soient $v_1, \ldots, v_p$ des vecteurs linéairement indépendants dans un espace vectoriel finiment engendré V. Il existe alors des vecteurs $v_{p+1}, \ldots, v_n$ tels que $v_1, \ldots, v_n$ forment une base de E.*

Preuve. Il existe des vecteurs $u_{p+1}, \ldots, u_q$ qui engendrent V. On applique le théorème aux vecteurs $v_1, \ldots, v_p, u_{p+1}, \ldots, u_q$ qui engendrent évidemment V. □

Corollaire 5.3. *Si $v_1, \ldots, v_q$ engendrent V, il existe parmi ces vecteurs des vecteurs qui forment une base de V.*

Preuve. On applique le théorème avec $p = 0$. □

Théorème 5.2. *Dans un espace vectoriel finiment engendré, toutes les bases ont le même nombre d'éléments.*

Proof. Si on avait deux bases de cardinalités différentes, cela contredirait la proposition 5.3. □

Définition 5.6. *Soit E un espace vectoriel finiment engendré. On appelle* dimension *de E le nombre d'éléments d'une base de E. Notation:* $\dim(E)$.

Théorème 5.3. *Soient E un espace vectoriel de dimension n et $e_1, \ldots, e_n \in E$. Les trois conditions suivantes sont équivalentes:*
(i) $e_1, \ldots, e_n$ forment une base;
(ii) $e_1, \ldots, e_n$ sont linéairement indépendants;
(iii) $e_1, \ldots, e_n$ engendrent E.

Il peut paraître bizarre que (ii) ou (iii) suffise. Mais c'est parce que le nombre des vecteurs est égal à la dimension de l'espace. Pour appliquer le théorème, il ne faut pas oublier de vérifier cette condition, qui suppose entre autres qu'on connaisse la dimension.

Proof. Il est clair que (i) implique (ii) et (iii).

(ii) $\Rightarrow$ (i): on peut compléter ces n vecteurs en une base de E (corollaire 5.2). Mais une base de E a n éléments. Donc $e_1, \ldots, e_n$ forment une base.

(iii) $\Rightarrow$ (i): si les e_i étaient linéairement dépendants, l'un d'eux serait linéairement dépendants des autres. On trouverait alors un système de $n-1$ vecteurs qui engendrent E (proposition 5.2). Donc une base avec au plus $n-1$ éléments (corollaire 5.3), contradiction. □

5.3 Bases des sous-espaces

Proposition 5.4. *Tout sous-espace d'un espace vectoriel de dimension fini n est un espace vectoriel de dimension finie $\leq n$.*

Proof. Soit F un sous-espace d'un espace vectoriel E de dimension finie n. Soit p maximum tel qu'il existe dans F p vecteurs linéairement indépendants. Ce p existe (c'est-à-dire, ce n'est pas l'infini) car dans E (donc dans F) $n+1$ vecteurs sont toujours linéairement indépendants, d'après la proposition 5.3 et le fait que tout vecteur dans E dépend linéairement des n vecteurs d'une base de E.

Soit donc $v_1, \ldots, v_p$ p vecteurs linéairement indépendants dans F. Montrons qu'ils engendrent F (et on en conclura qu'ils forment une base de F, qui est donc de dimension finie $p \leq n$). Soit v un vecteur quelconque dans F. Alors $v_1, \dots, v_n, v$ sont linéairement indépendants, par maximalité de p. Donc v est linéairement dépendant de $v_1, \dots, v_n$, d'après la proposition 5.1. □

Exercice 5.3. *Quelle est la dimension de l'espace vectoriel des matrices triangulaires supérieures dans* $M_n(\mathbb{R})$*?*

Exercice 5.4. *Montrer que dans* $M_n(\mathbb{R})$*, le sous-espace des matrices triangulaires supéurieres et celui des matrices strictement inférieures (c'est-à-dire, avec des 0 sur la diagonale) sont supplémentaires.*

Exercice 5.5. *Montrer que* F, G *sous-espaces de* E *de dimension finie, sont supplémentaires si et seulement s'il existe une base de* F *et une base* G *dont la réunion est une base de* E*. Montrer que c'est aussi vraie pour toutes les bases.*

Exercice 5.6. *Soit* $E = \mathbb{R}[x]$*,* F *le sous-espace engendré par les* x^n *avec* n *multiple de* 3*, et* G *le sous-espace engendré par les* x^n*,* n *pas multiple de* 3*. Montrer que* F, G *sont supplémentaires.*

Exercice 5.7. *Soit* F*,* G *deux sous-espaces supplémentaires d'un espace vectoriel* E *de dimension* n*. Onsuppose que* F *est de dimension* $n-2$*. Quelle est la dimension de* G*? Soit* u, v *une base de* G*,* $f \in F$ *et* G' *le sous-espace engendré par* $u+f, v+f$*. Quelle est las dimension de* G'*? Montrer que* F *et* G' *sont supplémentaires.*

Exercice 5.8. *Soit* E *l'ensemble des suites réelles* $(a_n)_{n \in \ltimes}$*. Soit* a *un réel. Soit* F *le sous-ensemble* E *constitué des suites qui satisfont* $\forall n \in \mathbb{N}, a_{n+1} = aa_n$*. Montrer que* F *est un sous-espace de* E*. Montrer que toute suite* (a_n) *dans* F *satisfait* $\forall n \in \mathbb{N}$*,* $a_{n+1} = a^{n+1}a_0$*. Montrer que* F *est de dimension 1.*

Exercice 5.9. *Avec* E *comme dans l'exercic précédent, et* a, b *des réels, on considère l'ensemble* G *des suites* (a_n) *dans* E *qui satisfont* $\forall n \in \mathbb{N}, a_{n+2} = aa_{n+1} + ba_n$*. Montrer que* G *est un sous-espace de dimension de 2 de* E*. Montrer que pour un réel* r*, la suite* $(1, r, r^2, r^3, \dots)$ *est dans* G *si et seuleemnt si* $r^2 = ar + b$*.*

6 Applications linéaires

6.1 Exemples

Les exemples sont légion (comme les cônes oranges à Montréal).

Exemple 6.1. *La fonction* identité *d'un espace vectoriel dans lui-même, c'est-à-dire la fonction qui envoie tout vecteur sur lui même, est une application linéaire. Si E est l'espace en question, on la note* id_E*,ou simplement* id*. On a donc:* $\forall x \in E$, $\mathrm{id}(x) = x$*.*

Exemple 6.2. *Soient E, F deux espaces vectoriels. La fonction qui envoie tout vecteur sur le vecteur nul (de F), est une application linéaire. On la note simplement* 0*. On a donc* $0(x) = 0$*,* $\forall x \in E$*. On appelle cette fonction la* fonction nulle *de E vers F.*

Exemple 6.3. *La fonction qui envoie une matrice sur sa transposée est une application linéaire de $M_{np}(\mathbb{R})$ dans $M_{pn}(\mathbb{R})$.*

Exemple 6.4. *La fonction de $\mathbb{R}^3$ dans lui-même qui envoie (x, y, z) sur $(x, x + y, x + y + z)$ est une application linéaire.*

Exemple 6.5. *La dérivation, qui envoie toute fonction sur sa dérivée, est une application linéaire de l'espace vectoriel des fonctions dérivables $I \to \mathbb{R}$ dans lui-même (I est un intervalle de $\mathbb{R}$).*

Exemple 6.6. *La fonction $X \mapsto AX$ est une applicaction linéaire de $M_{pq}(\mathbb{R})$ dans $M_{nq}(\mathbb{R})$ (ici $A \in M_{np}(\mathbb{R})$ est fixée).*

Exercice 6.1. *Soit E un espace vectoriel et $e \in E$. Montrer que la fonction qui à $a \in R$ associe ae est une application linéaire $\mathbb{R} \to E$.*

Exercice 6.2. *Montrer que toute application linéaire $f : \mathbb{R} \to E$ est de la forme ci-dessus (prendre $e = f(1)$).*

Exercice 6.3. *Montrer que la fonction $M_{np}(\mathbb{R})$ dans $\mathbb{R}$ qui à la matrice M associe la somme de tous ses coefficients est une application linéaire.*

6.2 Propriétés

Proposition 6.1. *Soit f une application linéaire de E vers F. On a $f(0) = 0$ et $f(\sum_i a_i x_i) = \sum_i a_i f(x_i)$ quels que soient les vecteurs $x_1, \ldots, x_n$ dans E les scalaires $a_1, \ldots, a_n$.*

Autrement dit, f envoie le vecteur nul (de E) sur le vecteur nul (de F). La seconde propriété exprime que f préserve les combinaisons linéaires.

Preuve. On a $f(0_E) = f(0_{\mathbb{R}} 0_E) = 0_{\mathbb{R}} f(0_E) = 0_F$.

Pour les combinaisons linéaires, on raisonne par récurrence sur n. Pour $n = 0$, c'est vrai, car la combinaison linéaire de longueur nulle vaut 0 et on applique ce qu'on vient de voir. Passage de n à $n+1$: on a $f(\sum_{1\leq i\leq n+1} a_i x_i) = f((\sum_{1\leq i\leq n} a_i x_i) + a_{n+1}x_{n+1}) = f(\sum_{1\leq i\leq n} a_i x_i) + f(a_{n+1}x_{n+1})$ (par linéarité de f) $= (\sum_{1\leq i\leq n} a_i f(x_i)) + a_{n+1} f(x_{n+1})$ (par hypothèse de récurrence et par linéarité de f) $= \sum_{1\leq i\leq n+1} a_i f(x_i)$. $\square$

Définition 6.1. *Si f, g sont des applications linéaires de E vers F, on appelle* somme *de f et g, notée $f+g$, la fonction $E \to F$ qui envoie tout vecteur x dans E sur $f(x)+g(x)$. Si a est un scalaire, on appelle* produit externe de a par f *la fonction $E \to F$ qui envoie tout vecteur x de E sur $af(x)$.*

On a donc les formules

$$(f+g)(x) = f(x) + g(x),\ (af)(x) = af(x).$$

Proposition 6.2. *Avec cette définition, $f+g$ et af sont des applications linéaires. L'ensemble des applications linéaires de E vers F, muni de ces deux opérations, est un espace vectoriel. Le vecteur nul de cet espace est la fonction nulle. L'opposée de f est la fonction $-f = (-1)f$.*

Preuve. Il faut vérifier les huit axiomes. C'est routinier, mais long. $\square$

Rapellons la notion de *composition* de deux fonctions. Si f est une fonction de E vers F et si g est une fonction de F vers G, alors la *composée* de ces deux fonctions est la fonction $g \circ f$ définie par

$$\forall x \in E, g \circ f(x) = g(f(x)).$$

Exemple 6.7. *La composée de $f(x) = x^2 + 1$, $\mathbb{R} \to \mathbb{R}_+$, par la fonction $g(x) = \sqrt{x}$, $\mathbb{R}_+ \to \mathbb{R}$, est la fonction $g \circ f(x) = \sqrt{x^2+1}$.*

Proposition 6.3. *La composée de deux applications linéaires est une application linéaire.*

Autrement dit, si f est une application linéaire de E vers F et si g est une application linéaire de F vers G, alors $g \circ f$ est une application linéaire de E vers G.

Preuve. Soient x, y dans E et a dans $\mathbb{R}$. On $g \circ f(x+y) = g(f(x+y)) = g(f(x)+f(y)) = g(f(x))+g(f(y)) = g\circ f(x)+g\circ f(y)$. De plus, $g\circ f(ax) = g(f(ax)) = g(af(x)) = ag(f(x)) = ag \circ g(x)$. $\square$

Exemple 6.8. *Soit I un intervalle de $\mathbb{R}$ qui contient 0. Soit E l'espace vectoriel des fonctions dérivables $I \to \mathbb{R}$. La fonction qui envoie toute fonction sur sa dérivée est une application linéaire de E dans lui-même. La fonction de E dans $\mathbb{R}$ qui envoie f sur $f(0)$ est une application linéaire. La composée de ces deux fonctions est la fonction qui envoie tout $f \in E$ sur $f'(0)$: elle est linéaire par la proposition précédente.*

6.3 Applications linéaires et sous-espaces

Proposition 6.4. *Une application linéaire envoie tout sous-espace sur un sous-espace, par image directe et inverse.*

Preuve. Soit $f : E \to F$ une application linéaire. On applique ci-dessous, systématiquement, plusieurs fois, et sans le dire, la proposition 4.2.

1. Soit V un sous-espace de E. Montrons que $f(V)$ est un sous-espace de F. On a $0 = f(0) \in F(V)$, car $0 \in V$. Si $u, v \in f(V)$, il existe $x, y \in V$ tels que $u = f(x)$ et $v = f(y)$; alors $u+v = f(x)+f(y) = f(x+y) \in f(V)$. Si $v \in f(V)$ et $a \in \mathbb{R}$, alors il existe $x \in V$ tel que $v = f(x)$; alors $av = af(x) = f(ax) \in f(V)$.

2. Soit V un sous-espace de F. Montrons que $f^{-1}(V)$ est un sous-espace de F. On a $0 \in f^{-1}(V)$ car $f(0) = 0 \in V$. Soient $x, y \in f^{-1}(V)$, c'est-à-dire $f(x), f(y) \in V$; alors $f(x+y) = f(x)+f(y) \in V$, donc $x+y \in f^{-1}(V)$. Soit $x \in f^{-1}(V)$ et $a \in \mathbb{R}$; alors $f(x) \in V$, donc $f(ax) = af(x) \in V$, donc $x \in f^{-1}(V)$. $\square$

Définition 6.2. *Le* noyau *d'une application linéaire est l'ensemble des vecteurs qu'elle envoie sur* 0.

Si f est une application linéaire de E vers F, on note $\mathrm{Ker}(f)$ son noyau. On a donc $\mathrm{Ker}(f) = \{x \in E | f(x) = 0\}$.

Corollaire 6.1. *Si $f : E \to F$ est une application linéaire, alors* $\mathrm{Ker}(f)$ *est un sous-espace de E.*

Preuve. Le noyau est en effet l'image réciproque du sous-espace nul. $\square$

Définition 6.3. *L'*image *d'une application linéaire $f : E \to F$ est l'ensemble $f(E)$.*

Notation: $\Im(f)$. On a donc $\Im(f) = \{y \in F | \exists x \in E, f(x) = y\}$.

Corollaire 6.2. *Si $f : E \to F$ est une application linéaire, alors $\Im(f)$ est un sous-espace de F.*

Preuve. $\Im(f)$ est en effet l'image de E (sous-espace de lui-même) par f. $\square$

Théorème 6.1. *(théorème du rang) Soit E, F des espaces vectoriels de dimensions finies et $f : E \to F$ une application linéaire. On* $\dim(E) = \dim(\mathrm{Ker}(f)) + \dim(\Im(f))$.

Ce théorème s'appelle théorème du rang car on appelle *rang* de f la dimension de $\Im(f)$. Notation $\mathrm{rg}(f)$.

Preuve. $\mathrm{Ker}(f)$ est un sous-espace de E. Il est de dimension finie p. Il existe une base $e_1, \dots, e_n$ de E tel que $e_1, \dots, e_p$ est une base de $\mathrm{Ker}(f)$. Alors $f(e_{p+1}), \dots, f(e_n)$ est une base $\Im(f)$. $\square$

6.4 Injections, surjections, isomorphismes

Rappelons qu'une fonction $f : E \to F$ est dite *injective* si elle satisfait à l'une des conditions équivalentes suivantes:

1. Si $x, y \in E$ et si $x \neq y$, alors $f(x) \neq f(y)$;
2. si $x, y \in E$ et si $f(x) = f(y)$, alors $x = y$;
3. si $v \in F$, il existe au plus un $x \in E$ tel que $f(x) = v$ (on dit que tout v dans E a au plus un *antécédant* (ou *image inverse*) par f dans E).

On dit alors aussi que f est une *injection.*

Exemple 6.9. *La fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe x^2 n'est pas injective; en effet, on a $1 \neq -1$ mais $f(1) = f(-1)$.*

Exemple 6.10. *La fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe x^3 est injective; en effet, tout réel a une unique racine cubique. Donc, $\forall v \in R$, il existe au plus un $x \in R$ tel que $f(x) = v$.*

Exemple 6.11. *La fonction de $\mathbb{N}$ dans $\mathbb{N}$ qui à n associe $2n$ est injective; en effet, pour tout entier naturel n il existe au plus un entier naturel p tel que $2p = n$.*

Proposition 6.5. *Une application linéaire est injective si et seulement si son noyau est nul.*

Proof. Le noyau de f contient toujours 0. Supposons que f soit injective et soit $x \in \text{Ker}(f)$. Alors $f(x) = 0 = f(0)$. Par injectivité (propriété 2. ci-dessus), on doit avoir $x = 0$. Donc $\text{Ker}(f) = \{0\}$.

Réciproquement, supposons que $\text{Ker}(f) = \{0\}$. Suposons que $f(x) = f(y)$. Alors $f(x - y) = f(x) - f(y) = 0$. Donc $x - y \in \text{Ker}(f)$. Donc $x - y = 0$ et par suite $x = y$. Donc f est injective. $\square$

Rappelons qu'une fonction $f : E \to F$ est dite *surjective* si elle satisfait à l'une des conditions équivalentes suivantes:

1. Si $y \in F$, il existe $x \in E$ tel que $y = f(x)$;
2. $f(E) = F$;
3. si $v \in F$, il existe au moins un $x \in E$ tel que $f(x) = v$ (on dit que tout v dans E a au moins un *antécédant* (ou *image inverse*) par f dans E).

On dit alors aussi que f est une *surjection*.

Exemple 6.12. *La fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe x^2 n'est pas surjective; en effet, il n'existe pas de $x \in \mathbb{R}$ tel que $-1 = x^2$.*

Exemple 6.13. *La fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe x^3 est surjective; en effet, tout réel a une unique racine cubique. Donc, $\forall v \in R$, il existe au moins un $x \in R$ tel que $f(x) = v$.*

Exemple 6.14. *La fonction de $\mathbb{N}$ dans $\mathbb{N}$ qui à $2n$ et $2n + 1$ associe n est surjective; en effet, tout entier naturel n est l'image par cette fonction de $2n$ (et aussi de $2n + 1$).*

Une fonction est dite *bijective* si elle est injective est surjective. Dans ce cas, on dit que c'est une *bijection*. De plus, la fonction *inverse* (ou *réciproque*) de f, notée f^{-1}, est la fonction de $F \to E$ définie par: $\forall x \in E, \forall y \in F$, $y = f(x)$ si et seulement si $x = f^{-1}(y)$. L'existence et l'unicité de la fonction inverse provient de ce qu'on a: pour tout $y \in F$, il existe un unique $x \in E$ tel que $y = f(x)$.

Exemple 6.15. *La fonction exponentielle est une bijection de $\mathbb{R}$ vers $\mathbb{R}_+^*$; la bijection réciproque est la fonction logarithme.*

Définition 6.4. *Un* isomorphisme *est une application linéaire bijective. S'il existe un isomorphisme de E vers F, on dit que E est F sont* isomorphes*.*

Proposition 6.6. *Soient E, F des espaces vectoriels de dimension finie et $f : E \to F$ une application linéaire. Les conditions suivantes sont équivalentes:*

(i) f est un isomorphisme;

(ii) toute base de E est envoyée sur une base de F;

(iIi) il existe une base de E qui est envoyée par f sur une base de F.

Lemme 6.1. *Soit $f \in L(E, F)$.*

1. Si f est injectif, alors f envoie toute famille de vecteurs linéairement indépendants sur une famille de vecteurs linéairement indépendants.

2. Si f est surjectif, alors f envoie toute famille génératrice sur une famille génératrice.

Proof. 1. Soient $e_1, \dots e_n$ linéairement indépendants dans E. Montrons que $f(e_1) \dots, f(e_n)$ sont linéairement indépendants. Supposons qu'il existe $a_1, \dots, a_n$ dans $\mathbb{R}$ tels que $a_1 f(e_1) + \cdots + a_n f(e_n) = 0$. Alors $f(a_1 e_1 + \cdots + a_n e_n) = 0$. Donc $a_1 e_1 + \cdots + a_n e_n \in \mathrm{Ker}(f)$. Comme f est injective, on a $a_1 e_1 + \cdots + a_n e_n = 0$. Par suite, les a_i sont tous nuls, car les e_i sont linéairement indépendants. On conclut donc que les $f(e_i)$ sont linéairement indépendants.

2. Soient $e_1, \dots e_n$ qui engendrent E. Soit y dans F; comme f est surjective, il existe $x \in E$ tel que $y = f(x)$. Il existe alors $a_1, \dots, a_n$ dans $\mathbb{R}$ tels que $x = a_1 e_1 + \cdots + a_n e_n$. Alors $y = f(x) = f(a_1 e_1 + \cdots + a_n e_n) = a_1 f(e_1) + \cdots + a_n f(e_n)$ et par suite $f(e_1), \dots, f(e_n)$ engendrent F. □

Preuve de la proposition 6.6. (i) implique (ii): on suppose que f est un isomorphisme. Soit $e_1, \dots, e_n$ une base de E. Le lemme 6.1 implique que $f(e_1), \dots, f(e_n)$ est une base de F.

(ii) implique (iii) est évident.

(iii) implique (i): soit $e_1, \dots, e_n$ une base de E telle que $f(e_1), \dots, f(e_n)$ est une base de F. Montrons que f est un isomorphisme. Soit $x \in \mathrm{Ker}(f)$, On peut écrire $x = a_1 e_1 + \cdots + a_n e_n$, $a_i \in F$. Alors $0 = f(x) = a_1 f(e_1) + \cdots + a_n f(e_n)$. Comme les $f(e_i)$ sont linéairement indépendants, les a_i sont tous nuls, donc x est nul. Donc f est injective.

Soit maintenant $y \in F$. On peut écrire $y = a_1 f(e_1) + \cdots + a_n f(e_n)$. Soit $x = a_1 e_1 + \cdots + a_n e_n$. Alors $y = f(x)$. □

On appelle *endomorphisme* de V une application linéaire de l'espace vectoriel V dans lui-même. Un *automorphisme* de V est un endomorphisme de V qui est bijectif.

Corollaire 6.3. *Les conditions suivantes sont équivalentes, pour un endomorphisme f d'un espace vectoriel V de dimension finie:*

(i) f est un automorphisme de V;

(ii) f est injectif;

(iii) f est surjectif.

Proof. Il est clair que (i) implique (ii) et (iii). Supposons que (ii) soit vrai. Soit $e_1, \ldots, e_n$ une base de E. Alors les vecteurs $f(e_1, \ldots, f(e_n)$ sont linéairement indépendants par le lemme 6.1. Ils forment donc une base de V par le théorème 5.3. Donc f est un automorphisme de V, c'est-à-dire (i) est vrai.

Pour prouver l'implication (iii) $\Rightarrow$ (i), c'est analogue. □

Définition 6.5. *Un endomorphisme f de V est* inversible à gauche *(resp.* inversible à droite *s'il existe un endomorphisme g de V (resp. un endomorphisme h de V) tel que $g \circ f = \mathrm{id}_V$ (resp. $f \circ h = \mathrm{id}_V$).*

Il est tout-à-fait analogue au cas des matrices (voir proposition 2.2) de montrer que si f a un inverse à gauche et un inverse à droite, alors les deux inverses sont égaux. On dit alors que f est *inversible*.

Corollaire 6.4. *Soit f un endomorphisme de V, espace vectoriel de dimension finie. Alors f est inversible à gauche si et seulement s'il est inversible à droite.*

Proof. Si f est inversible à gauche, il existe un endomorphisme g tel que $g \circ f = \mathrm{id}$. Alors f est injectif: en effet, si $f(x) = f(y)$, alors $g \circ f(x) = g \circ f(y)$, donc $x = y$. D'après le corollaire 6.3, f est donc un automorphisme, donc inversible à droite.

L'implication réciproque est similaire, en prouvant que f est surjective. □

Exercice 6.4. *Montrer que si F, G sont supplémentaires dans E, alors la fonction $F \times G \to E$, $(x, y) \mapsto x + y$, est un isomorphisme.*

Exercice 6.5. *Prouver les implications réciproques des deux implications du lemme 6.1.*

6.5 Applications linéaires et bases

Rappelons que $L(E, F)$ est un espace vectoriel (proposition 6.2) et que F^n est un espace vectoriel (voir 3.3).

Proposition 6.7. *Soit E un espace vectoriel de dimension finie, de base $(e_1, \ldots, e_n)$. La fonction $L(E, F) \to F^n$, qui à $f \in L(E, F)$ associe le n-uplet $(f(e_1), \ldots, f(e_n)$ est un isomorphisme d'espaces vectoriels.*

Proof. Notons $\psi : L(E, F) \to F^n$ cettte fonction. Elle préserve la somme: en effet, $\psi(f+g) = ((f+g)(e_1), \ldots, (f+g)(e_n)) = (f(e_1)+g(e_1), \ldots, f(e_n)+g(e_n)) = (f(e_1), \ldots, f(e_n)) + (g(e_1), \ldots, g(e_n)) = \psi(f) + \psi(g)$.

De plus, si $a \in \mathbb{R}$, alors $\psi(af) = (af(e_1), \ldots, af(e_n)) = a(f(e_1), \ldots, f(e_n)) = a\psi(f)$. Donc ψ préserve le produit externe. C'est une application linéaire.

Montrons que ψ est un isomorphisme. Si $\psi(f) = 0$, alors les $f(e_i)$ sont tous nuls. Donc $f(x) = 0$ pour tout $x \in E$ (car x est combinaison linéaire des e_i). Donc ψ est injective.

Soit maintenant $(v_1, \ldots, v_n) \in F^n$. Posons $f(\sum_i a_i e_i) = \sum_i a_i v_i$. Comme tout vecteur dans E est de manière unique combinaison linéaire des e_i, f est bien définie sur E. C'est pure routine que de vérifier que f est linéaire. Et on a $\psi(f) = (v_1, \ldots, v_n)$. Donc ψ est surjective. □

Corollaire 6.5. *Si $e_1, \ldots, e_n$ est une base E et si $v_1, \ldots v_n$ sont des vecteurs dans F, il existe une unique application linéaire de E vers F qui envoie chaque e_i sur v_i.*

Corollaire 6.6. *Deux espaces vectoriels de dimension finie sont isomorphes si et seulement s'il ont la même dimension.*

Proof. Si E, F sont isomorphes, il existe un isomorphisme $f : E \to F$. Soit $e_1, \ldots, e_n$ une base de E. Alors $f(e_1), \ldots, f(e_n)$ est une base de F, d'après la proposition 6.6. Donc $\dim(F) = n = \dim(E)$.

Réciproquement, supposons que E, F aient la même dimension. Soient $e_1, \ldots, e_n$ une base de E et $v_1, \ldots, v_n$ une base de F. Il existe d,après le corollaire 6.5, une application linéaire qui envoie chaque e_i sur v_i. D'après la proposition 6.6, c'est un isomorphisme. □

6.6 Matrices des applications linéaires et changement de base

Proposition 6.8. *Soient E, F des espaces vectoriels de dimension p, n, et $e_1, \ldots, e_p$, $v_1, \ldots, v_n$ des bases de ces espaces, respectivement. Il y a isomorphisme entre l'espace des applications linéaires $L(E, F)$ et l'espace des* *matrices $M_{np}(\mathbb{R})$. Cet espace associe à toute application linéaire la matrice $[a_{ij}]$ définie par $f(e_j) = \sum_i a_{ij} v_i$.*

Remarquez que la matrice est bien définie, car $f(e_j)$ est de manière unique combinaions linéaire des v_j. Aussi que la taille de la matrice est $n \times p$ car les lignes correspondent aux éléments de la base v_j et les colonnes à ceux de la base e_i.

Proof. On sait qu'il y a un isomorphisme de $L(E, F)$ vers F^p, par la proposition 6.7 (attention: ici c'est p, là-bas c'est n): l'isomorphisme envoie f sur le p-uplet $(f(e_1), \ldots, f(e_p))$. On fait suivre cet isomorphisme par la fonction qui envoie tout p-uplet $(u_1, \ldots, u_p)$ dans F^p sur la matrice $[a_{ij}]$ définie par les égalités $\forall j = 1, \ldots, p$, $u_j = \sum_{1 \leq i \leq n} a_{ij} v_i$; cette fonction est clairement un isomorphisme. Le produit de ces deux isomorphismes est l'isomorphisme cherché. $\square$

Définition 6.6. *On appelle* matrice de l'application linéaire $f \in L(E, F)$ dans les bases e_i et v_j de E et F respectivement *la matrice définie dans la proposition ci-dessus.*

Proposition 6.9. *Soit $f \in L(E, F)$, $g \in L(F, G)$, et (e_k), (f_j), (g_i) des bases de E, F, G respectivement. Soient A la matrice de f dans les bases (e_k), (f_j), et B la matrice de g dans les bases (f_j), (g_i). Alors la matrice de $g \circ f$ dans les bases (e_i), (g_k) est BA.*

On peut même dire que le produit des matrices a été défini pour que ce résultat soit vrai.

Proof. Soient $A = [a_{jk}], B = [b_{ij}], C = [c_{ik}]$ les matrices de $f, g, g \circ f$ respectivement. On a $f(e_k) = \sum_j a_{jk} f_j$ et $g(f_j) = \sum_i b_{ij} g_i$. Donc $g \circ f(e_k) = g(\sum_j a_{jk} f_j) = \sum_j a_{jk} g(f_j) = \sum_j a_{jk} \sum_i b_{ij} g_i = \sum_i (\sum_j b_{ij} a_{jk}) g_i$. Ceci prouve que $c_{ik} = \sum_j b_{ij} a_{jk}$, donc que $C = BA$. $\square$

Corollaire 6.7. *La matrice de l'inverse d'un isomorphisme est l'inverse de sa matrice.*

Proof. On applique la proposition précédente au cas où $E = G$, $e = g$, $g = f^{-1}$. $\square$

Définition 6.7. *Soient deux bases $(e_i), (e'_i)$ d'un espace vectoriel E. La* matrice de passage de la base de (e_i) vers la base (e'_i) *est la matrice, notée $P_{ee'} = [p_{ij}]$, définie par $e'_j = \sum_i p_{ij} e_i$.*

Définition 6.8. *La* matrice d'un vecteur v dans la base (v_i) *de* V *est le vecteur colonne des coefficients de* v *dans la base* v_i*. C'est-à-dire le vecteur* $(a_, \dots, a_n)$ *où les coefficients* a_i *sont définis par* $v = \sum_i a_i v_i$*.*

Proposition 6.10. *Soient deux base de* E *comme ci-dessus.*

1. Si C, C' *sont les matrices d'un vecteur* x *de* E *dans les bases* (e_i) *et* (e'_i) *respectivement, alors* $C = P_{ee'}C'$*.*

2. L'inverse de $P_{ee'}$ *est* $P_{e'e}$*.*

3. Soient f*,* f' *deux bases de* F*. Soit* $f \in L(E, F)$ *et* A, A' *ses matrices dans les bases* e, f *d'une part, et* e', f' *d'autre part. Alors* $A' = P_{f'f} A P_{ee'}$*.*

Proof. En comparant les définitions 6.7 et 6.6, on observe que $P_{ee'}$ est égal à la matrice de l'identité de E dans les bases (e'_i), (e_i) (dans cet ordre!). L'assertion 2. découle donc de cette observation et du corollaire 6.7.

Pour 1. On observe aussi que la matrice de x est égal à la matrice de l'application linéaire $\mathbb{R} \to E$ qui envoie 1 sur x. On applique la proposition 6.9.

Pour 3. la même observation et la même proposition s'appliquent: la matrice $P_{f'f} A P_{ee'}$ est la matrice de l'application linéaire composée de 3 applications linéaire: $\mathrm{id}_F \circ f \circ \mathrm{id}_E$. □

Définition 6.9. *La* matrice $[a_{ij}]$ d'un endomorphisme *de* V *dans la base* $v_1, \dots, v_n$ *de* V *est définie par* $f(v_j) = \sum_i a_{ij} v_i$*,* $j = 1, \dots, n$*.*

C'est donc un cas particulier de la définition ci-dessus, avec $E = F = V$, avec la même base au départ et à l'arrivée. De même, le corollaire suivant est un cas particulier de la proposition 6.9.

Corollaire 6.8. *Si* A, B *sont les matrices des endomorphismes* f *et* g *respectivement dans une base de* V*, alors la matrice de* $g \circ f$ *dans cette base est* BA*.*

Corollaire 6.9. *Une matrice carrée est inversible à gauche si et seulement si elle inversible à droite.*

Proof. Cela découle du corollaire précédent et de la proposition 6.4. □

7 Déterminants

7.1 Formule de Laplace

Définition 7.1. *Soit* $A = [a_{ij}]$ *une matrice carrée d'ordre* n*. Le* déterminant *de* A*, noté* $\det(A)$ *ou aussi* $|A|$*, est défini comme suit: si* $n = 1$*, c'est* a_{11}*; si* $n \geq 2$*, c'est* $a_{11}\Delta_{11} - a_{21}\Delta_{21} + \dots + (-1)^{n+1} a_{n1}\Delta_{n1}$*, où* Δ_{ij} *est*

Proposition 7.1. *Le déterminant d'une matrice triangulaire est égal au produit de ses éléments diagonaux.*

Proof. Si la matrice est triangulaire supérieure, on a par définition du déterminant $\det(A) = a_{11}\Delta_{11}$. Par hypothèse de récurrence $\Delta_{11} = a_{22}\cdots a_{nn}$. D'où le résultat.

Si la matrice est triangulaire inférieure, alors Δ_{i1} est pour $i \geq 2$ le déterminant d'une matrice triangulaire dont la première ligne est nulle; par hypothèse de récurrence son déterminant est nul. On a donc $\det(A) = a_{11}\Delta_{11}$ et on conclut comme dans le cas précédent. □

Corollaire 7.1. *Soit E une matrice élémentaire, obtenue en appliquant à la matrice identité une opération de ligne (voir la définition 2.4). Alors son déterminant est* 1 *si c'est une opération $l_i + aL_j$, et a si c'est une opération al_i.*

Proof. Dans le premier cas, E est une matrice triangulaire inférieure ou supérieure avec des 1 sur la diagonale. Donc son déterminant est 1 par la proposition précédente.

Dans le deuxième cas, E est une matrice diagonale avec des 1 sur la diagonale, sauf un a quelque part; son déterminant est donc a. □

Proposition 7.2. *Soient A, A', A'' trois matrices carrées qui ne diffèrent que par leur i-ème ligne.*

1. *Si $l_i = l'_i + l''_i$; alors $\det(A) = \det(A') + \det(A'')$.*
2. *Si $l_i = al'_i$, alors $\det(A) = a\det(A')$.*

Proof. □

Proposition 7.3. *Soit B une matrice carrée obtenue de A par une opération de ligne. Si c'est une opération $l_i + al_j$, elles ont même déterminant. Si c'est une opération al_i, on a $\det(B) = a\det(A)$. Si c'est une opération (ij), on a $\det(B) = -\det(A)$.*

Proof. □

Exercice 7.1. *Montrer que le déterminant de la matrice* $\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}$ *est égal à $(b-a)(c-a)(c-b)$. Généraliser.*

Exercice 7.2. *Soient $v = (a, b, c)$ et $v = (a', b', c')$ deux vecteurs non nuls dans $\mathbb{R}^3$. Montrer que si v, v' sont linéairement dépendants, si et seulement si* $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix} = \begin{vmatrix} a & c \\ a' & c' \end{vmatrix} = 0.$

7.2 Formule du produit et déterminant des endomorphismes

Théorème 7.1. *Soient A, B des matrices carrées de même taille. Alors det*$(AB) = \det(A)\det(B)$.

Lemme 7.1. *Si $B = EA$ où A, B, E sont des matrices carrées de même taille, avec E élémentaire alors,* $\det(B) = \det(P)\det(A)$.

Preuve du théorème 7.1. 1. Si A est une matrice élémentaire, alors la formule est vraie par le lemme 7.1.

2. Supposons que A soit inversible. Alors, grâce au corollaire 2.4, A est un produit de k matrices élémentaires, $A = E_1 \cdots E_k$. Si $k = 1$ on conclut grâce à 1. Supposons que $k \geq 2$. On alors $\det(AB) = \det(E_1(E_2 \cdots E_k B)) = \det(E_1)\det(E_2 \cdots E_k B)$ (par 1.) $= \det(E_1)\det(E_2 \cdots E_k)\det(B)$ (par récurrence sur k) $= \det(A)\det(B)$ (par le cas 1).

3, Supposons que A ne soit pas inversible. Alors il existe une matrice inversible (produit de matrice élémentaires) P telle que PA soit échelonnée réduite. Cette matrice n'est pas inversible (sinon A le serait), donc sa dernière ligne est nulle (si elle ne l'était pas, ça serait la matrice identité). Alors la dernière ligne de PAB est nulle. Le déterminant d'une matrice dont la dernière ligne est nulle est nul (multiplier par -1 cette dernière ligne et appliquer le lemme 7.1). Donc $\det(PA) = \det(PAB) = 0$. Mais nous savons déjà que $\det(PA) = \det(P)\det(A)$ et $\det(PAB) = \det(P)\det(AB)$. Comme $\det(P)$ est non nul, car P est produit de matrices élémentaires, dont le déterminant est le produit de leurs déterminants (par 2.), qui sont non nuls par le corollaire 7.1. Donc $\det(A) = 0 = \det(AB)$. D'où $\det(AB) = \det(A)\det(B)$. □

7.3 Inversion des matrices et déterminants

Théorème 7.2. *Une matrice carrée A est inversible si et seulement si son déterminant est non nul. Dans ce cas, l'élément en position i, j de A^{-1} est* $\frac{(-1)^{i+j}\Delta_{ji}}{\det(A)}$.

Lemme 7.2. $\sum_j (-1)^{j+k} a_{ij}\Delta_{jk} = \delta_{ik}\det(A)$.

Exercice 7.3. *Déduire du théorème l'inverse d'une matrice carrée d'ordre 2.*

7.4 Système de Cramer (suite)

Théorème 7.3. *L'unique solution du système de Cramer $AX = B$ est déterminée par: soit D_i le déterminant de la matrice obtenue à partir de A en y remplaçant la i-ème colonne par B; alors $x_i = D_i/\det(A)$.*

8 Diagonalisation

8.1 Valeurs et vecteurs propres d'un endomorphisme

Définition 8.1. *Soit f une endomorphisme de l'espace vectoriel V. On dit $\lambda \in \mathbb{R}$ est une* valeur propre *de V s'il existe un vecteur v non nul dans V tel que $f(v) = \lambda v$. Un tel vecteur est alors appelé* vecteur propre*, et on dit qu'il est* attaché à la valeur propre λ.

Remarquez qu'un vecteur propre est, par définition, toujours non nul.

Exemple 8.1. *Si v non nul est dans le noyau de f, alors v est un vecteur propre, attaché à la valeur propre 0.*

Proposition 8.1. *Un endomorphisme a la valeur propre λ si et seulement si $f - \lambda\,\mathrm{id}$ n'est pas un automorphisme de V.*

Corollaire 8.1. *Un endomorphisme a la valeur propre λ si et seulement le déterminant de $f - \lambda\,\mathrm{id}$ est nul.*

8.2 Polynôme caractéristique

On va parler dans cette partie de matrices carrées dont les coefficients sont des polynômes en la variable x, et de leur déterminants. Cela est un tantinet plus avancé que les déterminants des matrices carrées à coefficients réels, comme vu en section 7.

Pour calculer ces déterminants, on procède par la formule de Laplace, et aussi par les formules usuelles pour les matrices carrées d'ordre 2 ou 3. La principale propriété dont nous aurons besoin est la formule du produit.

Exemple 8.2. *Soit la matrice $M = \begin{bmatrix} x-1 & -2 \\ -3 & x-4 \end{bmatrix}$. Son déterminant est* $(x-1)(x-4) - (-2)(-3) = x^2 - x - 4x + 4 - 6 = x^2 - 5x + 6$.

Définition 8.2. *Soit $A = [a_{ij}]$ une matrice carrée d'ordre n. Son* polynôme caractéristique *est le déterminant de la matrice $xI_n - A$.*

Notons que xI_n désigne la matrice dont tous les coefficients diagonaux valent x, et les autres sont nuls.

Exemple 8.3. *Le polynôme caractéristique de la matrice* $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *est* $x^2 - 5x + 6$*, conformément à l'exemple précédent.*

Proposition 8.2. *Si deux matrices carrées* A, B *sont conjuguées, elles ont même polynôme caractérisque.*

Corollaire 8.2. *Soit* V *un espace vectoriel de dimension finie,* f *une endomorphisme de* V *et* A, B *ses matrices dans deux bases de* V*. Alors ces deux matrices ont même polynôme caractéristique.*

Le corollaire implique la cohérence de la définition qui suit.

Définition 8.3. *Le* polynôme caractéristique d'un endomorphisme *de* V *est le polynôme caractéristique de sa matrice dans une base de* V*.*

Théorème 8.1. *Soit* f *un endomorphisme de* V *de dimension finie. Alors* λ *est une valeur propre de* V *si et seulement si* λ *est une racine de son polynôme caractéristique.*

8.3 Endomorphisme diagonalisable

Définition 8.4. *Un endomorphisme de* V *est dit* diagonalisable *s'il existe une base de* V *formé de vecteurs propres de* f*.*

Proposition 8.3. *Un endomorphisme* f *de* V*, espace vectoriel de dimension finie, est diagonalisable, si et seulement si* V *a une base où la matrice de* f *est diagonale.*

Théorème 8.2. *Soit* f *un endomorphisme d'un espace vectoriel* E *de dimension* n*. Si le polynôme caractéristique de* f *a* n *racines distinctes, alors* f *est diagonalisable.*

Cette condition suffisante, mais elle n'est pas nécessaire, comme le montre l'exemple de la fonction identité d'un espace de dimension au moins 2. De même la fonction nulle.

Proof. f a n valeurs propres $\lambda_1, \ldots, \lambda_n$, qui sont distinctes. Soient $e_1, \ldots, e_n$ des vecteurs attachés respectivement à ces valeurs propres. Montrons que ces vecteurs forment une base de E. Il suffit, en vertu du théorème 5.3, de montrer qu'ils sont linéairement indépendants.

Ecrivons, par l'absurde, qu'une combinaison linéaire non triviale des e_i est nulle. En se restreignant aux e_i dont le coefficient est non nul, on peut écrire $a_1 e_{i_1} + \cdots a_k e_{i_k} = 0$, avec des coefficients non nuls $a_1, \ldots, a_k$ et des indices i_j dans $\{1, \ldots, n\}$ et de plus $i_1 < \ldots < i_k$. On peut supposer que k est le plus petit possible. On ne peut avoir $k = 1$, car les vecteurs propres son non nuls. On a donc $k \geq 2$. Appliquons f à la relation précédente. Ça donne $a_1 \lambda_{i_1} e_{i_1} + a_2 \lambda_{i_2} e_{i_2} + \cdots a_k \lambda_{i_k} e_{i_k} = 0$. Si λ_{i_1} est nul, on obtient une relation plus courte (car les autres λ_j sont alors non nuls), ce qui contredit la minimalité de k. Si λ_{i_1} est non nul, divisons par λ_{i_1}: ça donne $a_1 e_{i_1} + a_2 \frac{\lambda_{i_2}}{\lambda_{i_1}} e_{i_2} + \cdots + a_k \frac{\lambda_{i_k}}{\lambda_{i_1}} e_{i_k} = 0$. Soustrayons la relation ci-dessus: nous obtenons $a_2 (\frac{\lambda_{i_2}}{\lambda_{i_1}} - 1) e_{i_2} + \cdots + a_k (\frac{\lambda_{i_k}}{\lambda_{i_1}} - 1) e_{i_k} = 0$. C'est une relation plus courte (car les λ_j sont distincts, donc les nouveaux coefficients sont non nuls), ce qui contredit la minimalité de k aussi. □

8.4 Diagonalisation des matrices

Définition 8.5. *Soit A une matrice carrée d'ordre n. Soit f l'endomorphisme de l'espace vectoriel des colonnes $V = M_{n1}$ dont M est la matrice. Autrement dit, si $v \in V$, alors $f(v) = Mv$. Les* valeurs propres *(resp.* vecteurs propres*) de M sont les valeurs propres (resp. vecteurs propres) de f. La matrice A est dite* diagonalisable *si f est diagonalisable*

Proposition 8.4. *1. Les valeurs propres de A sont les racines du polynôme caractéristique de A.*

2. Une matrice carrée est diagonalisable si et seulement si elle est conjuguée à une matrice diagonale.

Ceci découle des sections précédentes.

9 Espaces euclidiens

9.1 Produits scalaires et bases othonormales

Définition 9.1. *Soit E une espace vectoriel de dimension finie. Une* produit scalaire *sur E est une fonction de $E \times E$ vers $\mathbb{R}$, qu'on note $(x, y) \mapsto \langle x, y \rangle$, avec les propriétés suivantes:*

- *si on fixe $x \in E$, la fonction $y \mapsto \langle x, y \rangle$ est linéaire;*
- *si on fixe $y \in E$, la fonction $x \mapsto \langle x, y \rangle$ est linéaire;*

- $\forall x, y \in E$, $\langle x, y\rangle = \langle y, x\rangle$;
- $\forall x, y \in E$, $\langle x, y\rangle \geq 0$;
- $\forall x \in E$, $\langle x, x\rangle = 0$ *si et seulement si* $x = 0$.

Un espace euclidien *est un espace vectoriel avec un produit scalaire.*

Exemple 9.1. *Avec* $E = \mathbb{R}^n$, *on définit pour* $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$, $lx, y\rangle = x_1y_1 + \cdots + x_ny_n$. *C'est un produit scalaire sur* E.

Exercice 9.1. *Avec* $E = M_{np}(\mathbb{R})$, *montrer que la fonction* $(A, B) \mapsto \mathrm{Tr}(AB^t)$ *est un produit scalaire.*

9.2 Orthonormalisation de Gram-Schmidt

Définition 9.2. *Une* base orthonormale *d'un espace euclidien est une base* $e_1, \ldots, e_n$ *telle que* $\forall i, j$, *on a* $\langle e_i, e_j\rangle = \delta_{ij}$.

Notez que δ_{ij} vaut 1 si $i = j$ et 0 sinon.

Théorème 9.1. *Soit* $e_1, \ldots, e_n$ *une base de* E *espace euclidien. Il existe une unique base orthonormale* $v_1, \ldots, v_n$ *de* E *telle que:* $\forall j = 1, \ldots, n$, $Vect(e_1, \ldots, e_j) = Vect(v_1, \ldots, v_j)$ *et que* $\langle e_j, v_j\rangle > 0$.

Proof. 1. Construisons d'abord une base orthogonale $u_1, \ldots, u_n$ qui satisfait les conditions de l'énoncé. Posons $u_1 = e_1$. Supposons que $j \leq 2$ et qu'on ait construit $u_1, \ldots, u_{j-1}$ et construisons u_j de la forme $u_j = e_j + a_1e_1 + \ldots + a_{j-1}e_{j-1}$. On veut que $\langle u_j, u_k\rangle = 0$ pour $k = 1, \ldots, j-1$. C'est-à-dire $\langle e_j, u_k\rangle + a_k\langle u_k, u_k\rangle = 0$. Ceci détermine a_k de manière unique. Et on a bien $\langle e_j, u_j\rangle = \langle u_j - a_1e_1 - \ldots - a_{j-1}u_{j-1}, u_j\rangle = \langle u_j, u_j\rangle > 0$.

2. On pose maintenant $v_j = u_j/\sqrt{\langle u_j, u_j\rangle}, j = 1, \ldots, n$.

3. L'unicité se vérifie en suivant la construction précédente. □

Corollaire 9.1. *Un espace euclidien possède une base orthonormale.*

9.3 Diagonalisation des matrices symétriques

Théorème 9.2. *Soit* u *une endomorphisme d'un espace euclidien* E. *Il existe un unique endomorphisme* u^* *de* E *tel que:* $\forall x, y \in E$, $\langle u(x), y\rangle = \langle x, u^*(y)\rangle$. *Dans toute base othonormale de* E, *la matrice de* u^* *est la transposée de celle de* u

Proof. 1. Soit $e_1, \ldots, e_n$ une base orthonormale de E. Définissons u^* par: (*) $u^*(e_j) = \sum_i \langle u(e_i), e_j\rangle e_i$. La matrice de u^* dans cette base est $[\langle u(e_i), e_j\rangle]_{ij}$. Celle de u est $[m_{ij}]$ avec $u(e_j) = \sum_i m_{ij} e_i$. Par orthonormalité, on a $\langle u(e_j), e_i\rangle = m_{ij}$. Donc les matrices de u et u^* sont transposées l'une de l'autre.

2. Considérons les deux fonctions $E \times E \to \mathbb{R}$: $(x, y) \mapsto \langle u(x), y\rangle$ et $(x, y) \mapsto \langle x, u^*(y)\rangle$. Pour $x = e_i$ et $y = e_j$, on a $\langle u(e_i), e_j\rangle = \langle u^*(e_j), e_i\rangle$, en utulisant (*). Ceci implique que les deux fonction considérées coïncident sur la base. Il s'ensuit qu'elles sont égales (voir exercice **??**). D'où l'existence et l'unicité de u^* et l'assertion sur les matrices.

Comme l'égalité dans le théorème est indépendant des bases, on conclut que les matrices de u et $u*$ sont transposées dans toute base orthonormale. □

Définition 9.3. *1. On appelle* adjoint *de u l'endomorphisme, noté u^*, défini dans le théorème précédent.*

2. On dit que u est symétrique *si $u = u^*$.*

Corollaire 9.2. *u est symétrique si et seulement sa matrice dans une base orthonormale est symétrique (et alors elle l'est dans toutes).*

Théorème 9.3. *Si u est un endomorphisme symétrique de E euclidien, il existe une base orthonormale de E où la matrice de u est diagonale.*

Preuve par récurrence sur la dimension de E. Soit λ une valeur propre complexe de E. On va montrer que $\lambda \in \mathbb{R}$. Il existe un vecteur colonne sur $\mathbb{C}$ tel que $Mv = \lambda v$. On a donc $M\bar{v} = \bar{\lambda}\bar{v}$. Donc $\bar{\lambda} v^t \bar{v} = v^t(\bar{\lambda}\bar{v}) = v^t M \bar{v} = (v^t M \bar{v})^t = \bar{v}^t M v = \bar{v}^t \lambda v = \lambda \bar{v}^t v$.

Ecrivons $v = (a_1, \ldots, a_n)^t$. Alors $v^t \bar{v} = \sum_i |a_i|^2 = \bar{v}^t v$. Ce nombre est nul, car $v \neq 0$, et par suite $\bar{\lambda} = \lambda$. Donc λ est réel.

Soit $v \in \mathbb{R}^n$ un vecteur propre de M pour la valeur propre λ. Soit D la droite $\mathbb{R}v$ et $F = \{x \in \text{E}, \langle x, v\rangle = 0\}$. Nous montrons que $u(F) \subset F$: si $x \in F$, alors $\langle u(x), v\rangle = \langle x, u(v)\rangle = \langle x, \lambda v\rangle = \lambda\langle x, v\rangle = 0$; donc $u(x) \in F$. Donc u est un endomorphisme de F, qui est, par restriction du produit scalaire de E, un espace euclidien. De plus, la restriction de u à F est symétrique. Il existe donc, par hypothèse de récurrence une base orthonormale de F où la matrice de u est orthonormale. En complétant par v cette base, on obtient une base orthonormale de E, où la matrice de u est diagonale. □

Corollaire 9.3. *Une matrice symétrique est diagonalisable et toutes ses valeurs propres sont réelles.*

Proof. On considère l'endomorphisme de $E = \mathbb{R}^n$ qui a M comme matrice dans la base canonique. E est un espace euclidien avec le produit scalaire usuel. Par le corollaire 9.2, c'est un endomorphisme symétrique. On conclut par le théorème 9.3. $\square$

References

[1] P. Leroux, Algèbre linéaire, une approche matricielle, Modulo, 1978.

[2] F. Liret, D. Martinais, Algèbre 1ère année, Dunod, 2003.