Algèbre linéaire

Sciences pour l'Ingénieur L2 (1er semestre, 2^{e} année)

72h de cours/TD

Notes rédigées par Alexey Muranov

1er avril 2025

Préface

Ces notes sont basées sur un cours enseigné à l'Université d'Aix-Marseille dans les années 2013–2017 aux étudiants de la deuxième année de Licence « Sciences pour l'Ingénieur ».

En fonction des connaissances de l'audience, des interactions avec l'audience (beaucoup de questions ou peu de questions) et du style d'enseignement (exemples intégrés dans le cours ou laissés pour les TDs), la partie théorique du cours est estimée de prendre entre 24 et 36 heures. Cependant, certaines parties de ces notes contiennent plus de la matière que ce que serait possible à inclure dans le cours sur lequel les notes sont basées.

Ce cours ne suit aucun livre, mais peut-être une majorité de son contenu est couvert par un n'importe quel livre classique d'algèbre linéaire, comme, par exemple, [1, 2, 3, 4].

Table des matières

I. Systèmes d'équations linéaires[3h]

I.1. Exemples et définitions de base

I.1.1. Équations linéaires

Définition. Une *équation linéaire* en variables $x_1, \dots, x_n$ est une équation de la forme

$$a_1x_1 + \cdots + a_nx_n = b,$$

où $a_1, \dots, a_n, b \in \mathbf{R}$. Le nombre b dans cette équation s'appelle le *second membre*. Si $b = 0$, on dit que l'équation est *homogène* (ou *sans second membre*).

Définition. Un *système d'équations linéaires* en variables $x_1, \dots, x_n$ est un système d'équations linéaires en variables $x_1, \dots, x_n$, c'est-à-dire, un système de la forme

$$\begin{cases} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{cases}$$

où $a_{i,j}, b_i \in \mathbf{R}$. Les nombre $b_1, \dots, b_m$ dans ce système s'appellent les *seconds membres*. Si $b_1 = \cdots = b_m = 0$, on dit que le système est *homogène*.

Observation : un système homogène admet toujours la solution nulle.

I.1.2. Résolution d'un système d'équations linéaires sans méthode particulière

Écrire et résoudre quelque systèmes d'équations linéaires par toute méthode qui semble bonne.

I.1.3. Introduction à la méthode d'élimination de Gauss sur un exemple

Écrire et résoudre un système d'équations linéaires par la méthode d'élimination de Gauss.

I.1.4. Les cas de R, C, Q

Parler de la différence entre les systèmes d'équations linéaires à coefficients dans $\mathbf{R}$, à coefficients dans $\mathbf{C}$, et à coefficients dans $\mathbf{Q}$ (qu'il n'y en a pas).

I.2. Le système homogène associé

I.2.1. Définitions et exemples

Définition. Soit (E) une équation linéaire :

$$a_1x_1 + \cdots + a_nx_n = b. \qquad (E)$$

L'équation *homogène associée* à (E) est l'équation

$$a_1x_1 + \cdots + a_nx_n = 0 \qquad (E_0)$$

(où les nombres $a_1, \dots, a_n$ sont les mêmes que dans (E)).

Définition. Soit (S) un système d'équations linéaires :

$$\begin{cases} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{cases} \qquad (S)$$

Le système *homogène associé* à (S) est le système

$$\begin{cases} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & 0 \end{cases} \qquad (S_0)$$

(avec les mêmes $a_{i,j}$).

I.2.2. Propriétés

Proposition. *Soient (S) un système d'équations linéaires et (S_0) son système homogène associé. Alors*

(1) *une « combinaison linéaire » de solutions de (S_0) est une solution de (S_0),*

(2) *la somme d'une solution de (S) avec une solution de (S_0) est une solution de (S),*

(3) *la différence de deux solutions de* (S) *est une solution de* (S_0).

L'idée d'une démonstration : additionner deux systèmes d'égalités obtenues en remplaçant les variables par des solutions, ou soustraire un système de l'autre, etc.

I.3. L'algorithme d'élimination de Gauss

I.3.1. Systèmes échelonnés

Soit (S) un système d'équations linéaires où on a choisi et « fixé » l'ordre des variables (ainsi que l'ordre des équations). On va supposer que les équations sont ordonnées de haut en bas, et que les variables dans chaque équation apparaissent dans l'ordre de gauche à droit :

$$\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}\right. \qquad (S)$$

Définition. On dit que (S) est *échelonné* si et seulement si les conditions suivantes sont satisfaites[1] :

(1) s'il y a des équation de la forme « $0 = b$ », alors elles sont toutes regroupées à la fin du système (en bas),

(2) si une variable apparaît comme la première variable (avec coefficient non nul) dans une équation, alors ni cette variable, ni les variables précédentes, n'apparaissent pas dans les équations suivantes.

Définition. Si (S) est échelonné, alors le premier (à gauche) coefficient non nul de chaque équation de (S) s'appelle le *pivot* de cette équation, et la variable auprès de laquelle il se trouve s'appelle la *variable pivotale* de cette équation. Les variables qui ne sont pas pivotales s'appellent parfois *variables libres*.

Définition. Un système échelonné est dit *échelonné réduit* si et seulement si tous les pivots sont 1, et chaque variable pivotale n'apparaît que dans une seule équation.

[1] Cette notion d'un *système échelonné* n'est pas vraiment standard. En revanche, la notion d'une *matrice échelonnée*, à voir plus tard, est assez standard.

Exemple. Le système suivant est échelonné (par rapport à l'ordre des variables x, y, z, w, t) :

$$\left\{\begin{array}{ccccccccccc} \boldsymbol{x} & + & 2y & + & 3z & + & 4w & & & = & 1 \\ & & \boldsymbol{5y} & + & 6z & & & + & 7t & = & 2 \\ & & & & & & \boldsymbol{8w} & + & 9t & = & 3 \end{array}\right.$$

Ici, les variables z et t sont libres.

Le système suivant est échelonné réduit (par rapport à l'ordre des variables x, y, z, w, t) :

$$\left\{\begin{array}{ccccccccc} \boldsymbol{x} & & + & 2z & & + & 3t & = & 1 \\ & \boldsymbol{y} & + & 4z & & + & 5t & = & 2 \\ & & & & \boldsymbol{w} & & & = & 3 \end{array}\right.$$

Les variables z et t sont libres.

Remarque. On peut montrer que tout système d'équations linéaires qui admet au moins une solution est équivalent à un *unique* système échelonné réduit — par rapport à l'ordre choisi des variables, et en supposant que les équations de la forme « $0 = 0$ » sont supprimées à la fin. (Pour démontrer l'unicité, on peut utiliser la récurrence sur le nombre des variables.)

I.3.2. Opérations élémentaires

Opérations élémentaires de la méthode d'élimination de Gauss :

(1) $E_i \leftarrow E_i + \alpha E_j$, $i \neq j$,

(2) $E_i \leftarrow \alpha E_i$, $\alpha \neq 0$,

(3) $E_i \leftrightarrow E_j$, $i \neq j$.

Ces opérations sont toutes inversibles : pour chacune de ces opérations, il y a une opération inverse du même type.

Proposition. *Les opérations élémentaires transforment un système donné en un système équivalent.*

En pratique, lorsque on cherche à simplifier un système, en plus d'appliquer ces trois type d'opération on s'autorise d'habitude de supprimer des équations de la forme « $0 = 0$ ». Plus généralement, si une équation résulte des autres, on peut la supprimer sans que cela aurait changé l'ensemble des solutions.

I.3.3. Présentation informelle de l'algorithme

L'*algorithme d'élimination de Gauss* est un algorithme pour résoudre un système quelconque d'équations linéaires. On effectue cet algorithme en deux étapes.

(1) D'abord on transforme le système donné en un système échelonné par des opérations élémentaires (le système échelonné ainsi obtenu sera équivalent au système initial).

(2) Si en effectuant l'étape précédente on a obtenu une équation « absurde » de la forme « $0 = b$ » avec $b \neq 0$, alors on conclut qu'il n'y a pas de solutions. Sinon, on prend des paramètres différents comme les valeurs des variables libres, et on trouve ensuite une par une les valeurs des variables pivotales en commençant par la dernière et en « remontant » vers la première.

Il y a une version légèrement modifiée de cet algorithme qui s'appelle l'*algorithme de Gauss-Jordan*. Dans cette version, après avoir obtenu un système échelonné, au lieu de chercher la solution générale par la « remontée », on continue à transformer le système par des opérations élémentaires jusqu'à obtenir un système échelonné réduit.

I.3.4. Conséquences sur le nombre de solutions

Si l'on étudie toutes les possibilités qu'on peut rencontrer en appliquant la méthode d'élimination de Gauss, on peut constater qu'il n'y a que trois possibilités pour l'ensemble des solutions : il peut

(1) être vide (si l'on déduit une équation « absurde » de la forme « $0 = b$ » avec $b \neq 0$),

(2) avoir un seul élément (si aucune équation « absurde » n'a été déduite, alors que toutes les variables dans le système échelonné sont pivotales), ou

(3) être infini (si aucune équation « absurde » n'a été déduite, alors que dans le système échelonné il y a au moins une variable libre).

Théorème. *Un système d'équations linéaires soit n'admet aucune solution, soit admet une solution unique, soit admet une infinité des solutions.*

I.4. Matrices de systèmes et vocabulaire « vectoriel »

Soit (S) un système d'équations linéaires :

$$\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}\right. \qquad (S)$$

Définition. La *matrice* de (S) est le tableau des coefficients :

$$\begin{array}{ccc} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{array}$$

Le *vecteur second membre* de (S) est le m-uple $(b_1, \ldots, b_m)$, souvent écrit en colonne :

$$\begin{array}{c} b_1 \\ \vdots \\ b_m \end{array}$$

La *matrice augmentée* de (S) est la matrice de taille $m \times (n+1)$ obtenue en juxtaposant la matrice du système avec le vecteur second membre écrit en colonne, comme suit :

$$\begin{array}{ccccc} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & b_m \end{array}$$

Si

$$\left\{\begin{array}{ccc} x_1 & = & s_1 \\ & \vdots & \\ x_n & = & s_n \end{array}\right.$$

est une solution de (S), alors le n-uple $(s_1, \ldots, s_n)$ s'appelle un *vecteur solution* de (S) ; il est souvent écrit en colonne :

$$\begin{array}{c} s_1 \\ \vdots \\ s_n \end{array}$$

Lorsque on effectue une élimination de Gauss, on peut au lieu de systèmes utiliser leurs matrices augmentées.

Définition. Une matrice est dite *échelonnée* si et seulement si les conditions suivantes sont satisfaites pour elle :

(1) si une ligne est nulle, alors toutes les lignes suivantes sont nulles,

(2) si le premier élément non zéro d'une ligne se trouve dans la colonne numéro j, alors tous les éléments dans les colonnes des numéros de 1 à j dans toutes les lignes suivantes sont zéro.

Pour toute ligne non nulle d'une matrice échelonnée, son premier élément non zéro s'appelle le *pivot* de cette ligne, et la colonne où se trouve ce pivot s'appelle la *colonne pivotale* de cette ligne.

Définition. Une matrice échelonnée est dite *échelonnée réduite* si et seulement si tous les pivots sont 1, et dans chaque colonne pivotale tous les élément sauf un (le pivot) sont zéro.

Exemple. Soient

$$A = \begin{pmatrix} \mathbf{1} & 2 & 3 & 4 & 0 \\ 0 & \mathbf{5} & 6 & 0 & 7 \\ 0 & 0 & 0 & \mathbf{8} & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} \mathbf{1} & 0 & 2 & 0 & 3 \\ 0 & \mathbf{1} & 4 & 0 & 5 \\ 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

La matrice A est échelonnée. La matrice B est échelonnée réduite.

II. Espaces vectoriels réels[5h]

II.1. Définitions et propriétés de base

Définition. Un *espace vectoriel réel* est un ensemble E non vide avec deux opération :

- une addition des éléments de E

$$u, v \mapsto u + v \in E, \quad \text{pour} \quad u, v \in E,$$

- une multiplication des éléments de E par les nombres réels

$$\alpha, u \mapsto \alpha u \in E, \quad \text{pour} \quad \alpha \in \mathbf{R}, u \in E.$$

Ce n'est pas tout : ils doivent satisfaire les 8 axiomes suivants.

(1) Pour tous $u, v, w \in E$,

$$(u + v) + w = u + (v + w).$$

(2) Pour tous $u, v \in E$,

$$u + v = v + u.$$

(3) Il existe un élément $\vec{0} \in E$ tel que pour tout $u \in E$,

$$u + \vec{0} = \vec{0} + u = u.$$

(4) Pour tout $u \in E$, il existe un élément dans E noté $-u$ tel que

$$u + (-u) = (-u) + u = \vec{0}.$$

(5) Pour tous $\alpha \in \mathbf{R}$ et $u, v \in E$,

$$\alpha(u + v) = \alpha u + \alpha v.$$

(6) Pour tous $\alpha, \beta \in \mathbf{R}$ et $u \in E$,

$$(\alpha + \beta)u = \alpha u + \beta u.$$

(7) Pour tous $\alpha, \beta \in \mathbf{R}$ et $u \in E$,

$$(\beta\alpha)u = \beta(\alpha u).$$

(8) Pour tout $u \in E$,

$$1u = u.$$

Éléments des espaces vectoriels s'appellent *vecteurs*.

On va considérer uniquement les espaces vectoriels *réels*, donc on ne va pas dire « réel » chaque fois.

Exemples. (1) $\mathbf{R}$, $\mathbf{R}^n$, $\mathbf{C}$, $\mathbf{C}^n$, $\{0\}$.

(2) Fonctions réelles sur un ensemble donné A : $\mathbf{R}^A$.

(3) Fonctions réelles continues sur un intervalle réel I : $C(I, \mathbf{R})$.

(4) Polynômes à coefficients réels ou complexes : $\mathbf{R}[X]$, $\mathbf{C}[X]$.

(5) Solutions d'un système d'équations linéaires homogènes.

Proposition. *Soit E un espace vectoriel. Alors*

(1) *il existe un seul vecteur $\vec{0}$ dans l'axiome* (3),

(2) *pour tout vecteur u, il existe un seul vecteur $-u$ dans l'axiome* (4),

(3) *pour tout $\alpha \in \mathbf{R}$, $\alpha\vec{0} = \vec{0}$,*

(4) *pour tous $u \in E$ et $\alpha \in \mathbf{R}$, $\alpha(-u) = -\alpha u$,*

(5) *pour tout $u \in E$, $0u = \vec{0}$,*

(6) *pour tous $u \in E$ et $\alpha \in \mathbf{R}$, $(-\alpha)u = -\alpha u$.*

Définition. L'élément $\vec{0}$ s'appelle l'*élément neutre*. Pour tout u, l'élément $-u$ s'appelle l'*opposé* de u.

Définition. Une *combinaison linéaire* de vecteurs $v_1, \ldots, v_n$ est tout vecteur de la forme $a_1v_1 + \cdots + a_nv_n$ avec $a_1, \ldots, a_n \in \mathbf{R}$.

Observation : une combinaison linéaire de combinaisons linéaires des vecteurs d'un ensemble S est encore une combinaison linéaire des vecteurs de S.

II.2. Sous-espaces vectoriels

Définition. Soit E un espace vectoriel réel. Une partie F de E s'appelle un *sous-espace vectoriel* si et seulement si

(1) $\vec{0}_E \in F$,

(2) pour tout $u \in F$, $-u \in F$,

(3) pour tous $\alpha \in \mathbf{R}$ et $u \in F$, $\alpha u \in F$,

(4) pour tous $u, v \in F$, $u + v \in F$.

Observation : un sous-espace vectoriel est lui-même un espace vectoriel.

Proposition. *Une partie F d'un espace vectoriel E est un sous-espace vectoriel si et seulement si*

(1) $F \neq \varnothing$,

(2) *pour tous $\alpha \in \mathbf{R}$ et $u \in F$, $\alpha u \in F$,*

(3) *pour tous $u, v \in F$, $u + v \in F$.*

Définition. Si A et B sont deux parties d'un espace vectoriel E, on définit leur *somme* $A + B$ par

$$A + B \stackrel{\text{déf}}{=} \{\, a + b \mid a \in A,\ b \in B \,\}.$$

Proposition. *La somme de deux sous-espaces vectoriels de E est aussi un sous-espace vectoriel de E.*

Observons que pour tous sous-espaces vectoriels U, V, W d'un espace vectoriel E, on a :

$$U + (V + W) = (U + V) + W.$$

Ainsi, on peut écrire « $U + V + W$ » (sans parenthèses) pour désigner l'espace vectoriel $U + (V + W) = (U + V) + W$.

Définition. La *somme directe* (interne) de deux sous-espaces vectoriels F et G, notée $F \oplus G$, est la même chose que la somme $F + G$ à la condition que $F \cap G = \{\vec{0}\}$, sinon, il n'y a pas de somme directe (interne) de F et G. Ainsi,

$$F \oplus G \stackrel{\text{déf}}{=} F + G \quad \text{si} \quad F \cap G = \{\vec{0}\}.$$

Si $F \oplus G = E$, alors G est dit un sous-espace *supplémentaire* de F dans E.

Proposition. *Soient U, V, W trois sous-espaces vectoriels d'un espace vectoriel E. Alors la somme directe composée $U \oplus (V \oplus W)$ est définie si et seulement si $(U \oplus V) \oplus W$ est définie, et lorsque les deux sont définies, elle sont égales.*

Grâce à cette proposition, on peut écrire « $U \oplus V \oplus W$ » au lieux de « $U \oplus (V \oplus W)$ » et de « $(U \oplus V) \oplus W$ ».

Esquisse d'une démonstration. On peut montrer que l'existence de la somme directe composée $U \oplus (V \oplus W)$ équivaut à la condition suivante :

$$\text{si } u \in U, v \in V, w \in W, \text{ et que } u+v+w = \vec{0}, \text{ alors } u = v = w = \vec{0}.$$

On peut également montrer que cette condition est équivalente à l'existence de la somme directe composée $(U \oplus V) \oplus W$. Ainsi, $U \oplus (V \oplus W)$ est définie si et seulement si $(U \oplus V) \oplus W$ est définie.

Si $U \oplus (V \oplus W)$ et $(U \oplus V) \oplus W$ sont définies, alors

$$U \oplus (V \oplus W) = U + (V + W) = (U + V) + W = (U \oplus V) \oplus W. \qquad \square$$

II.3. Familles de vecteurs

Rappel : différence entre familles et ensembles

II.3.1. Familles génératrices

Théorème. *Soit S un ensemble ou une famille de vecteurs dans un espace vectoriel E. Alors l'ensemble des combinaisons linéaires de vecteurs de S est un sous-espace vectoriel de E. En plus, c'est le plus petit sous-espace qui contient tous les éléments de S.*

Définition. Le sous-espace composé de toutes les combinaisons linéaires de vecteurs d'une famille (ou d'un ensemble) S dans un espace vectoriel E est dit le *sous-espace* de E *engendré* par S.

Le sous-espace de E engendré par S est noté $\langle S \rangle$, ou $\langle S \rangle_E$, ou $\mathrm{Vect}(S)$, ou $\mathrm{Vect}_E(S)$.

Définition. Une famille $\mathcal{F}$ dans un espace vectoriel E est dite *génératrice* de E si et seulement si $\mathrm{Vect}(\mathcal{F}) = E$.

Proposition. *Les opérations des types suivants appliquées à une famille de vecteurs ne changent pas le sous-espace engendré par cette famille* :

(1) $E_i \leftarrow E_i + \alpha E_j$, $i \neq j$,

(2) $E_i \leftarrow \alpha E_i$, $\alpha \neq 0$,

(3) $E_i \leftrightarrow E_j$, $i \neq j$,

(4) *ajouter ou supprimer un élément qui est une combinaison linéaire des autres* (*par exemple* $\vec{0}$).

II.3.2. Dépendances linéaires

Définition. Une *dépendance linéaire* entre n vecteurs $u_1, \ldots, u_n$ est un élément $(a_1, \ldots, a_n)$ de $\mathbf{R}^n$ tel que

$$a_1 u_1 + \cdots + a_n u_n = \vec{0}.$$

La dépendance linéaire nulle $(0, \ldots, 0)$ est dite *triviale* (car elle est toujours présente : $0u_1 + \cdots + 0u_n = \vec{0}$).

Définition. Une famille de vecteurs $(u_1, \ldots, u_n)$ est dite *liée* si et seulement si elle admet une dépendance linéaire non triviale (elle est « liée » par ses dépendances linéaires non triviales). Si $(u_1, \ldots, u_n)$ est liée, on dit aussi que les vecteurs $u_1, \ldots, u_n$ sont *linéairement dépendants*.

Définition. Une famille de vecteurs $(u_1, \ldots, u_n)$ est dite *libre* si et seulement si elle n'est pas liée, c'est-à-dire, s'il n'y a pas de dépendances linéaires non triviales entre $u_1, \ldots, u_n$. On dit aussi dans ce cas que les vecteurs $u_1, \ldots, u_n$ sont *linéairement indépendants*.

Une famille infinie $\mathcal{F}$ est dite *liée* si et seulement si elle possède une sous-famille finie liée ; $\mathcal{F}$ est dite *libre* si et seulement si toute sa sous-famille finie est libre.

II.3.3. Bases

Théorème. *Si S est un ensemble de vecteurs dans un espace vectoriel et $\mathcal{F}$ est une famille libre maximale d'éléments de S, alors*

$$S \subset \mathrm{Vect}(\mathcal{F}).$$

Corollaire. *Toute famille libre maximale dans un espace vectoriel est génératrice de cet espace.*

Théorème. *Toute famille génératrice minimale d'un espace vectoriel est libre.*

Théorème. *Soit $(u_1, \dots, u_n)$ une famille de vecteurs dans un espace vectoriel E. Alors les propriétés suivantes sont équivalentes* :

(1) *$(u_1, \dots, u_n)$ est une famille libre maximale,*

(2) *$(u_1, \dots, u_n)$ est une famille génératrice de E minimale,*

(3) *$(u_1, \dots, u_n)$ est une famille génératrice de E libre.*

Démonstration. Les implications (1) $\Rightarrow$ (3) et (2) $\Rightarrow$ (3) résultent de deux théorèmes précédents.

Les implications (3) $\Rightarrow$ (1) et (3) $\Rightarrow$ (2) sont faciles à vérifier. □

Définition. Une *base* d'un espace vectoriel est une famille génératrice libre de cet espace.

Exemple. Posons

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

n éléments de $\mathbf{R}^n$. Il est facile à vérifier que la famille $(e_1, \dots, e_n)$ est libre et génératrice de $\mathbf{R}^n$. Donc elle est une base de $\mathbf{R}^n$. Elle s'appelle la *base canonique* de $\mathbf{R}^n$.

Proposition. *Une famille $\mathcal{B} = (u_1, \dots, u_n)$ dans un espace vectoriel E est une base de E si et seulement si pour tout $v \in E$, il existe un unique n-uple $(x_1, \dots, x_n) \in \mathbf{R}^n$ tel que*

$$v = x_1 u_1 + \cdots + x_n u_n.$$

Les nombres $x_1, \dots, x_n$ dans cette proposition s'appellent les *coordonnées* de v dans la base $\mathcal{B}$, et on va écrire

$$[v]_{\mathcal{B}} \overset{\text{déf}}{=} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Théorème. *Tout espace vectoriel a une base (finie ou infinie). En plus, toute famille libre peut être « complétée » en une base, et de toute famille génératrice une base peut être extraite.*

Ce théorème est facile a démontrer si l'espace en question est engendré par une famille finie de vecteurs. Dans le cas contraire, les démonstrations reposent sur l'*axiome du choix* de la théorie des ensembles.

II.3.4. Matrices de passage

Définition. Si $\mathcal{B} = (e_1, \dots, e_m)$ et $\mathcal{C} = (f_1, \dots, f_n)$ sont deux bases d'un espace vectoriel, alors la *matrice de passage* de $\mathcal{B}$ à $\mathcal{C}$ (ou la *matrice de changement de base* de $\mathcal{B}$ à $\mathcal{C}$) est la matrice avec la famille des colonnes $([f_1]_{\mathcal{B}}, \dots, [f_n]_{\mathcal{B}})$.

Ainsi, si $(e_1, \dots, e_n)$ et $(f_1, \dots, f_n)$ sont deux bases d'un espace vectoriel, et

$$\begin{cases} f_1 &= a_{1,1}e_1 + \cdots + a_{m,1}e_m, \\ &\vdots \\ f_n &= a_{1,n}e_1 + \cdots + a_{m,n}e_m, \end{cases}$$

alors la matrice de passage de $(e_1, \dots, e_m)$ à $(f_1, \dots, f_n)$ est cette matrice :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

Dans la prochaine section on verra qu'ici nécessairement $m = n$.

II.4. Le rang et la dimension

Le théorème suivant va nous permettre de définir le *rang* d'une famille de vecteurs et la *dimension* d'un espace vectoriel.

Théorème. *Si $\{v_1, \dots, v_n\} \subset \mathrm{Vect}(u_1, \dots, u_m)$ et $n > m$, alors la famille $(v_1, \dots, v_n)$ est liée.*

Démonstration. On cherche à montrer qu'il existe $x_1, \dots, x_n \in \mathbf{R}$ pas tous zéro tels que

$$x_1 v_1 + \cdots + x_n v_n = \vec{0}. \tag{E}$$

L'idée est d'utiliser le fait qu'un système de m équations linéaires homogènes avec n inconnues admet (une infinité) de solutions non nulles, car $n > m$.

Décomposons $v_1, \dots, v_n$ en combinaisons linéaires de $u_1, \dots, u_m$:

$$\begin{cases} v_1 &= a_{1,1}u_1 + \cdots + a_{m,1}u_m, \\ &\vdots \\ v_n &= a_{1,n}u_1 + \cdots + a_{m,n}u_m, \end{cases}$$

Alors, pour tous $x_1, \dots, x_n \in \mathbf{R}$,

$$\begin{aligned} &x_1 v_1 + \cdots + x_n v_n \\ &\quad = x_1(a_{1,1}u_1 + \cdots + a_{m,1}u_m) + \cdots + x_n(a_{1,n}u_1 + \cdots + a_{m,n}u_m) \\ &\quad = (a_{1,1}x_1 + \cdots + a_{1,n}x_n)u_1 + \cdots + (a_{m,1}x_1 + \cdots + a_{m,n}x_n)u_m \end{aligned}$$

Ainsi, pour trouver une solution non nulle de l'équation (E), il suffit de trouver une solution non nulle du système

$$\left\{ \begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & 0 \end{array} \right. \qquad (S)$$

Une solution non nulle de (S) existe car le système est homogène et le nombre de variables n est plus grand que le nombre d'équations m. □

Corollaire. *Si $(u_1, \dots, u_m)$ est une famille génératrice d'un espace vectoriel E et $(v_1, \dots, v_n)$ est une famille libre dans E, alors $m \geqslant n$.*

Corollaire. *Si $(u_1, \dots, u_m)$ et $(v_1, \dots, v_n)$ sont deux familles libres maximales dans un même ensemble (ou famille), alors $m = n$.*

Corollaire. *Si $(u_1, \dots, u_m)$ et $(v_1, \dots, v_n)$ sont deux bases d'un espace vectoriel, alors $m = n$.*

Définition. Le *rang* d'une famille ou d'un ensemble de vecteurs S, noté $\operatorname{rk} S$ (en anglais) ou $\operatorname{rang} S$ (en français), est le nombre d'éléments dans n'importe quelle famille libre maximale d'éléments de S.

Définition. La *dimension* d'un espace vectoriel E, notée $\dim E$, est le nombre d'éléments dans n'importe quelle base de E.

Si E est un espace vectoriel, alors

$$\dim E = \operatorname{rk} E.$$

Si S est un ensemble ou une famille dans un espace vectoriel, alors

$$\operatorname{rk} S = \dim \operatorname{Vect} S.$$

Comme la base canonique de $\mathbf{R}^n$ compte n éléments, on conclut que

$$\dim \mathbf{R}^n = n.$$

Proposition. *Si E est un espace vectoriel de dimension finie n, alors toute famille libre dans E a au plus n éléments, et toute famille génératrice de E a au moins n éléments. En plus, toute famille libre à n éléments est une base, et toute famille génératrice à n éléments est une base.*

Proposition. *Si E est un espace vectoriel de dimension finie n, F est un sous-espace de E, et $F \neq E$, alors* $\dim F \leqslant n - 1$.

Démonstration. Comme F est un sous-espace de E, $\dim F \leqslant \dim E = n$. Soient $(f_1, \dots, f_k)$ une base de F et $v \in E \backslash F$. Alors $\dim F = k$ et $(f_1, \dots, f_k, v)$ est libre. Donc $n = \dim E \geqslant k + 1 = \dim F + 1$. □

II.5. Le rang d'une matrice

Exemple. Soit

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

Soient u_1, u_2, u_3, u_4 les colonnes de A et v_1, v_2, v_3 les lignes de A. Alors

$$\operatorname{rk}(u_1, u_2, u_3, u_4) = 2 = \operatorname{rk}(v_1, v_2, v_3).$$

Proposition. *Si $(u_1, \dots, u_n)$ est une famille de vecteurs dans un espace vectoriel E et $(v_1, \dots, v_n)$ est une famille de vecteurs dans un espace vectoriel F, et que les équations*

$$x_1 u_1 + \cdots + x_n u_n = \vec{0} \quad \textit{et} \quad x_1 v_1 + \cdots + x_n v_n = \vec{0}$$

ont le même ensemble des solutions pour $(x_1, \dots, x_n)$ (c'est-à-dire, les dépendances linéaires entre $u_1, \dots, u_n$ sont les mêmes qu'entre $v_1, \dots, v_n$), alors les deux familles ont le même rang :

$$\operatorname{rk}(u_1, \dots, u_n) = \operatorname{rk}(v_1, \dots, v_n).$$

Par exemple, si les deux systèmes

$$\left\{ \begin{array}{ccc} a_{1,1}x_1 + \cdots + a_{1,n}x_n & = & 0 \\ \vdots & & \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n & = & 0 \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{ccc} b_{1,1}x_1 + \cdots + b_{1,n}x_n & = & 0 \\ \vdots & & \\ b_{q,1}x_1 + \cdots + b_{q,n}x_n & = & 0 \end{array} \right.$$

sont équivalents, alors

$$\operatorname{rk}\left(\begin{bmatrix}a_{1,1}\\ \vdots \\ a_{p,1}\end{bmatrix},\dots,\begin{bmatrix}a_{1,n}\\ \vdots \\ a_{p,n}\end{bmatrix}\right)=\operatorname{rk}\left(\begin{bmatrix}b_{1,1}\\ \vdots \\ b_{q,1}\end{bmatrix},\dots,\begin{bmatrix}b_{1,n}\\ \vdots \\ b_{q,n}\end{bmatrix}\right)$$

Proposition. *Les opérations des types suivants appliquées à une matrice ne changent pas l'espace engendré par les lignes ni les dépendances linéaires entre les colonnes* :

(1) $L_i \leftarrow L_i + \alpha L_j$, $i \neq j$,

(2) $L_i \leftarrow \alpha L_i$, $\alpha \neq 0$,

(3) $L_i \leftrightarrow L_j$, $i \neq j$,

(4) *ajouter ou supprimer une ligne nulle.*

Par symétrie, les opérations analogiques sur les colonnes ne changent pas l'espace engendré par les colonnes ni les dépendances linéaires entre les lignes. En particulier, toutes ces opérations ne changent ni le rang de la famille des colonnes, ni le rang de la famille des lignes.

Lemme. *Par les opérations élémentaires décrites dans la dernière proposition, on peut transformer n'importe quelle matrice A en une matrice* (carrée) *B telle que dans chaque ligne et dans chaque colonne de B, il se trouve exactement un élément non zéro.*

Par exemple, on peut ainsi transformer

$$\begin{pmatrix}1 & 3 & 1 & 1\\ 2 & 2 & 1 & 0\\ 3 & 1 & 1 & -1\end{pmatrix} \rightsquigarrow \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Lemme. *Si dans chaque ligne et dans chaque colonne d'une matrice, il y a au plus un élément non zéro, alors le rang de la famille des lignes est égal au rang de la famille des colonnes et égal au nombre d'éléments non zéros de cette matrice.*

Théorème. *Le rang de la famille des lignes d'une matrice est égal au rang de sa famille des colonnes.*

Démonstration. Soit A une matrice. Par les opérations élémentaires sur les lignes et sur les colonnes, on peut transformer A en une matrice (carrée) B telle que dans chaque ligne et dans chaque colonne de B, il se trouve exactement un élément non zéro. Le rang de la famille des lignes de A est égal au rang de la famille des lignes de B, et le rang de la famille des colonnes de A est égal au rang de la famille des colonnes de B. Or, B a le même rang de lignes que le rang de colonnes. □

Définition. Le *rang* d'une matrice A, noté $\operatorname{rk} A$ (en anglais) ou $\operatorname{rang} A$, est le rang de la famille des colonnes de A et simultanément le rang de la famille des lignes de A.

Proposition. *Le rang d'une matrice échelonnée est égal à son nombre des pivots.*

Remarque. Si $A \in \mathbf{R}^{n\times n}$ est échelonnée réduite et de rang n, alors $A = I$.

Définition. Le *rang* d'un système d'équations linéaires est le rang de sa matrice (non augmentée).

Théorème. *Soit (S) un système d'équations linéaires.*

(1) *Si le rang de (S) est égal au nombre d'équations, alors (S) admet au moins une solution.*

(2) *Si le rang de (S) est égal au nombre de variables, alors (S) admet au plus une solution.*

Démonstration. Soient (S) le système

$$\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1\\ & & & & \vdots & & \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}\right. \qquad (S)$$

et A sa matrice :

$$A=\begin{pmatrix}a_{1,1} & \cdots & a_{1,n}\\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n}\end{pmatrix}.$$

Soient $a_1,\dots,a_n$ les colonnes de A vues comme éléments de $\mathbf{R}^m$:

$$a_1=(a_{1,1},\dots,a_{m,1}),\quad \dots,\quad a_n=(a_{1,n},\dots,a_{m,n}).$$

Soit $b = (b_1, \dots, b_m)$ le vecteur second membre de (S). Alors le système (S) est équivalent à l'équation

$$x_1 a_1 + \cdots + x_n a_n = b.$$

Si $\mathrm{rk}(a_1, \dots, a_n) = m$, alors la famille $(a_1, \dots, a_n)$ est génératrice de $\mathbf{R}^m$, et donc b s'écrit comme une combinaison linéaire de $a_1, \dots, a_n$.

Si $\mathrm{rk}(a_1, \dots, a_n) = n$, alors la famille $(a_1, \dots, a_n)$ est libre, et donc il n'est pas possible d'écrire b comme une combinaison linéaire de $a_1, \dots, a_n$ de deux façons différentes. □

II.6. Sous-espaces affines

Définition. Soient E un espace vectoriel et F un sous-espace vectoriel de E. Une partie non vide A de E est dite un *sous-espace affine de E avec espace vectoriel associé F* si et seulement si

(1) $\forall x \in A,\ \forall u \in F,\ x + u \in A$,

(2) $\forall x, y \in A,\ y - x \in F$.

Éléments d'un espace vectoriel s'appellent « vecteurs », mais éléments d'un espace affine s'appellent « points ». Si A est un sous-espace affine d'un espace vectoriel et $x, y \in A$, alors

$$\overrightarrow{xy} \overset{\text{déf}}{=} y - x.$$

Exemples. (1)

$$E = \mathbf{R}^2,$$
$$A = \{\, (x, y) \in E \mid x + y = 1 \,\},$$
$$F = \{\, (x, y) \in E \mid x + y = 0 \,\}.$$

(2) Soient (S) un système d'équations linéaires et (S_0) son système homogène associé. Alors l'ensemble des solutions de (S) est un espace affine, son espace vectoriel associé est l'ensemble des solutions de (S_0).

(3) Soient (E) une équation différentielle linéaire et (E_0) son équation homogène associée, par exemple :

$$y'' + \sin(x) y' + \cos(x) y = e^x, \qquad (E)$$
$$y'' + \sin(x) y' + \cos(x) y = 0. \qquad (E_0)$$

Alors l'ensemble des solutions de (E) est un espace affine, son espace vectoriel associé est l'ensemble des solutions de (E_0).

III. Applications linéaires [1h]

III.1. Définition et exemples

Définition. Soient E et F deux espaces vectoriels et $h\colon E \to F$ une application. Alors h est dite *linéaire* si et seulement si

(1) pour tous $x, y \in E$,

$$h(x + y) = h(x) + h(y),$$

(2)

$$h(\vec{0}) = \vec{0},$$

(3) pour tout $x \in E$,

$$h(-x) = -h(x),$$

(4) pour tous $x \in E$ et $\alpha \in \mathbf{R}$,

$$h(\alpha x) = \alpha h(x).$$

Proposition. *Une application $h\colon E \to F$ entre deux espaces vectoriels est linéaire si les conditions (1) et (4) de la définition sont satisfaites.*

Exemple. Opérateurs différentiels : $\frac{d}{dx}$, $\frac{d^2}{dx^2}$, etc.

III.2. Propriétés basiques et opérations

Proposition. *Soit $h\colon E \to F$ une application linéaire.*

(1) *Si G est un sous-espace vectoriel de E, alors $h(G)$ est un sous-espace vectoriel de F.*

(2) *Si G est un sous-espace vectoriel de F, alors $h^{-1}(G)$ est un sous-espace vectoriel de E.*

Proposition. *Si $f\colon E \to F$ et $g\colon F \to G$ sont deux applications linéaires, alors l'application composée $g \circ f\colon E \to G$ est aussi linéaire.*

Proposition. *Si $f\colon E \to F$ est une application linéaire bijective entre E et F, alors la bijection réciproque $f^{-1}\colon F \to E$ est aussi linéaire.*

Proposition. *Soient E et F deux espaces vectoriels. Alors l'ensemble des applications linéaires $E \to F$ est un espace vectoriel par rapport aux opérations définies par les formules*

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

L'espace vectoriel des applications linéaires $E \to F$ est noté $\mathcal{L}(E, F)$.

III.3. Image, noyau, rang

Définition. Soit $h\colon E \to F$ une application linéaire. L'*image* et le *noyau* de h, notés $\operatorname{img} h$ et $\ker h$ (« *kernel* » en anglais), sont définis comme suit :

$$\operatorname{img} h \overset{\text{déf}}{=} \{\, h(x) \mid x \in E \,\},$$
$$\ker h \overset{\text{déf}}{=} \{\, x \in E \mid h(x) = \vec{0} \,\}.$$

Proposition. *Soit $h\colon E \to F$ une application linéaire. Alors $\ker h$ est un sous-espace vectoriel de E et $\operatorname{img} h$ est un sous-espace vectoriel de F.*

Démonstration. $\ker h = h^{-1}(\{\vec{0}\})$, $\operatorname{img} h = h(E)$. □

Proposition. *Soient $h\colon E \to F$ une application linéaire et $x, y \in E$. Alors*

$$h(x) = h(y) \quad \Leftrightarrow \quad x - y \in \ker h.$$

Corollaire. *Une application linéaire h est injective si et seulement si*

$$\ker h = \{\vec{0}\}.$$

Définition. Soit $h\colon E \to F$ une application linéaire. Le *rang* de h, noté $\operatorname{rk} h$, est la dimension de l'image :

$$\operatorname{rk} h \overset{\text{déf}}{=} \dim \operatorname{img} h.$$

Lemme. *Si $h\colon E \to F$ est une application linéaire et $x_1, \dots, x_n \in E$, alors*

$$h(\operatorname{Vect}(x_1, \dots, x_n)) = \operatorname{Vect}(h(x_1), \dots, h(x_n)).$$

Il en résulte facilement que si $h\colon E \to F$ est une application linéaire, alors $\operatorname{rk} h \leqslant \dim E$. D'ailleurs, il est clair que $\operatorname{rk} h \leqslant \dim F$.

Théorème (Théorème du rang)**.** *Soit $h\colon E \to F$ une application linéaire. Alors*

$$\operatorname{rk} h + \dim \ker h = \dim E.$$

On peut aussi écrire cette formule comme suit :

$$\dim \operatorname{img} h + \dim \ker h = \dim \operatorname{dom} h,$$

où $\operatorname{dom} h \overset{\text{déf}}{=} E$, le « domaine » de h.

Démonstration. Soit $(v_1, \dots, v_s)$ une base de $\ker h$. Complétons la en une base $(v_1, \dots, v_s, u_1, \dots, u_r)$ de E. On va vérifier que $(h(u_1), \dots, h(u_r))$ est une base de $\operatorname{img} h$.

La famille $(h(u_1), \dots, h(u_r))$ est génératrice de $\operatorname{img} h$ car

$$\begin{aligned}\operatorname{img} h = h(E) &= h(\operatorname{Vect}(v_1, \dots, v_s, u_1, \dots, u_r)) \\ &= \operatorname{Vect}(\vec{0}, \dots, \vec{0}, h(u_1), \dots, h(u_r)) = \operatorname{Vect}(h(u_1), \dots, h(u_r)).\end{aligned}$$

Montrons que $(h(u_1), \dots, h(u_r))$ est libre. Supposons pour cela que

$$\alpha_1 h(u_1) + \cdots + \alpha_r h(u_r) = \vec{0}, \quad \alpha_1, \dots, \alpha_r \in \mathbf{R}.$$

Alors

$$h(\alpha_1 u_1 + \cdots + \alpha_r u_r) = \vec{0},$$

donc il existe $\beta_1, \dots, \beta_s \in \mathbf{R}$ tels que

$$\alpha_1 u_1 + \cdots + \alpha_r u_r = \beta_1 v_1 + \cdots + \beta_s v_s.$$

Comme la famille $(v_1, \dots, v_s, u_1, \dots, u_r)$ est libre, forcement

$$\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_s = 0.$$

Ainsi $(h(u_1), \dots, h(u_r))$ est une base de $\operatorname{img} h$. Donc

$$\dim E = r + s = \dim \operatorname{img} h + \dim \ker h.$$ □

Exemple. Soit

$$E = \{\, (x, y, z) \in \mathbf{R}^3 \mid x + y + z = 0 \,\}$$

un sous-espace vectoriel de $\mathbf{R}^3$. Trouver sa dimension.

Une façon de faire est la suivante. Soit $h\colon \mathbf{R}^3 \to \mathbf{R}$ l'application linéaire définie par la formule :

$$h(x, y, z) = x + y + z.$$

Alors $E = \ker h$. Observons que $h(1, 0, 0) \neq 0$, donc $\operatorname{img} h = \mathbf{R}$. Ainsi, par le théorème du rang,

$$\dim E = \dim \mathbf{R}^3 - \dim \operatorname{img} h = 3 - 1 = 2.$$

Avant d'énoncer le théorème suivant, voici un rappel pertinent : si A et B sont deux ensembles finis avec le même nombre d'éléments et $f\colon A \to B$ une fonction (application), alors f est injective si et seulement si f est surjective.

Théorème. *Soient E et F deux espaces vectoriels de la même dimension finie* ($\dim E = \dim F < \infty$). *Alors pour tout application linéaire $h\colon E \to F$, les énoncés suivants sont équivalents* :

(1) *h est injective sur E* ($\ker h = \{\vec{0}\}$),

(2) *h est surjective sur F* ($\operatorname{img} h = F$).

Démonstration. h est injective sur $E \Leftrightarrow \ker h = \{\vec{0}\} \Leftrightarrow \dim \ker h = 0 \Leftrightarrow \dim \operatorname{img} h = \dim E$ (par le théorème du rang) $\Leftrightarrow \dim \operatorname{img} h = \dim F$ (car $\dim E = \dim F$) $\Leftrightarrow \operatorname{img} h = F$ (car si $\operatorname{img} h$ était un sous-espace *propre* de F, $\dim \operatorname{img} h$ serait inférieur au égale à $\dim F - 1$) $\Leftrightarrow$ h est surjective sur F. □

III.4. Applications linéaires comme « morphismes »

Une application linéaire s'appelle aussi un *homomorphisme* d'espaces vectoriels. Une application linéaire injective s'appelle un *monomorphisme*. Une application linéaire surjective s'appelle un *epimorphisme*. Une application linéaire bijective s'appelle un *isomorphisme*.

En fait, une meilleur façon de définir un *isomorphisme* est comme un homomorphisme *inversible*.

Une application linéaire $h\colon E \to F$ est dite *inversible* (comme une application linéaire) si et seulement si il existe une application linéaire $g\colon F \to E$ telle

que $g \circ h = \mathrm{id}_E$ et $h \circ g = \mathrm{id}_F$. Dans ce cas, telle application g est unique et s'appelle l'application *réciproque* de h, notée h^{-1}. Comme on le sait déjà, une application linéaire $h\colon E \to F$ est inversible comme une application linéaire entre E et F si et seulement si elle est bijective entre E et F.

Une application linéaire d'un espace linéaire dans lui-même s'appelle un *endomorphisme*. Les endomorphismes linéaires s'appellent aussi *opérateurs* (linéaires). Un endomorphisme qui est au même temps un isomorphisme (d'un espace avec lui-même) s'appelle un *automorphisme*.

III.5. Matrices d'applications linéaires

Soient $h\colon E \to F$ une application linéaire et $(e_1, \ldots, e_n)$ une base de E. Alors h est complètement déterminée par les valeurs $h(e_1), \ldots, h(e_n)$.

Définition. Soient $h\colon E \to F$ une application linéaire, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, et $\mathcal{C} = (f_1, \ldots, f_m)$ une base de F. Alors la *matrice* de h *relativement* aux bases $\mathcal{B}$ et $\mathcal{C}$ (ou *par rapport* aux bases $\mathcal{B}$ et $\mathcal{C}$) est la matrice avec la famille des colonnes $([h(e_1)]_{\mathcal{C}}, \ldots, [h(e_n)]_{\mathcal{C}})$. C'est-à-dire, si

$$\begin{cases} h(e_1) &= a_{1,1}f_1 + \cdots + a_{m,1}f_m, \\ &\vdots \\ h(e_n) &= a_{1,n}f_1 + \cdots + a_{m,n}f_m, \end{cases}$$

alors la matrice de h relativement aux bases $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_m)$ est

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

Elle sera notée $[h]_{\mathcal{C},\mathcal{B}}$.

En particulier, si $F = E$, alors $[\mathrm{id}]_{\mathcal{C},\mathcal{B}}$ est la matrice de passage (matrice de changement de base) de $\mathcal{C}$ à $\mathcal{B}$.

Proposition. *Soient $h\colon E \to F$ une application linéaire entre deux espaces vectoriels E et F des dimensions finies, $\mathcal{B}$ une base de E et $\mathcal{C}$ une base de F. Alors*

$$\operatorname{rk} h = \operatorname{rk}[h]_{\mathcal{C},\mathcal{B}}.$$

Démonstration. Pour tous $\alpha_1, \dots, \alpha_n \in \mathbf{R}$,

$$\alpha_1 h(e_1) + \cdots + \alpha_n h(e_n) = \vec{0} \quad \Leftrightarrow \quad \alpha_1 [h(e_1)]_{\mathcal{C}} + \cdots + \alpha_n [h(e_n)]_{\mathcal{C}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

donc,

$$\operatorname{rk}(h(e_1), \dots, h(e_n)) = \operatorname{rk}([h(e_1)]_{\mathcal{C}}, \dots, [h(e_n)]_{\mathcal{C}}).$$

Or,

$$\operatorname{rk}(h(e_1), \dots, h(e_n)) = \dim \operatorname{Vect}(h(e_1), \dots, h(e_n)) = \dim \operatorname{img} h = \operatorname{rk} h$$

et

$$\operatorname{rk}([h(e_1)]_{\mathcal{C}}, \dots, [h(e_n)]_{\mathcal{C}}) = \operatorname{rk}[h]_{\mathcal{C},\mathcal{B}}.$$ □

IV. Endomorphismes d'espaces vectoriels[1h]

IV.1. Définition et exemples

Définition. Un *endomorphisme* d'un espace vectoriel E est une application linéaire $E \to E$.

[...]

IV.2. Propriétés basiques et opérations

[...]

IV.3. Sous-espaces stables

Soient E un espace vectoriel et $h\colon E \to E$ un endomorphisme de E.

Définition. Un sous-espace vectoriel F de E est dit *stable* par h si et seulement si

$$h(x) \in F \quad \text{pour tout} \quad x \in F.$$

En utilisant une autre notation, on peut dire que F est stable par h si et seulement si $h(F) \subset F$.

Exercice. (1) Montrer que $\ker h$ est stable par h.

(2) Montrer que $\ker h^2$ est stable par h.

(3) Montrer que pour tout $m \in \mathbf{N}$, $\ker h^m$ est stable par h.

Exercice. (1) Montrer que $\operatorname{img} h$ est stable par h.

(2) Montrer que $\operatorname{img} h^2$ est stable par h.

(3) Montrer que pour tout $m \in \mathbf{N}$, $\operatorname{img} h^m$ est stable par h.

Exercice. Soient $h, k\colon E \to E$ deux endomorphismes de E tels que $h \circ k = k \circ h$.

(1) Montrer que $\ker k$ est stable par h.

(2) Montrer que $\operatorname{img} k$ est stable par h.

IV.4. Spectre, vecteurs et valeurs propres

Soient E un espace vectoriel et $h\colon E \to E$ un endomorphisme de E.

Définition. Le *spectre* de h, noté $\sigma(h)$ ou $\operatorname{Sp} h$, est l'ensemble de tous les $\alpha \in \mathbf{R}$ tels que l'application $h - \alpha \operatorname{id}_E$ n'est pas inversible :

$$\sigma(h) \overset{\text{déf}}{=} \{\, \alpha \in \mathbf{R} \mid h - \alpha \operatorname{id}_E \text{ n'est pas inversible} \,\}.$$

Observation : l'application h est inversible si et seulement si $0 \notin \sigma(h)$.

Exemple (important)**.**

$$\sigma(\operatorname{id}) = \{1\}, \qquad \sigma(\lambda \operatorname{id}) = \{\lambda\} \quad \text{pour tout} \quad \lambda \in \mathbf{R}.$$

Définition. Un *vecteur propre* de h est un vecteur non nul v tel qu'il existe $\alpha \in \mathbf{R}$ tel que

$$h(v) = \alpha v.$$

Définition. Une *valeur propre* de h est une valeur $\alpha \in \mathbf{R}$ telle qu'il existe un vecteur non nul v tel que

$$h(v) = \alpha v.$$

Autrement dit, α est une valeur propre de h si et seulement si l'application $h - \alpha \operatorname{id}_E$ n'est pas injective. En effet :

$$(h - \alpha \operatorname{id}_E)(v) = h(v) - \alpha \operatorname{id}_E(v) = h(v) - \alpha v,$$

donc $h - \alpha \operatorname{id}_E$ n'est pas injective si et seulement si il existe un $v \neq \vec{0}$ tel que $h(v) - \alpha v = \vec{0}$.

Proposition. *Si α est une valeur propre de h, alors $\alpha \in \sigma(h)$.*

Théorème. *Si $\dim E$ est finie, alors $\sigma(h)$ est l'ensemble des valeurs propres de h.*

Esquisse d'une démonstration. Appliquer le fait que tout endomorphisme injectif d'un espace de dimension finie est inversible. □

Exemples. (1) Soit $E = C^\infty(\mathbf{R})$ l'espace des fonctions $\mathbf{R} \to \mathbf{R}$ indéfiniment dérivables, « de variable t ». Soit $h = \frac{d}{dt}\colon E \to E$. Alors l'ensemble des valeurs propres de h est $\mathbf{R}$, et l'ensemble des vecteurs propres associés à une valeur propre λ est l'ensemble des fonctions de la forme $t \mapsto Ce^{\lambda t}$ avec $C \in \mathbf{R}^* = \mathbf{R} \setminus \{0\}$.

(2) Soient $E = \mathbf{R}^2$, $h\colon E \to E$, $h(x, y) = (x + 2y, 2x + y)$. Alors les valeurs propres de h sont 3 et -1. Un vecteur propre associé à la valeur 3 : $v_1 = (1, 1)$. Un vecteur propre associé à la valeur -1 : $v_2 = (1, -1)$. Une question supplémentaire : soit $\mathcal{B} = (v_1, v_2)$, trouver la représentation matricielle $[h]_{\mathcal{B},\mathcal{B}}$ de h relativement à la base $\mathcal{B}$.

Définition. Soit α une valeur propre d'un endomorphisme h de E. Alors le sous-espace $\ker(h - \alpha \operatorname{id}_E)$ de E s'appelle le *sous-espace propre* associé à α, et sa dimension s'appelle la *multiplicité géométrique* de la valeur propre α.

Proposition. *Soit $h\colon E \to E$ un endomorphisme d'un espace vectoriel réel E. Si $\alpha_1, \dots, \alpha_n \in \mathbf{R}$ sont des valeurs deux à deux différentes et que $v_1, \dots, v_n$ sont des éléments de E tels que*

$$h(v_i) = \alpha_i v_i, \quad \textit{pour tout } i = 1, \dots, n,$$

et

$$v_1 + \cdots + v_n = \vec{0},$$

alors

$$v_1 = \cdots = v_n = \vec{0}.$$

En d'autres termes, cette proposition affirme que si $\alpha_1, \dots, \alpha_n \in \mathbf{R}$ sont deux à deux différentes, alors

$$\ker(h - \alpha_1 \operatorname{id}) + \cdots + \ker(h - \alpha_n \operatorname{id}) = \ker(h - \alpha_1 \operatorname{id}) \oplus \cdots \oplus \ker(h - \alpha_n \operatorname{id}).$$

Démonstration. Montrons, par exemple, que, sous les hypothèses énoncées, $v_1 = \vec{0}$. Posons

$$P = \frac{(X - \alpha_2)(X - \alpha_3) \cdots (X - \alpha_n)}{(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \cdots (\alpha_1 - \alpha_n)} \in \mathbf{R}[X].$$

Alors

$$P(\alpha_1) = 1 \quad \text{et} \quad P(\alpha_2) = P(\alpha_3) = \cdots = P(\alpha_n) = 0.$$

On va maintenant substituer h pour X dans P et appliquer le résultat à la somme $v_1 + \cdots + v_n$.

On note $P(h)$ l'endomorphisme de E qui s'exprime comme :

$$P(h) = \frac{(h - \alpha_2 \operatorname{id}) \circ \cdots \circ (h - \alpha_n \operatorname{id})}{(\alpha_1 - \alpha_2) \cdots (\alpha_1 - \alpha_n)}$$

(toute autre écriture du polynôme P résulte en même endomorphisme $P(h)$ après la substitution de h pour X).

Observons que si $h(v) = \alpha v$, alors $P(h)(v) = P(\alpha)v$. Ainsi,

$$\begin{aligned} P(h)(v_1 + \cdots + v_n) &= P(h)(v_1) + \cdots + P(h)(v_n) \\ &= P(\alpha_1)v_1 + \cdots + P(\alpha_n)v_n = v_1. \end{aligned}$$

Or, $P(h)(v_1 + \cdots + v_n) = P(h)(\vec{0}) = \vec{0}$. Donc $v_1 = \vec{0}$.

De la même manière on peut montrer que $v_2 = \vec{0}, \ldots, v_n = \vec{0}$. □

V. Matrices[3h]

V.1. Définitions et exemples

Définition. Une *matrice* réelle de taille $m \times n$ est une famille de nombres réels $(a_{i,j})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n}$, d'habitude écrite en tableau :

$$\begin{matrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{matrix}$$

L'élément d'indice (i, j) d'une matrice A sera noté $[A]_{i,j}$.
L'ensemble des matrices réelles de taille $m \times n$ sera noté $\mathcal{M}_{m,n}(\mathbf{R})$ ou $\mathbf{R}^{m \times n}$.
Voici quelques matrices spéciales.

- Les *matrices nulles* $O = O_{m,n} \in \mathbf{R}^{m \times n}$: $[O]_{i,j} = 0$.
- Les *unités matricielles* $E_{k,l} = E_{k,l;m,n} \in \mathbf{R}^{m \times n}$:

$$[E_{k,l}]_{i,j} = \begin{cases} 1 & \text{si } i = k \text{ et } j = l, \\ 0 & \text{sinon.} \end{cases}$$

- Les *matrices identités* $I = I_n \in \mathbf{R}^{n \times n}$:

$$[I]_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

V.2. Opérations sur matrices

V.2.1. Matrices comme vecteurs

L'*addition* de matrices $A, B \in \mathbf{R}^{m \times n}$ est définie par la formule

$$[A + B]_{i,j} \overset{\text{déf}}{=} [A]_{i,j} + [B]_{i,j}.$$

La *multiplication par un scalaire* $\alpha \in \mathbf{R}$ d'une matrice A est définie par la formule

$$[\alpha A]_{i,j} \overset{\text{déf}}{=} \alpha [A]_{i,j}.$$

Proposition. *L'ensemble* $\mathbf{R}^{m\times n}$ *avec les opérations d'addition et de multiplication par les réels est un espace vectoriel réel de dimension* mn*. La famille* $(E_{i,j})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ *est une base de* $\mathbf{R}^{m\times n}$.

V.2.2. La transposition

Soit $A \in \mathbf{R}^{m\times n}$ une matrice. Alors la *matrice transposée* de A, notée ${}^{\mathbf{t}}A$, est la matrice dans $\mathbf{R}^{n\times m}$ définie par la formule :

$$[{}^{\mathbf{t}}A]_{i,j} = [A]_{j,i}.$$

Observations :

(1) pour toute matrice A, ${}^{\mathbf{t}}({}^{\mathbf{t}}A) = A$,

(2) l'application

$$A \mapsto {}^{\mathbf{t}}A, \quad \mathbf{R}^{m\times n} \to \mathbf{R}^{n\times m},$$

est un isomorphisme des *espaces vectoriels* $\mathbf{R}^{m\times n}$ et $\mathbf{R}^{n\times m}$.

V.2.3. La multiplication

Supposons que certaines valeurs ou variables $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p$ s'écrivent comme combinaisons linéaires de $\boldsymbol{y}_1, \dots, \boldsymbol{y}_q$, et que $\boldsymbol{y}_1, \dots, \boldsymbol{y}_q$ s'écrivent comme combinaisons linéaires de $\boldsymbol{z}_1, \dots, \boldsymbol{z}_r$:

$$\begin{cases} \boldsymbol{x}_1 &= a_{1,1}\boldsymbol{y}_1 + \cdots + a_{1,q}\boldsymbol{y}_q \\ &\vdots \\ \boldsymbol{x}_p &= a_{p,1}\boldsymbol{y}_1 + \cdots + a_{p,q}\boldsymbol{y}_q \end{cases} \qquad \begin{cases} \boldsymbol{y}_1 &= b_{1,1}\boldsymbol{z}_1 + \cdots + b_{1,r}\boldsymbol{z}_r \\ &\vdots \\ \boldsymbol{y}_q &= b_{q,1}\boldsymbol{z}_1 + \cdots + b_{q,r}\boldsymbol{z}_r \end{cases}$$

Soient A et B les matrices de ces deux « systèmes » :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,q} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,q} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,r} \\ \vdots & \ddots & \vdots \\ b_{q,1} & \cdots & b_{q,r} \end{pmatrix}.$$

Si on exprime $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p$ comme combinaisons linéaires de $\boldsymbol{z}_1, \dots, \boldsymbol{z}_r$, on obtient un « système » de la forme

$$\begin{cases} \boldsymbol{x}_1 &= c_{1,1}\boldsymbol{z}_1 + \cdots + c_{1,r}\boldsymbol{z}_r \\ &\vdots \\ \boldsymbol{x}_p &= c_{p,1}\boldsymbol{z}_1 + \cdots + c_{p,r}\boldsymbol{z}_r \end{cases}$$

La matrice C de ce « système » s'appelle le *produit* de A et B, noté AB.

Autrement dit, le produit de $A \in \mathbf{R}^{p\times q}$ avec $B \in \mathbf{R}^{q\times r}$ est la matrice $AB \in \mathbf{R}^{p\times r}$ définie par la formule :

$$[AB]_{i,j} = [A]_{i,1}[B]_{1,j} + [A]_{i,2}[B]_{2,j} + \cdots + [A]_{i,q}[B]_{q,j} = \sum_{k=1}^{q} [A]_{i,k}[B]_{k,j}.$$

Voici quelques observations.

(1) Pour $E_{i,j} \in \mathbf{R}^{p\times q}$ et $E_{k,l} \in \mathbf{R}^{q\times r}$,

$$E_{i,j}E_{k,l} = \begin{cases} E_{i,l} \in \mathbf{R}^{p\times r} & \text{si } j = k, \\ O \in \mathbf{R}^{p\times r} & \text{sinon.} \end{cases}$$

(2) Pour la matrice identité $I \in \mathbf{R}^{m\times m}$ et pour tout $A \in \mathbf{R}^{m\times n}$, $IA = A$. Pour la matrice identité $I \in \mathbf{R}^{n\times n}$ et pour tout $A \in \mathbf{R}^{m\times n}$, $AI = A$.

Proposition. *La multiplication des matrices vérifie les propriétés suivantes (lorsque les dimensions des matrices sont « compatibles ») :*

(1) $(AB)C = A(BC)$,

(2) $IA = A$, $AI = A$,

(3) $A(B + C) = AB + AC$, $(A + B)C = AC + BC$,

(4) ${}^{\mathbf{t}}(AB) = {}^{\mathbf{t}}B \cdot {}^{\mathbf{t}}A$.

Proposition. *Si* A, C*, et* L *sont trois matrices et* $A = CL$*, alors toute colonne de* A *est une combinaison linéaire des colonnes de* C*, et toute ligne de* A *est une combinaison linéaire des lignes de* L.

Application aux systèmes d'équations linéaires

Écriture d'un système d'équations linéaires à l'aide de matrices :

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m \end{cases} \quad \Leftrightarrow \quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Équations « matricielles » linéaires

Théorème. *Soient $A \in \mathbf{R}^{m\times n}$ et $B \in \mathbf{R}^{m\times k}$. Soit l'équation*

$$AX = B \qquad (E)$$

pour l'inconnue $X \in \mathbf{R}^{n\times k}$.

(1) *Si $\operatorname{rk} A = m$, alors l'équation (E) admet au moins une solution.*

(2) *Si $\operatorname{rk} A = n$, alors l'équation (E) admet au plus une solution.*

Démonstration. Soient $B_1, \dots, B_k$ les colonnes de la matrice B et $X_1, \dots, X_k$ les colonnes de la matrice inconnue X. Alors l'équation (E) pour X est équivalente au système suivante pour $X_1, \dots, X_k$:

$$\begin{cases} AX_1 &= B_1 \\ &\vdots \\ AX_k &= B_k \end{cases}$$

Chaque équation

$$AX_i = B_i$$

est équivalente à un système de m équations pour n inconnues réelles (pour les composantes de X_i). Comme démontré précédemment, si $\operatorname{rk} A = m$, alors ce système admet au moins une solution, et si $\operatorname{rk} A = n$, alors ce système admet au plus une solution.

Si chaque équation $AX_i = B_i$ admet au moins une solution, alors l'équation (E), elle aussi, admet au moins une solution. Si chaque équation $AX_i = B_i$ admet au plus une solution, alors l'équation (E) admet au plus une solution. □

On peut résoudre toute équation matricielle de la forme

$$AX = B,$$

où A et B sont données et X est l'inconnue, par une forme d'élimination de Gauss.

V.3. Matrices des applications linéaires

Proposition. *Soient E, F, G espaces vectoriels de dimensions finies, $\mathcal{A}$ une base de E, $\mathcal{B}$ une base de F, $\mathcal{C}$ une base de G.*

(1)
$$[\mathrm{id}_E]_{\mathcal{A},\mathcal{A}} = I.$$

(2) *Si $f\colon E \to F$ est une application linéaire et $v \in E$ un vecteur, alors*
$$[f(v)]_{\mathcal{B}} = [f]_{\mathcal{B},\mathcal{A}}[v]_{\mathcal{A}}.$$

(3) *Si $f\colon E \to F$ et $g\colon F \to G$ sont deux applications linéaires, alors*
$$[g \circ f]_{\mathcal{C},\mathcal{A}} = [g]_{\mathcal{C},\mathcal{B}}[f]_{\mathcal{B},\mathcal{A}}.$$

V.4. Rang de matrices

V.4.1. Propriétés du rang par rapport aux opérations

Voici quelques propriétés du rang de matrice par rapport aux opérations.

(1) Pour toute matrice A,
$$\operatorname{rk} {}^{\mathtt{t}}A = \operatorname{rk} A.$$

(2) Si A et B sont deux matrices de mêmes dimensions, alors
$$\operatorname{rk}(A+B) \leqslant \operatorname{rk} A + \operatorname{rk} B.$$

(3) Si A et B sont deux matrices et le produit AB est défini, alors
$$\operatorname{rk} AB \leqslant \min\{\operatorname{rk} A, \operatorname{rk} B\}.$$

L'égalité $\operatorname{rk} {}^{\mathtt{t}}A = \operatorname{rk} A$ est assez facile à démontrer.

Pour démontrer l'inégalité $\operatorname{rk}(A+B) \leqslant \operatorname{rk} A + \operatorname{rk} B$, on peut observer, par exemple, que toute colonne de $A+B$ est la somme d'une colonne de A et d'une colonne de B, d'où, chaque colonne de $A+B$ appartient à l'espace engendré par les colonnes de A et de B.

Pour démontrer que $\operatorname{rk} AB \leqslant \operatorname{rk} A$ et que $\operatorname{rk} AB \leqslant \operatorname{rk} B$, on peut appliquer la propriété que toute colonne de AB est une combinaison linéaire des colonnes de A, et toute ligne de AB est une combinaison linéaire des lignes de B.

V.4.2. Définition du rang en revue

Voici une autre démonstration que le rang de la famille des lignes d'une matrice est égal au rang de sa famille des colonnes.

Théorème. *Le rang de la famille des lignes d'une matrice est égal au rang de sa famille des colonnes.*

Démonstration. Soient A une matrice $m \times n$ et k un entier. Alors les énoncés suivants sont équivalents :

(1) le rang de la famille des colonnes de A est $\leqslant k$,

(2) il existe k colonnes $c_1, \dots, c_k$ (pas nécessairement dans A) de taille m telles que toute colonne de A est une combinaison linéaire de $c_1, \dots, c_k$,

(3) il existe une matrice C de taille $m \times k$ et une matrice L de taille $k \times n$ telles que $A = CL$,

(4) il existe k lignes $l_1, \dots, l_k$ (pas nécessairement dans A) de taille n telles que toute ligne de A est une combinaison linéaire de $l_1, \dots, l_k$,

(5) le rang de la famille des lignes de A est $\leqslant k$.

En effet, les équivalences suivantes sont évidentes :

$$(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4) \Leftrightarrow (5).$$

Il résulte de l'équivalence (1) $\Leftrightarrow$ (5) que la famille des colonnes et la famille des lignes d'une matrice ont le même rang. □

V.4.3. Les « factorisations du rang »

Théorème. *Si $A \in \mathbf{R}^{m\times n}$ est une matrice du rang $r > 0$, alors il existe $B \in \mathbf{R}^{m\times r}$ et $C \in \mathbf{R}^{r\times n}$ telles que $A = BC$.*

Corollaire. *Le rang d'une matrice non nulle $A \in \mathbf{R}^{m\times n}$ est le plus petit $r \in \mathbf{N}$ tel que A se décompose en produit $A = BC$ avec $B \in \mathbf{R}^{m\times r}$ et $C \in \mathbf{R}^{r\times n}$.*

Exemple. Soit

$$A = \begin{pmatrix} 0 & 1 & 3 & 4 \\ 0 & 2 & 2 & 4 \\ 0 & 3 & 1 & 4 \end{pmatrix}.$$

Alors $\operatorname{rk} A = 2$, et les colonnes $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ et $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ forment une famille libre maximale parmi les colonnes de A. En exprimant les autres colonnes de A comme combinaisons linéaires de ces deux, on trouve que

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

V.5. Le noyau et l'image d'une matrice

Par analogie avec les applications linéaires, on définit le *noyau* et l'*image* d'une matrice.

Définition. Soit $A \in \mathbf{R}^{m\times n}$. Le *noyau* de A est l'ensemble des éléments $x \in \mathbf{R}^n$ qui, écrits en colonne, satisfont l'équation

$$Ax = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

L'image de A est l'ensemble des éléments $y \in \mathbf{R}^m$ tels que si y est écrit en colonne, alors il existe $x \in \mathbf{R}^n$ écrit en colonne tel que

$$Ax = y.$$

On va noter $\ker A$ le noyau de A et $\operatorname{img} A$ l'image de A.

Si $A \in \mathbf{R}^{m\times n}$, alors $\ker A$ est un sous-espace vectoriel de $\mathbf{R}^n$ et $\operatorname{img} A$ est un sous-espace vectoriel de $\mathbf{R}^m$.

Comme pour les applications linéaires,

$$\dim \operatorname{img} A = \operatorname{rk} A$$

(cependant, pour les applications linéaires, c'était la définition du rang, alors que pour les matrices, c'est une propriété qui doit être démontrée).

Le *théorème du rang* pour les matrices est analogique à celui pour les applications linéaires : si $A \in \mathbf{R}^{m\times n}$, alors

$$\dim \ker A + \dim \operatorname{img} A = n.$$

VI. Matrices carrées[5h]

Définition. Une *matrice carrée* est une matrice de taille $n \times n$ (autrement dit, avec le même nombre des colonnes que des lignes).

VI.1. Matrices carrées de formes spéciales

Définition. Soit A une matrice carrée.

- A est dite *diagonale* si
$$[A]_{i,j} = 0 \quad \text{pour} \quad i \neq j.$$
- A est dite *triangulaire supérieure* si
$$[A]_{i,j} = 0 \quad \text{pour} \quad i > j.$$
- A est dite *triangulaire inférieure* si
$$[A]_{i,j} = 0 \quad \text{pour} \quad i < j.$$

Exemples. Soient
$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 4 & 5 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 5 & 0 & 2 & 0 \\ 0 & 6 & 0 & 3 \end{pmatrix}.$$

La matrice A est diagonale, les matrices A et B sont triangulaires supérieures, les matrices A et C sont triangulaires inférieures.

Définition. Une matrice carrée A est dite *symétrique* si ${}^{\mathbf{t}}A = A$, c'est-à-dire,
$$[A]_{i,j} = [A]_{j,i}.$$

Exemple. La matrice $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ est symétrique.

VI.2. Matrice inverse

VI.2.1. Définition et propriétés

Définition. Soient $A, B \in \mathbf{R}^{n\times n}$ deux matrices carrées. Elle sont dites *inverses* l'une de l'autre si et seulement si
$$AB = I \quad \text{et} \quad BA = I.$$

Proposition. *Si $A \in \mathbf{R}^{n\times n}$ et B et C sont deux matrices inverses de A, alors $B = C$.*

Démonstration. $B = BI = BAC = IC = C$. □

La matrice inverse d'une matrice carrée A est notée A^{-1}.
Voici quelques propriétés.

(1) Si $A \in \mathbf{R}^{n\times n}$ est inversible, alors A^{-1} est inversible et
$$(A^{-1})^{-1} = A,$$

(2) Si $A \in \mathbf{R}^{n\times n}$ est inversible, alors ${}^{\mathbf{t}}A$ est inversible et
$$({}^{\mathbf{t}}A)^{-1} = {}^{\mathbf{t}}(A^{-1}),$$

(3) Si $A, B \in \mathbf{R}^{n\times n}$ sont inversibles, alors AB est inversible et
$$(AB)^{-1} = B^{-1}A^{-1}.$$

Montrons que $({}^{\mathbf{t}}A)^{-1} = {}^{\mathbf{t}}(A^{-1})$:
$$\begin{aligned} {}^{\mathbf{t}}A\,{}^{\mathbf{t}}(A^{-1}) &= {}^{\mathbf{t}}(A^{-1}A) = {}^{\mathbf{t}}I = I, \\ {}^{\mathbf{t}}(A^{-1})\,{}^{\mathbf{t}}A &= {}^{\mathbf{t}}(AA^{-1}) = {}^{\mathbf{t}}I = I. \end{aligned}$$

Montrons que $(AB)^{-1} = B^{-1}A^{-1}$:
$$\begin{aligned} ABB^{-1}A^{-1} &= AIA^{-1} = AA^{-1} = I, \\ B^{-1}A^{-1}AB &= B^{-1}IB = B^{-1}B = I. \end{aligned}$$

Proposition. *Si $A, B \in \mathbf{R}^{n\times n}$ et $AB = I$, alors* $\operatorname{rk} A = \operatorname{rk} B = n$ *et* $BA = I$.

Démonstration. Par une proposition précédente,

$$\min(\operatorname{rk} A, \operatorname{rk} B) \geqslant \operatorname{rk} I = n.$$

Donc $\operatorname{rk} A = n$. Ainsi, par un théorème démontré dans le chapitre précédent, l'équation

$$AX = A \tag{E}$$

admet au plus une solution.

Il est clair que $X = I$ est une solution de (E). Or, $X = BA$ est aussi une solution de (E) :

$$ABA = IA = A.$$

Ainsi $BA = I$. □

Proposition. *Une matrice $n \times n$ est inversible si et seulement si son rang est n.*

Démonstration. Par la proposition précédente, si $A \in \mathbf{R}^{n\times n}$ est inversible, alors $\operatorname{rk} A = n$.

Si $A \in \mathbf{R}^{n\times n}$ et $\operatorname{rk} A = n$, alors l'équation matricielle

$$AX = I$$

admet (exactement) une solution B d'après un théorème du chapitre précédent, et par la proposition précédente, cette matrice B est l'inverse de A. □

VI.2.2. Calcul

Pour les matrices 1×1 :

$$(a)^{-1} = \left(\frac{1}{a}\right) = \frac{1}{a}\,(1)\,.$$

Pour les matrices 2×2 :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(Ici, $ad - bc$ est ce qui s'appelle le *déterminant* de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Le déterminant d'une matrice $(a) \in \mathbf{R}^{1\times 1}$ est a.)

Il existe des formes explicites pour toutes dimensions, mais elles ne fournissent pas d'habitude de façon efficace de calcul de la matrice inverse d'une matrice générale de grande taille. En pratique, on peut calculer A^{-1} en résolvant l'équation

$$AX = I.$$

VI.3. Matrices des applications linéaires et changement de bases

Proposition. *Soit $f\colon E \to F$ une application linéaire entre deux espaces vectoriels tels que $\dim E = \dim F < \infty$. Soient $\mathcal{A}$ une base de E et $\mathcal{B}$ une base de F. Alors f est inversible si et seulement si la matrice $[f]_{\mathcal{B},\mathcal{A}}$ est inversible, et dans ce cas,*

$$[f^{-1}]_{\mathcal{A},\mathcal{B}} = [f]_{\mathcal{B},\mathcal{A}}^{-1}.$$

Rappel : soient $\mathcal{A}$ et $\mathcal{B}$ deux bases d'un espace vectoriel E de dimension finie, alors la matrice de changement de base de $\mathcal{A}$ à $\mathcal{B}$ est la même que la matrice de l'application identité id_E par rapport à $\mathcal{B}$ et $\mathcal{A}$: $[\mathrm{id}_E]_{\mathcal{A},\mathcal{B}}$.

Si $h\colon E \to E$ est un endomorphisme d'un espace vectoriel E et $\mathcal{B}$ est une base de E, alors la matrice de h relativement à $\mathcal{B}$ et $\mathcal{B}$ sera aussi notée simplement $[h]_{\mathcal{B}}$, au lieu de $[h]_{\mathcal{B},\mathcal{B}}$.

Proposition. *Soient $\mathcal{A}$ et $\mathcal{B}$ deux bases d'un espace vectoriel E de dimension finie, et soit P la matrice de passage de $\mathcal{A}$ à $\mathcal{B}$. Alors*

(1) *la matrice de passage de $\mathcal{B}$ à $\mathcal{A}$ est P^{-1},*

(2) *pour tout $v \in E$,*

$$[v]_{\mathcal{A}} = P[v]_{\mathcal{B}} \quad \text{et} \quad [v]_{\mathcal{B}} = P^{-1}[v]_{\mathcal{A}},$$

(3) *pour tout endomorphisme $h\colon E \to E$,*

$$[h]_{\mathcal{A}} = P[h]_{\mathcal{B}}P^{-1} \quad \text{et} \quad [h]_{\mathcal{B}} = P^{-1}[h]_{\mathcal{A}}P.$$

Démonstration.

$$P^{-1} = [\mathrm{id}_E]_{\mathcal{A},\mathcal{B}}^{-1} = [\mathrm{id}_E^{-1}]_{\mathcal{B},\mathcal{A}} = [\mathrm{id}_E]_{\mathcal{B},\mathcal{A}}.$$

$$P[v]_{\mathcal{B}} = [\mathrm{id}_E]_{\mathcal{A},\mathcal{B}}[v]_{\mathcal{B}} = [\mathrm{id}_E(v)]_{\mathcal{A}} = [v]_{\mathcal{A}}.$$

$$P[h]_{\mathcal{B},\mathcal{B}}P^{-1} = [\mathrm{id}_E]_{\mathcal{A},\mathcal{B}}[h]_{\mathcal{B},\mathcal{B}}[\mathrm{id}_E]_{\mathcal{B},\mathcal{A}} = [\mathrm{id}_E \circ h \circ \mathrm{id}_E]_{\mathcal{A},\mathcal{A}} = [h]_{\mathcal{A},\mathcal{A}}.$$ □

Proposition. *Soient E et F deux espaces vectoriels de dimension finie, $\mathcal{A}$ et $\mathcal{B}$ deux bases de E, et $\mathcal{C}$ et $\mathcal{D}$ deux bases de F. Soient P la matrice de passage de $\mathcal{A}$ à $\mathcal{B}$, et Q la matrice de passage de $\mathcal{C}$ à $\mathcal{D}$. Alors pour toute application linéaire (homomorphisme) $h\colon E \to F$,*

$$[h]_{\mathcal{C},\mathcal{B}} = Q[h]_{\mathcal{D},\mathcal{A}}P.$$

Démonstration.

$$Q[h]_{\mathcal{D},\mathcal{A}}P = [\mathrm{id}_F]_{\mathcal{C},\mathcal{D}}[h]_{\mathcal{D},\mathcal{A}}[\mathrm{id}_E]_{\mathcal{A},\mathcal{B}} = [\mathrm{id}_F \circ h \circ \mathrm{id}_E]_{\mathcal{C},\mathcal{B}} = [h]_{\mathcal{C},\mathcal{B}}.$$ □

VI.4. Déterminant

VI.4.1. Déterminant d'une matrice carrée

Théorème. *Pour tout entier $n > 0$, il existe une unique fonction*

$$\det\colon \mathbf{R}^{n\times n} \to \mathbf{R}$$

qui satisfait les propriétés suivantes :

(1) *si B est obtenue de $A \in \mathbf{R}^{n\times n}$ par une opération $L_i \leftarrow L_i + \alpha L_j$, $i \neq j$, alors*

$$\det B = \det A,$$

(2) *si B est obtenue de $A \in \mathbf{R}^{n\times n}$ par une opération $L_i \leftarrow \alpha L_i$, alors*

$$\det B = \alpha \det A,$$

(3)

$$\det I = 1.$$

Esquisse d'une démonstration. L'unicité. Considérons d'abord une matrice échelonnée réduite $A \in \mathbf{R}^{n\times n}$. Il n'y a que deux possibilité : $A = I$ ou la dernière ligne de A est nulle. Si $A = I$, alors $\det A = 1$ par la condition (3). Si la dernière ligne de A est nulle, alors l'opération $L_n \leftarrow 0L_n$ ne change pas A, et donc $\det A = 0 \det A = 0$ par la condition (2).

Soit maintenant $A \in \mathbf{R}^{n\times n}$ une matrice quelconque. Par les opérations de type $L_i \leftarrow L_i + \alpha L_j$, $i \neq j$, et $L_i \leftarrow \alpha L_i$, $\alpha \neq 0$, on peut transformer A en une matrice échelonnée réduite $B \in \mathbf{R}^{n\times n}$ (comme dans l'algorithme de Gauss-Jordan). Comme déjà démontré, la valeur de $\det B$ est déterminée uniquement (et elle est 0 ou 1). La valeur de $\det A$ est uniquement déterminée par la valeur de $\det B$ grâce aux conditions (1) et (2).

L'existence. Pour montrer qu'une telle fonction det existe, on peut en donner une formule et vérifier que les propriétés sont satisfaites. Par exemple, on peut vérifier que pour les matrices 1×1 on peut définir det par la formule

$$\det\begin{pmatrix} a \end{pmatrix} = a.$$

Pour les matrices 2×2, on peut la définir par la formule

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Pour vérifier l'existence pour les matrices $n \times n$ en général, on aurait besoin d'utiliser des propriétés du *groupe* S_n des *permutations* de l'ensemble $\{1, \dots, n\}$, et en particulière une définition et des propriétés de la *signature* d'une permutation $\sigma \in S_n$. Si l'on admet tout cela, une formule qui convient à tout n est la suivante :

$$\det\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}. \qquad \square$$

Définition. Cette fonction det s'appelle *déterminant*.

Le déterminant d'une matrice A est noté $\det A$ ou $|A|$.
Voici quelques observations :

(1) Si une ligne de A est nulle, alors $\det A = 0$ (par la condition (2)). Plus généralement, si A est une matrice $n \times n$ et $\operatorname{rk} A < n$, alors $\det A = 0$.

(2) $\det \alpha A = \alpha^n \det A$ pour $A \in \mathbf{R}^{n\times n}$ (par la condition (2)).

(3) Le déterminant d'une matrice diagonale est égal au produit des éléments sur la diagonale (par les conditions (2) et (3)).

(4) Si B est obtenue de $A \in \mathbf{R}^{n\times n}$ par une opération $L_i \leftrightarrow L_j$, $i \neq j$, alors $\det B = -\det A$. En effet, l'opération $L_i \leftrightarrow L_j$, $i \neq j$, est la composition des opérations suivantes :

1. $L_i \leftarrow L_i + L_j$,
2. $L_j \leftarrow L_j - L_i$,
3. $L_i \leftarrow L_i + L_j$,
4. $L_j \leftarrow -L_j$.

Proposition. *Le déterminant d'une matrice triangulaire est égal au produit des éléments sur sa diagonale.*

Exemple.

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 5 & 6 \end{vmatrix} = 4 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 6 \end{vmatrix} = 4 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{vmatrix} = 24 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 24.$$

Lemme. *Si t est une opération élémentaire sur les lignes de matrice ou si t est une composition de telles opérations, et qu'une matrice A' est obtenue de A par t, alors t transforme tout produit AB en $A'B$. Si t est une opération élémentaire sur les colonnes de matrice ou si t est une composition de telles opérations, et qu'une matrice B' est obtenue de B par t, alors t transforme tout produit AB en AB'.*

Théorème. *Si $A, B \in \mathbf{R}^{n\times n}$, alors*

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

Démonstration. Appliquons une récurrence sur le nombre d'opérations élémentaires (sur les lignes) suffisant pour transformer A en une matrice échelonnée réduite par l'algorithme de Gauss-Jordan.

La base de récurrence : supposons que A est déjà échelonné réduite. Alors il n'y a que deux cas possibles : soit $A = I$, soit la dernière ligne de A est nulle. Si $A = I$, alors $AB = B$. Si la dernière ligne de A est nulle, alors la dernière ligne de AB est nulle aussi, et donc $\det A = 0$ et $\det AB = 0$. Dans les deux cas, l'égalité $\det AB = (\det A)(\det B)$ est claire.

Supposons maintenant qu'il y a une opération élémentaire reversible sur les lignes qui transforme A en une matrice A' telle que $\det A'B = (\det A')(\det B)$. Alors l'opération inverse transforme A' en A, et, d'après le lemme précédent, elle transforme aussi $A'B$ en AB. Il en résulte qu'il existe un $\alpha \in \mathbf{R}$ tel que $\det A = \alpha \det A'$ et $\det AB = \alpha \det A'B$, d'où, $\det AB = (\det A)(\det B)$. □

Théorème. *La fonction déterminant* det *satisfait les propriétés suivantes par rapport aux opérations élémentaires sur les colonnes* :

(1) *si B est obtenue de $A \in \mathbf{R}^{n\times n}$ par une opération $C_i \leftarrow C_i + \alpha C_j$, $i \neq j$, alors*

$$\det B = \det A,$$

(2) *si B est obtenue de $A \in \mathbf{R}^{n\times n}$ par une opération $C_i \leftarrow \alpha C_i$, alors*

$$\det B = \alpha \det A.$$

Démonstration. Écrivons $A = AI$. Soit T la matrice obtenue de I par l'opération élémentaire sur les colonnes par laquelle B est obtenue de A. Alors $B = AT$. Comme le déterminant respecte la multiplication des matrices, $\det B = (\det A)(\det T)$. Maintenant observons que quelque soit l'opération élémentaire sur les colonnes qu'on a utilisé pour obtenir T à partir de I, on pourrait obtenir T de I par une opération élémentaire du même type sur les lignes :

(1) si T est obtenu de I par l'opération $C_i \leftarrow C_i + \alpha C_j$ (où $i \neq j$), alors l'opération $L_j \leftarrow L_j + \alpha L_i$ appliqué à I va aussi donner T, et donc dans ce cas $\det T = \det I = 1$ et

$$\det B = 1 \det A = \det A,$$

(2) si T est obtenu de I par l'opération $C_i \leftarrow \alpha C_i$, alors l'opération $L_i \leftarrow \alpha L_i$ appliqué à I va aussi donner T, et donc dans ce cas $\det T = \alpha \det I = \alpha$ et

$$\det B = \alpha \det A.$$ □

Corollaire. *Pour tout matrice carrée A, $\det {}^{\mathbf{t}}A = \det A$.*

Voici un résumé des propriétés algébriques les plus importantes du déterminant. Toutes les matricés sont supposées d'être carrées d'une même taille.

(1)

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

(2)

$$A \text{ est inversible } \Leftrightarrow \det A \neq 0.$$

(3) Si A est inversible, alors

$$\det A^{-1} = \frac{1}{\det A}.$$

(4)

$$\det {}^{\mathbf{t}}A = \det A.$$

(5) Si deux matrices A et B coïncident sauf dans leurs lignes numéro i, et la ligne numéro i de B est la ligne numéro i de A multipliée par $\alpha \in \mathbf{R}$, alors

$$\det B = \alpha \det A.$$

Par exemple :

$$\begin{vmatrix} 1 & 3 \\ \mathbf{2} & \mathbf{4} \end{vmatrix} = 2 \begin{vmatrix} 1 & 3 \\ \mathbf{1} & \mathbf{2} \end{vmatrix}.$$

De même pour les colonnes.

(6) Si trois matrices A, B, C coïncident sauf dans leurs lignes numéro i, et la ligne numéro i de C est la somme des lignes numéro i de A et de B, alors

$$\det C = \det A + \det B.$$

Par exemple :

$$\begin{vmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} \mathbf{1} & \mathbf{0} \\ 3 & 4 \end{vmatrix} + \begin{vmatrix} \mathbf{0} & \mathbf{2} \\ 3 & 4 \end{vmatrix}.$$

De même pour les colonnes. Voici une indication pour une démonstration de cette propriété : traiter séparément le cas où la famille des lignes sans la ligne numéro i est liée. Dans ce cas, les trois déterminants sont nuls. Dans le cas contraire (si la famille des lignes sans la ligne numéro i est libre), la ligne numéro i d'une des deux matrices A ou B est une combinaison linéaire des lignes de l'autre.

Formules pour petites dimensions

(1) Pour une matrice 1×1 :

$$|(a)| = a.$$

(2) Pour une matrice 2×2 :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

(3) Pour une matrice 3×3 :

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$
$$= aei + bfg + cdh - afh - bdi - ceg.$$

VI.4.2. Comatrice

Définition. Soit A une matrice carrée. Le *cofacteur* de A d'indice (i, j), noté $\mathrm{Cof}_{i,j} A$, est $(-1)^{i+j}$ fois le déterminant de la sous-matrice de A obtenue en supprimant la ligne i et la colonne j. La *comatrice* de A, notée $\mathrm{com}\, A$, est la matrice des cofacteurs de A :

$$\mathrm{com}\, A = \begin{pmatrix} \mathrm{Cof}_{1,1} A & \cdots & \mathrm{Cof}_{1,n} A \\ \vdots & \ddots & \vdots \\ \mathrm{Cof}_{n,1} A & \cdots & \mathrm{Cof}_{n,n} A \end{pmatrix}.$$

Théorème. *Soit $A \in \mathbf{R}^{n \times n}$. Alors pour tout $k \in \{1, \ldots, n\}$,*

$$\sum_{i=1}^{n} [A]_{k,i} \,\mathrm{Cof}_{k,i} A = \det A,$$

et aussi

$$\sum_{i=1}^{n} [A]_{i,k} \,\mathrm{Cof}_{i,k} A = \det A.$$

Ce théorème permet de *développer* un déterminant *suivant une ligne* ou *suivant une colonne*.

Corollaire. *Pour toute $A \in \mathbf{R}^{n \times n}$,*

$$A\, {}^{\mathbf{t}}\mathrm{com}\, A = ({}^{\mathbf{t}}\mathrm{com}\, A) A = (\det A) I.$$

D'où,

$$A^{-1} = \frac{{}^{\mathbf{t}}\mathrm{com}\, A}{\det A}.$$

Parfois la matrice ${}^{\mathbf{t}}\mathrm{com}\, A$ est appelée *matrice complémentaire* de A.

VI.4.3. Déterminant d'une famille de vecteurs

Si $\mathcal{V} = (v_1, \ldots, v_n)$ est une famille de vecteurs dans un espace vectoriel E, on définit les opérations suivantes applicables à $\mathcal{V}$:

(1) $E_i \leftarrow E_i + \alpha E_j$, $i \neq j$:
remplacer l'élément v_i en position i par $v_i + \alpha v_j$,

(2) $E_i \leftarrow \alpha E_i$:
remplacer v_i en position i par αv_i,

(3) $E_i \leftrightarrow E_j$, $i \neq j$:
remplacer v_i en position i par v_j, et v_j en position j par v_i.

En fait, l'opération $E_i \leftrightarrow E_j$, $i \neq j$, est la composition des opérations suivantes :

(1) $E_i \leftarrow E_i + E_j$,

(2) $E_j \leftarrow E_j - E_i$,

(3) $E_i \leftarrow E_i + E_j$,

(4) $E_j \leftarrow -E_j$.

Théorème. *Soient E un espace vectoriel de dimension $n < \infty$ et $\mathcal{B}$ une base de E. Alors il existe une unique fonction $\det_{\mathcal{B}} \colon E^n \to \mathbf{R}$ qui satisfait les propriétés suivantes :*

(1) *si une famille $\mathcal{U} = (u_1, \dots, u_n)$ est obtenue d'une famille $\mathcal{V} = (v_1, \dots, v_n)$ dans E par une opération $E_i \leftarrow E_i + \alpha E_j$, $i \neq j$, alors*

$$\det_{\mathcal{B}} \mathcal{U} = \det_{\mathcal{B}} \mathcal{V},$$

(2) *si $\mathcal{U}$ est obtenue de $\mathcal{V} \in E^n$ par une opération $E_i \leftarrow \alpha E_i$, alors*

$$\det_{\mathcal{B}} \mathcal{U} = \alpha \det_{\mathcal{B}} \mathcal{V},$$

(3)

$$\det_{\mathcal{B}} \mathcal{B} = 1.$$

Ce théorème est admis dans ce cours.

Définition. Si $\mathcal{B}$ est une base d'un espace vectoriel E de dimension n et $(v_1, \dots, v_n)$ est une famille de n vecteurs dans E, alors $\det_{\mathcal{B}}(v_1, \dots, v_n)$ s'appelle le *déterminant* de la famille $(v_1, \dots, v_n)$ *relativement* à la base $\mathcal{B}$.

Observation : si un vecteur de la famille est nul, alors le déterminant est zéro (à cause de la condition (2)).

Proposition. *Si $\mathcal{U}$ est obtenue de $\mathcal{V} \in E^n$ par une opération $E_i \leftrightarrow E_j$, $i \neq j$, alors*

$$\det_{\mathcal{B}} \mathcal{U} = -\det_{\mathcal{B}} \mathcal{V}.$$

Théorème. *Soient E un espace vectoriel de dimension $n < \infty$, $\mathcal{B}$ une base de E, et $(v_1, \dots, v_n)$ une famille de n vecteurs dans E. Alors*

$$(v_1, \dots, v_n) \text{ est liée } \Leftrightarrow \det_{\mathcal{B}}(v_1, \dots, v_n) = 0.$$

Théorème. *Soient E un espace vectoriel de dimension n, $\mathcal{A}$ et $\mathcal{B}$ deux bases de E, et $\mathcal{F} = (v_1, \dots, v_n)$ une famille dans E. Alors*

$$\big(\det_{\mathcal{A}} \mathcal{B}\big)\big(\det_{\mathcal{B}} \mathcal{A}\big) = 1 \quad \text{et} \quad \big(\det_{\mathcal{A}} \mathcal{B}\big)\big(\det_{\mathcal{B}} \mathcal{F}\big) = \det_{\mathcal{A}} \mathcal{F}.$$

VI.4.4. Déterminant d'un endomorphisme linéaire

Proposition. *Si $h \colon E \to E$ est un endomorphisme d'un espace vectoriel E de dimension $n < \infty$, et $\mathcal{A} = (e_1, \dots, e_n)$ et $\mathcal{B} = (f_1, \dots, f_n)$ sont deux bases de E, alors*

$$\det_{(e_1,\dots,e_n)}(h(e_1), \dots, h(e_n)) = \det_{(f_1,\dots,f_n)}(h(f_1), \dots, h(f_n)).$$

Définition. Le *déterminant* d'un endomorphisme $h \colon E \to E$ d'un espace vectoriel E de dimension $n < \infty$, noté $\det h$, est défini par l'équation

$$\det h = \det_{(e_1,\dots,e_n)}(h(e_1), \dots, h(e_n)),$$

où $(e_1, \dots, e_n)$ est une base de E.

Proposition. *Si $h \colon E \to E$ est un endomorphisme d'un espace vectoriel E de dimension finie et $\mathcal{B}$ est une base de E, alors*

$$\det h = \det[h]_{\mathcal{B}}.$$

VI.4.5. Sens géométrique du déterminant

Si $\mathcal{B}$ est une base d'un espace vectoriel réel E de dimension $n < \infty$, et $\mathcal{F}$ est une famille de n vecteurs dans E, alors

(1) le volume (n-dimensionnel) du parallélépipède porté par la famille $\mathcal{F}$ est $|\det_{\mathcal{B}} \mathcal{F}|$ fois le volume du parallélépipède porté par la famille $\mathcal{B}$,

(2) si $\mathcal{F}$ est une base, alors $\det_{\mathcal{B}} \mathcal{F} > 0$ si et seulement si $\mathcal{F}$ et $\mathcal{B}$ ont la même orientation.

Pour un endomorphisme d'un espace vectoriel réel de dimension finie, la valeur absolue du déterminant représente le changement de volume, et le signe du déterminant représente le changement d'orientation. Plus précisément, si $h \colon E \to E$ est un endomorphisme d'un espace vectoriel E, alors

(1) h multiplie le volume par $|\det h|$,

(2) si $\det h > 0$, alors h préserve l'orientation, et si $\det h < 0$, alors h inverse l'orientation.

Volume dans un espace vectoriel réel

Pour toute base $\mathcal{B}$ d'un espace vectoriel réel E de dimension n, il existe une unique façon « naturelle » de définir le *volume* de tout parallélépipède n-dimensionnel dans E de façon que le volume du parallélépipède « porté » par les vecteurs de $\mathcal{B}$ soit 1.

Orientations de bases d'un espace vectoriel réel

Définition. Deux bases $(e_1, \dots, e_n)$ et $(f_1, \dots, f_n)$ d'un espace vectoriel réel sont dites avoir la même *orientation* si $(e_1, \dots, e_n)$ peut être transformée continûment en $(f_1, \dots, f_n)$ de telle façon, qu'en tout moment de sa transformation, la famille reste une base.

Théorème. *Tout espace vectoriel réel de dimension finie admet exactement deux orientations différentes.* (*Elle sont dites* opposées *l'une de l'autre*).

VI.5. Valeurs et vecteurs propres, spectre

Définition. Un *vecteur propre réel* de $A \in \mathbf{R}^{n\times n}$ est un $v \in \mathbf{R}^n$ non nul tel qu'il existe $\alpha \in \mathbf{R}$ tel que

$$Av = \alpha v.$$

Une *valeur propre réelle* de $A \in \mathbf{R}^{n\times n}$ est une valeur $\alpha \in \mathbf{R}$ telle qu'il existe un $v \in \mathbf{R}^n$ non nul tel que

$$Av = \alpha v.$$

On peut également définir les vecteurs et valeurs propres complexes.

Définition. Le *spectre réel* de $A \in \mathbf{R}^{n\times n}$, noté $\sigma_{\mathbf{R}}(A)$ ou $\mathrm{Sp}_{\mathbf{R}}(A)$, est l'ensemble de tous les $\alpha \in \mathbf{R}$ tels que la matrice $A - \alpha I$ n'est pas inversible :

$$\sigma_{\mathbf{R}}(A) \overset{\text{déf}}{=} \{\, \alpha \in \mathbf{R} \mid A - \alpha I \text{ n'est pas inversible} \,\}.$$

Le *spectre complexe* de $A \in \mathbf{C}^{n\times n}$, noté $\sigma_{\mathbf{C}}(A)$ ou $\mathrm{Sp}_{\mathbf{C}}(A)$, est l'ensemble de tous les $\alpha \in \mathbf{C}$ tels que la matrice $A - \alpha I$ n'est pas inversible :

$$\sigma_{\mathbf{C}}(A) \overset{\text{déf}}{=} \{\, \alpha \in \mathbf{C} \mid A - \alpha I \text{ n'est pas inversible} \,\}.$$

Le spectre complexe sera aussi appelé *spectre* tout court et noté σ ou Sp.

Exemples.

$$\sigma(I) = \{1\}, \qquad \sigma(\lambda I) = \{\lambda\}, \qquad \sigma\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \{1, 2, 3\}.$$

Proposition. *Le spectre d'une matrice carrée coïncide avec l'ensemble de ses valeurs propres. À la condition qu'on parle soit du spectre réel et valeurs propres réelles, soit du spectre complexe et valeurs propres complexes.*

Démonstration. Une valeur $\alpha \in \mathbf{R}$ est dans le spectre de $A \in \mathbf{R}^{n\times n} \Leftrightarrow A - \alpha I$ n'est pas inversible $\Leftrightarrow \mathrm{rk}(A - \alpha I) < n \Leftrightarrow$ l'équation $(A - \alpha I)v = 0$ admet une solution non nulle dans $\mathbf{R}^n \Leftrightarrow$ il existe $v \in \mathbf{R}^n$ non nul tel que $Av = \alpha v$.

Pareil pour $\mathbf{C}$. □

VI.6. Polynôme caractéristique

Pour chercher le spectre d'une matrice, on peut utiliser la propriété que une matrice carrée n'est pas inversible si et seulement si son déterminant est nul, d'où

$$\sigma_{\mathbf{R}}(A) = \{\, \alpha \in \mathbf{R} \mid \det(A - \alpha I) = 0 \,\},$$
$$\sigma(A) = \{\, \alpha \in \mathbf{C} \mid \det(A - \alpha I) = 0 \,\}.$$

Définition. Le *polynôme caractéristique* d'une matrice $A \in \mathbf{R}^{n\times n}$, noté χ_A, est défini par l'équation :

$$\chi_A(\alpha) = \det(A - \alpha I) \quad \text{pour tout nombre } \alpha.$$

Observations :

(1) $\chi_A(0) = \det A$,

(2) $\deg \chi_A = n$,

(3) le *terme dominant* (le terme du plus haut degré) de $\chi_A(X)$ est $(-1)^n X^n$,

(4) $\chi_{{}^t A} = \chi_A$.

Remarque. On peut aussi définir le polynôme caractéristique d'un endomorphisme d'un espace vectoriel de dimension finie.

Proposition. *Le spectre* (*complexe*) *d'une matrice carrée A est l'ensemble des racines* (*complexes*) *de χ_A.*

Définition. Si α est une valeur propre de $A \in \mathbf{R}^{n\times n}$, alors la multiplicité de α comme racine de χ_A s'appelle la *multiplicité algébrique* de α comme valeur propre de A.

Exemple. Soit

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Alors

$$\chi_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 2 \cdot 3 \\ = -2 - 5\lambda + \lambda^2.$$

Pour trouver le spectre de A, il faut résoudre l'équation

$$-2 - 5\lambda + \lambda^2 = 0.$$

On trouve ainsi :

$$\sigma(A) = \left\{ \frac{5 \pm \sqrt{33}}{2} \right\}.$$

Par ailleurs, on peut vérifier que $-2I - 5A + A^2$ est la matrice nulle.

Remarque. Un théorème remarquable (appelé *théorème de Cayley-Hamilton*) affirme que pour toute matrice carrée A, $\chi_A(A)$ est la matrice nulle. Ainsi, dans un certain sens, toute matrice carrée est une racine de son polynôme caractéristique.

VI.7. Trace

Définition. La *trace* d'une matrice carrée A, notée $\operatorname{tr} A$, est la somme des élément sur la diagonale de A :

$$\operatorname{tr} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \overset{\text{déf}}{=} a_{1,1} + a_{2,2} + \cdots + a_{n,n}.$$

Voici trois simples observations.

(1) L'application $A \mapsto \operatorname{tr} A$, $\mathbf{R}^{n\times n} \to \mathbf{R}$, est linéaire :

$$\operatorname{tr}(\alpha A + \beta B) = \alpha \operatorname{tr} A + \beta \operatorname{tr} B.$$

(2) Si $I \in \mathbf{R}^{n\times n}$ est la matrice identité, alors

$$\operatorname{tr} I = n.$$

(3) Pour toute matrice carrée A,

$$\operatorname{tr} A = \operatorname{tr} {}^{\mathbf{t}}A.$$

Théorème. *Si A est une matrice $m \times n$ et B est une matrice $n \times m$, alors*

$$\operatorname{tr} AB = \operatorname{tr} BA.$$

Corollaire. *Si A et P sont deux matrices carrées de même taille et P est inversible, alors*

$$\operatorname{tr} PAP^{-1} = \operatorname{tr} A.$$

Démonstration.

$$\operatorname{tr} PAP^{-1} = \operatorname{tr} P^{-1}PA = \operatorname{tr} A. \qquad \square$$

Théorème. *La fonction trace $\mathbf{R}^{n\times n} \to \mathbf{R}$ est complètement déterminée par les trois propriétés suivantes* :

(1) *elle est linéaire*,

(2) *la trace de AB est égale à la trace de BA pour toutes $A, B \in \mathbf{R}^{n\times n}$*,

(3) *la trace de la matrice identité $I \in \mathbf{R}^{n\times n}$ est n.*

La caractérisation de la trace par ces trois propriétés peut être utilisée comme la *définition* de la trace. Elle peut aussi servir à définir la trace des endomorphismes d'un espace vectoriel réel (ou complexe, etc.) de dimension finie : la trace sur l'ensemble $\operatorname{End}(E)$ des endomorphismes d'un espace vectoriel E est la fonction $\operatorname{tr}\colon \operatorname{End}(E) \to \mathbf{R}$ telle que

(1) tr est linéaire,

(2) $\operatorname{tr}(f \circ g) = \operatorname{tr}(g \circ f)$ pour tous $f, g \in \operatorname{End}(E)$,

(3) $\operatorname{tr} \mathrm{id}_E = \dim E$.

Ainsi que le déterminant, la trace d'une matrice (ou d'un endomorphisme) apparaît dans son polynôme caractéristique : si A est une matrice $n \times n$, alors le coefficient du terme de degré $n-1$ de son polynôme caractéristique est $(-1)^{n-1} \operatorname{tr} A$. Le polynôme caractéristique de A donc s'écrit comme

$$\chi_A(X) = \det A + \cdots + (-1)^{n-1}(\operatorname{tr} A)X^{n-1} + (-1)^n X^n.$$

Remarque. Si on définit l'*exponentielle* d'une matrice carrée par la formule

$$\exp A \overset{\text{déf}}{=} \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

(où $A^0 \overset{\text{déf}}{=} I$ est la matrice identité), alors on peut montrer que

$$\det \exp A = \exp \operatorname{tr} A \quad (= e^{\operatorname{tr} A}).$$

Les exponentielles de matrices carrées sont utilisées, entre autre, pour écrire les solutions de systèmes d'équations différentielles linéaires à coefficients constants.

VI.8. Diagonalisation et triangularisation

Définition. *Diagonaliser* une matrice carrée A veut dire trouver une matrice diagonale D et une matrice inversible P telles que

$$A = PDP^{-1}.$$

Trianguliariser une matrice carrée A veut dire trouver une matrice triangulaire T (supérieure ou inférieure) et une matrice inversible P telles que

$$A = PTP^{-1}.$$

Proposition. *Soient $A, P, D \in \mathbf{R}^{n\times n}$, P inversible et D diagonale avec $\lambda_1, \dots, \lambda_n$ sur la diagonale* :

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Soient $v_1, \dots, v_n \in \mathbf{R}^n$ les colonnes de P dans leur ordre. Alors $A = PDP^{-1}$ si et seulement si

$$Av_i = \lambda_i v_i, \quad \textit{pour tout } i = 1, \dots, n.$$

(*on regarde $v_1, \dots, v_n$ comme des éléments de $\mathbf{R}^n$ et au même temps comme des matrices colonnes*).

Démonstration. Soit $\mathcal{B}$ la base canonique de $\mathbf{R}^n$. Soit $h\colon \mathbf{R}^n \to \mathbf{R}^n$ l'endomorphisme de $\mathbf{R}^n$ tel que $[h]_{\mathcal{B}} = A$. Autrement dit, h est défini par la formule

$$h(x) = Ax \quad \text{pour tout } x \in \mathbf{R}^n$$

(éléments de $\mathbf{R}^n$ sont traités comme des matrices colonnes ici).

Posons $\mathcal{V} = (v_1, \dots, v_n)$. Alors $\mathcal{V}$ est une base de $\mathbf{R}^n$ (car $\operatorname{rk} P = n$) et $P = [\mathrm{id}]_{\mathcal{B},\mathcal{V}}$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{V}$. Ainsi,

$$A = [h]_{\mathcal{B}} = P[h]_{\mathcal{V}}P^{-1} \quad \text{et} \quad [h]_{\mathcal{V}} = P^{-1}[h]_{\mathcal{B}}P = P^{-1}AP.$$

D'où, $A = PDP^{-1}$ si et seulement si $[h]_{\mathcal{V}} = D$. Or, $[h]_{\mathcal{V}} = D$ si et seulement si

$$h(v_i) = \lambda_i v_i, \quad \text{pour tout } i = 1, \dots, n. \qquad \square$$

Voici une autre façon de démontrer cette proposition, qui ne mentionne pas les applications linéaires.

Si $A = PDP^{-1}$ où D est une matrice diagonale, alors $AP = PD$, d'où il est assez facile de voir que les colonnes de P sont des vecteurs propres de A, alors que les éléments de la diagonale de D sont les valeurs propres associées.

Réciproquement, si

$$Av_i = \lambda_i v_i, \quad v_i = \begin{pmatrix} x_{1,i} \\ \vdots \\ x_{n,i} \end{pmatrix} \quad \text{pour tout } i = 1, \dots, n,$$

où

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix}$$

(P est donc la matrice de passage de la base canonique à la base $(v_1, \dots, v_n)$), alors

$$AP = \begin{pmatrix} \lambda_1 x_{1,1} & \cdots & \lambda_n x_{1,n} \\ \vdots & & \vdots \\ \lambda_1 x_{n,1} & \cdots & \lambda_n x_{n,n} \end{pmatrix} = PD,$$

et donc

$$A = PDP^{-1}.$$

Corollaire. *Une matrice $A \in \mathbf{R}^{n\times n}$ est diagonalisable par des matrices réelles si et seulement si il existe une base de $\mathbf{R}^n$ formée de vecteurs propres de A.*

On va admettre sans démonstration le théorème suivant, qui sera revu avec plus de précisions (mais toujours sans démonstration) dans le chapitre sur les espaces vectoriels euclidiens comme un *théorème spectral*.

Théorème. *Toute matrice réelle symétrique est diagonalisable par des matrices réelles.*

Exemple. La matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'est pas diagonalisable, car ses vecteurs propres n'engendrent pas $\mathbf{R}^2$.

Pour trianguliariser une matrice $A \in \mathbf{R}^{n\times n}$, au lieu d'une base composée de vecteurs propres de A, on cherche une base $(v_1, \dots, v_n)$ de $\mathbf{R}^n$ telle que :

$$\begin{cases} Av_1 &= \lambda_1 v_1, \\ Av_2 &= t_{1,2}v_1 + \lambda_2 v_2, \\ &\vdots \\ Av_n &= t_{1,n}v_1 + t_{2,n}v_1 + \cdots + \lambda_n v_n, \end{cases}$$

avec $\lambda_i \in \mathbf{R}$ et $t_{i,j} \in \mathbf{R}$. (Le vecteur v_1 est donc un vecteur propre de A.) Soit P la matrice de passage de la base canonique à la base $(v_1, \dots, v_n)$ (les colonnes de P sont les vecteurs $v_1, \dots, v_n$ écrits en colonnes). Alors

$$A = P \begin{pmatrix} \lambda_1 & t_{1,2} & \cdots & t_{1,n} \\ 0 & \lambda_2 & & t_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} P^{-1}.$$

Théorème. *Une matrice réelle carrée A est triangularisable par des matrices réelles si et seulement si toutes les racines de χ_A sont réelles.*

Théorème (Théorème de Jordan)**.** *Pour toute matrice carrée $A \in \mathbf{C}^{n\times n}$, il existe $P, T \in \mathbf{C}^{n\times n}$ telles que*

$$A = PTP^{-1}$$

et T est triangulaire supérieure et diagonale-par-block :

$$T = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_k \end{pmatrix},$$

et chaque block $B_i \in \mathbf{C}^{k_i\times k_i}$ est de la forme

$$B_i = J_{k_i}(\lambda_i) \overset{déf}{=} \begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \lambda_i & & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_i \end{pmatrix}.$$

Les matrices carres $n \times n$ réelles peuvent être vues comme représentations canoniques des endomorphismes de l'espace vectoriel $\mathbf{R}^n$ (car $\mathbf{R}^n$ possède une base canonique). On va définir maintenant la *diagonalisation* et la *triangularisation* des endomorphismes des espaces vectoriels quelconques de dimension finie.

Définition. *Diagonaliser* un endomorphisme h d'un espace vectoriel E de dimension finie veut dire trouver une base $\mathcal{B}$ de E telle que $[h]_{\mathcal{B}}$ soit diagonale, et calculer $[h]_{\mathcal{B}}$. *Trianguliariser* h veut dire trouver une base $\mathcal{B}$ de E telle que $[h]_{\mathcal{B}}$ soit triangulaire, et calculer $[h]_{\mathcal{B}}$.

VII. Systèmes d'équations différentielles linéaires à coefficients constants[2h]

Ce chapitre est consacré à une des « applications » de l'algèbre linéaire : la résolution des systèmes d'équations différentielles linéaires à coefficients constants.

VII.1. Généralités

Parfois on va confondre une famille $(f_1, \dots, f_n)$ de n fonctions $D \to \mathbf{R}$ avec une seule fonction $F\colon D \to \mathbf{R}^n$ définie par :

$$F(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \quad t \in D.$$

On définit la *dérivée* d'une famille $(f_1, \dots, f_n)$ de n fonctions par :

$$(f_1, \dots, f_n)' \overset{\text{déf}}{=} (f_1', \dots, f_n').$$

Pour une fonction $F\colon D \to \mathbf{R}^n$ où

$$F(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \quad t \in D,$$

on définit sa *dérivée* par la formule

$$F'(t) \overset{\text{déf}}{=} \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix}, \quad t \in D.$$

Proposition. *Si $f_1, \dots, f_n\colon D \to \mathbf{R}$ sont des fonctions dérivables, et $A \in \mathbf{R}^{m\times n}$, alors*

$$\left(A \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \right)' = A \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}'.$$

Remarque. Si $F\colon D \to \mathbf{R}^n$ est une fonction à valeurs dans $\mathbf{R}^n$ et $A\colon D \to \mathbf{R}^{m\times n}$ est une fonction à valeurs dans $\mathbf{R}^{m\times n}$ (« une matrice variable »), alors on a la *règle de Leibniz* :

$$(AF)' = A'F + AF'.$$

Pareil pour le produit de matrices : si $A\colon D \to \mathbf{R}^{p\times q}$ et $B\colon D \to \mathbf{R}^{q\times r}$, alors

$$(AB)' = A'B + AB'.$$

VII.2. Résolution d'un système d'ordre 1 par diagonalisation ou par triangularisation

On va étudier seulement le cas de systèmes avec le même nombre d'équations que de fonctions inconnues.

Considérons un système d'équations différentielle linéaires d'ordre 1 à coefficients constants :

$$\left\{ \begin{array}{rcccccccl} x_1'(t) & = & a_{1,1}x_1(t) & + & \cdots & + & a_{1,n}x_n(t) & + & f_1(t), \\ & \vdots & & & & & & & \\ x_n'(t) & = & a_{n,1}x_1(t) & + & \cdots & + & a_{n,n}x_n(t) & + & f_n(t), \end{array} \right. \qquad t \in I. \quad (S)$$

Posons

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}, \quad F(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \quad t \in I.$$

La matrice A s'appelle la *matrice* du système (S).

Alors résoudre (S) revient à la même chose que résoudre l'équation différentielle « vectorielle »

$$X'(t) = AX(t) + F(t), \quad t \in I, \qquad (E)$$

où X est une fonction inconnue $I \to \mathbf{R}^n$, parce que on peut poser

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad t \in I.$$

Le système (S) est d'habitude facile à résoudre si la matrice A est diagonale, car dans ce case les n équations sont toutes indépendantes les unes des autres, donc il suffit de les résoudre indépendamment.

Si la matrice A n'est pas diagonale mais est triangulaire, on peut résoudre le système en résolvant les équations une par une, en utilisant à chaque étape les solutions des équations résolues précédemment.

Si la matrice A n'est pas diagonale mais est diagonalisable, on peut obtenir un système dont la matrice est diagonale par un « changement des variables ».

Supposons $A = PDP^{-1}$ où D est diagonale. Alors l'équation (E) s'écrit

$$X' = PDP^{-1}X + F.$$

Faisons le « changement des variables »

$$X(t) = PY(t), \quad Y(t) = P^{-1}X(t).$$

Posons en plus, pour simplifier l'écriture,

$$F(t) = PG(t), \quad G(t) = P^{-1}F(t).$$

Alors

$$\begin{aligned} & X' = PDP^{-1}X + F \\ \Leftrightarrow \quad & PY' = PDP^{-1}PY + PG \\ \Leftrightarrow \quad & Y' = DY + G. \end{aligned}$$

On peut maintenant trouver Y et ensuite X.

Si la matrice A dans (S) n'est pas diagonalisable mais est triangularisable, on peut obtenir un système dont la matrice est triangulaire par un « changement des variables » comme dans le cas avec une matrice diagonalisable.

Remarque. Si F est nulle, alors la solution de (E) avec la condition initiale $X(0) = X_0$ s'écrit comme

$$X(t) = \exp(tA)X_0.$$

Exemple. Résoudre le système d'équations différentielles :

$$\left\{ \begin{array}{rcrcrcrcl} y_1' & = & -6y_1 & & & + & 6y_3 & - & t, \\ y_2' & = & 2y_1 & - & 2y_2 & - & 3y_3 & + & t, \\ y_3' & = & -2y_1 & & & + & y_3. & & \end{array} \right. \qquad (S)$$

Posons

$$A = \begin{pmatrix} -6 & 0 & 6 \\ 2 & -2 & -3 \\ -2 & 0 & 1 \end{pmatrix}, \quad F(t) = \begin{pmatrix} -t \\ t \\ 0 \end{pmatrix}, \quad Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix}.$$

Alors le système (S) est équivalent à

$$Y' = AY + F.$$

Voici le polynôme caractéristique de A :

$$\chi_A(\lambda) = -12 - 16\lambda - 7\lambda^2 - \lambda^3 = -(\lambda + 2)^2(\lambda + 3).$$

Comme un vecteur propre pour la valeur propre -3, on peut prendre $(2, -1, 1)$. Pour la valeur propre -2, on peut prendre 2 vecteurs propres linéairement indépendants : $(3, 0, 2)$ et $(0, 1, 0)$. Les trois vecteurs choisis sont linéairement indépendants, donc

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

Posons

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad P = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix}.$$

Alors $A = PDP^{-1}$. Ainsi le système (S) est équivalent à

$$Y' = PDP^{-1}Y + F \quad \Leftrightarrow \quad (P^{-1}Y)' = DP^{-1}Y + P^{-1}F.$$

Posons

$$X(t) = P^{-1}Y(t), \quad G(t) = P^{-1}F(t).$$

Alors le système (S) est équivalent à

$$X' = DX + G.$$

On calcule P^{-1} et l'on trouve :

$$P^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 2 & 1 & -3 \\ 2 & 0 & -3 \end{pmatrix}.$$

D'où,

$$G(t) = \begin{pmatrix} t \\ -t \\ -2t \end{pmatrix}.$$

Ainsi, le système (S) pour Y est équivalent au système suivant pour X :

$$\begin{cases} x_1' &= -2x_1 &+ &t, \\ x_2' &= -2x_2 &- &t, \\ x_3' &= -3x_3 &- &2t. \end{cases}$$

La solution générale pour x_1, x_2, x_3 :

$$\begin{cases} x_1(t) &= \dfrac{t}{2} - \dfrac{1}{4} &+ &C_1 e^{-2t}, \\ x_2(t) &= -\dfrac{t}{2} + \dfrac{1}{4} &+ &C_2 e^{-2t}, \\ x_3(t) &= -\dfrac{2t}{3} + \dfrac{2}{9} &+ &C_3 e^{-3t}, \end{cases} \qquad \text{où } C_1, C_2, C_3 \in \mathbf{R}.$$

En utilisant le changement des variables

$$Y(t) = PX(t),$$

on trouve la solution générale pour y_1, y_2, y_3 :

$$\begin{cases} y_1(t) &= \dfrac{t}{6} - \dfrac{11}{36} &+ &3C_1 e^{-2t} + 2C_3 e^{-3t}, \\ y_2(t) &= \dfrac{t}{6} + \dfrac{1}{36} &+ &C_2 e^{-2t} - C_3 e^{-3t}, \\ y_3(t) &= \dfrac{t}{3} - \dfrac{5}{18} &+ &2C_1 e^{-2t} + C_3 e^{-3t}, \end{cases} \qquad \text{où } C_1, C_2, C_3 \in \mathbf{R}.$$

VII.3. Transformation d'une équation d'ordre n à une inconnue en un système d'ordre 1 à n inconnues

Considérons une équation différentielle linéaire d'ordre n à coefficients constants :

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \cdots + a_1 x'(t) + a_0 x(t) = f(t), \quad t \in I. \qquad (E)$$

Pour résoudre cette équation, on peut poser

$$\begin{cases} y_0 &= x, \\ y_1 &= x', \\ &\vdots \\ y_{n-1} &= x^{(n-1)}, \end{cases}$$

et résoudre le système :

$$\begin{cases} y_0' &= y_1, \\ y_1' &= y_2, \\ &\vdots \\ y_{n-1}' &= -a_0 y_0 - a_1 y_1 - \cdots - a_{n-1} y_{n-1} + f. \end{cases} \qquad (S)$$

Cependant, cette méthode n'est pas la plus directe.

Soit A la matrice du système (S) :

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}.$$

Pour diagonaliser ou trianguler A, il faudrait d'abord calculer son polynôme caractéristique χ_A et trouver ses racines. Il se trouve que

$$\chi_A(\lambda) = (-1)^n (a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n),$$

ce qui ressemble beaucoup au *polynôme caractéristique* de l'équation (E) défini dans un cours d'analyse ou dans un cours d'équations différentielles :

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

En fait,

$$\chi_A(\lambda) = (-1)^n P(\lambda).$$

(D'ailleurs, la matrice A s'appelle parfois la *matrice compagnon* du polynôme P.)

L'équation (E) s'écrit autrement comme

$$\left(\frac{d^n}{dt^n} + a_{n-1}\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_1\frac{d}{dt} + a_0\right) x(t) = f(t),$$

ou, en utilisant le polynôme P, comme

$$P\left(\frac{d}{dt}\right)x(t) = f(t).$$

Supposons qu'on a trouvé toutes les racines $\lambda_1, \dots, \lambda_n$ de χ_A et de P et a factorisé ces deux polynômes :

$$\chi_A(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda), \qquad P(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n).$$

En utilisant cette factorisation de P, on peut réécrire l'équation (E) comme

$$\left(\left(\frac{d}{dt} - \lambda_n\right)\cdots\left(\frac{d}{dt} - \lambda_1\right)\right)x(t) = f(t).$$

Posons

$$\begin{cases} y_0(t) &= x(t), \\ y_1(t) &= \left(\frac{d}{dt} - \lambda_1\right)y_0(t) &= y_0'(t) - \lambda_1 y_0(t), \\ y_2(t) &= \left(\frac{d}{dt} - \lambda_2\right)y_1(t) &= y_1'(t) - \lambda_2 y_1(t), \\ &\vdots \\ y_{n-1}(t) &= \left(\frac{d}{dt} - \lambda_{n-1}\right)y_{n-2}(t) &= y_{n-2}'(t) - \lambda_{n-1}y_{n-2}(t). \end{cases}$$

Alors on obtient le système suivant :

$$\begin{cases} \left(\frac{d}{dt} - \lambda_1\right)y_0(t) &= y_1(t) \\ &\vdots \\ \left(\frac{d}{dt} - \lambda_{n-1}\right)y_{n-2}(t) &= y_{n-1}(t) \\ \left(\frac{d}{dt} - \lambda_n\right)y_{n-1}(t) &= f(t) \end{cases} \Leftrightarrow \begin{cases} y_0' &= \lambda_1 y_0 + y_1 \\ &\vdots \\ y_{n-2}' &= \lambda_{n-1}y_{n-2} + y_{n-1} \\ y_{n-1}' &= \lambda_n y_{n-1} + f \end{cases}$$

On peut le résoudre en trouvant d'abord y_{n-1}, ensuite y_{n-2}, et ainsi de suite jusqu'à $y_0 = x$.

Par ailleurs, si f est nulle, alors il y a une formule assez simple pour la solution générale de (E) en fonction des racines de P (avec des exponentielles et des polynômes).

VIII. Applications bilinéaires[1h]

[...]

IX. Applications bilinéaires symétriques à valeurs réelles[1h]

[...]

X. Espaces vectoriels euclidiens[3h]

X.1. Produit scalaire

En géométrie euclidienne classique, on définit le *produit scalaire* de deux vecteurs u et v comme le nombre

$$\|u\|\,\|v\|\cos\theta,$$

où $\|u\|$ est la longueur de u, $\|v\|$ est la longueur de v, et θ est l'angle entre u et v. Or, dans les espaces vectoriels « purs », on n'a pas des notions de longueurs ni d'angles. En fait, en algèbre linéaire on définit les longueurs et les angles à partir du produit scalaire, qui est défini par des axiomes.

Définition. Soit E un espace vectoriel réel. Un *produit scalaire* dans E est n'importe quelle application $E \times E \to \mathbf{R}$ qui à chaque couple d'éléments $x, y \in E$ associe un nombre réel, appelé leur produit scalaire et noté $\langle x,\, y\rangle$, de telle façon, que les propriétés suivantes (axiomes) soient satisfaites :

(1) pour tous $x, y, z \in E$ et $\alpha, \beta \in \mathbf{R}$,

$$\begin{aligned}\langle x,\, \alpha y + \beta z\rangle &= \alpha\,\langle x,\, y\rangle + \beta\,\langle x,\, z\rangle\,,\\ \langle \alpha x + \beta y,\, z\rangle &= \alpha\,\langle x,\, z\rangle + \beta\,\langle y,\, z\rangle\end{aligned}$$

(on dit que $\langle\cdot,\,\cdot\rangle$ est *bilinéaire*),

(2) pour tous $x, y \in E$,

$$\langle y,\, x\rangle = \langle x,\, y\rangle$$

(on dit que $\langle\cdot,\,\cdot\rangle$ est *symétrique*),

(3) pour tout $x \in E$,

$$x \neq \vec{0} \;\Rightarrow\; \langle x,\, x\rangle > 0$$

(on dit que $\langle\cdot,\,\cdot\rangle$ est *défini positif*).

Exemples. (1) Le *produit scalaire canonique* dans $\mathbf{R}^n$ est défini comme suit : si $x, y \in \mathbf{R}^n$, $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$, alors

$$\langle x,\, y\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

(2) Soit $E = C([0,1])$ l'espace des fonctions continues $[0,1] \to \mathbf{R}$. Pour $f, g \in E$, on peut définir leur produit scalaire par la formule :

$$\langle f,\, g\rangle = \int_{t=0}^{1} f(t)g(t)\,dt\,.$$

Définition. Soit E un espace vectoriel réel muni d'un produit scalaire. La *norme*, ou la *longueur*, d'un vecteur $x \in E$, notée $\|x\|$, est définie par :

$$\|x\| \stackrel{\text{déf}}{=} \sqrt{\langle x,\, x\rangle}.$$

L'*angle* entre deux vecteurs non nuls $x, y \in E$ est l'unique nombre réel θ tel que

$$\cos\theta = \frac{\langle x,\, y\rangle}{\|x\|\,\|y\|} \quad \text{et} \quad 0 \leqslant \theta \leqslant \pi.$$

Définition. On dit que deux vecteurs u et v sont *orthogonaux*, et on écrit $u \perp v$, si $\langle u,\, v\rangle = 0$. On dit que deux ensembles de vecteurs A et B sont *orthogonaux*, et on écrit $A \perp B$, si pour tous $a \in A$ et $b \in B$, $a \perp b$.

Définition. Un *espace vectoriel euclidien*[1] est un espace vectoriel réel, d'habitude de dimension finie, muni d'un produit scalaire.

X.2. Quelques identités remarquables

Les identités suivantes se vérifient de façon directe, en utilisant la définition de la norme et les axiomes du produit scalaire.

La *règle du parallélogramme* :

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2\,.$$

L'*identité de polarisation réelle*[2] :

$$\langle x,\, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4}.$$

[1] Un *espace euclidien* est un espace affine.

[2] Beaucoup plus intéressante est l'*identité de polarisation* pour le *produit scalaire complexe* dans les espaces complexes (en particulier, dans les *espaces de Hilbert*) :

$$\langle x,\, y\rangle = \frac{\|x+y\|^2 + \boldsymbol{i}\,\|x-\boldsymbol{i}y\|^2 - \|x-y\|^2 - \boldsymbol{i}\,\|x+\boldsymbol{i}y\|^2}{4}.$$

La définition du produit scalaire complexe est assez similaire à celle du produit scalaire réel. Une particularité du produit scalaire complexe est la propriété suivante, qui remplace en partie la bilinéarité du produit scalaire réel :

$$\langle \alpha x,\, \beta y\rangle = \bar{\alpha}\beta\,\langle x,\, y\rangle \quad \text{pour } \alpha, \beta \in \mathbf{C}$$

($\bar{\alpha}$ est le conjugué complexe de α).

X.3. Deux inégalités classiques

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Lemme (Théorème de Pythagore). *Si $x, y \in E$ et $x \perp y$, alors*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2\,.$$

Démonstration.

$$\begin{aligned}\|x+y\|^2 &= \langle x+y,\, x+y\rangle = \langle x,\, x\rangle + \langle x,\, y\rangle + \langle y,\, x\rangle + \langle y,\, y\rangle \\ &= \|x\|^2 + \|y\|^2\,,\end{aligned}$$

car $\langle x,\, y\rangle = \langle y,\, x\rangle = 0$. □

Théorème (Inégalité de Cauchy-Schwarz). *Si $x, y \in E$, alors*

$$|\langle x,\, y\rangle| \leqslant \|x\|\,\|y\|\,.$$

En plus, l'égalité a lieu si et seulement si les vecteurs x et y sont linéairement dépendants.

Démonstration. Si $x = \vec{0}$, alors $|\langle x,\, y\rangle| = 0 = \|x\|\,\|y\|$.

Supposons $x \neq \vec{0}$. Alors $\langle x,\, x\rangle > 0$. Posons

$$y_1 = \frac{\langle x,\, y\rangle}{\langle x,\, x\rangle}x \quad \text{et} \quad y_2 = y - y_1.$$

(Ici y_1 est le *projeté orthogonal de y sur* $\mathrm{Vect}(x)$.)

On vérifie que $\langle x,\, y_1\rangle = \langle x,\, y\rangle$:

$$\langle x,\, y_1\rangle = \left\langle x,\, \frac{\langle x,\, y\rangle}{\langle x,\, x\rangle}x \right\rangle = \frac{\langle x,\, y\rangle}{\langle x,\, x\rangle}\,\langle x,\, x\rangle = \langle x,\, y\rangle\,.$$

D'où,

$$\langle x,\, y_2\rangle = \langle x,\, y\rangle - \langle x,\, y_1\rangle = 0 \quad \text{et} \quad \langle y_1,\, y_2\rangle = \frac{\langle x,\, y\rangle}{\langle x,\, x\rangle}\,\langle x,\, y_2\rangle = 0.$$

Donc, par le théorème de Pythagore,

$$\|y\|^2 = \|y_1\|^2 + \|y_2\|^2\,.$$

En particulier, $\|y\| \geqslant \|y_1\|$, et l'égalité a lieu si et seulement si $y = y_1$.

Ainsi,

$$\|x\|\,\|y\| \geqslant \|x\|\,\|y_1\| = \|x\|\left|\frac{\langle x,\, y\rangle}{\langle x,\, x\rangle}\right|\|x\| = |\langle x,\, y\rangle|\,.$$

L'égalité a lieu si et seulement si $y = y_1$, c'est-à-dire, si et seulement si x et y sont linéairement dépendants (cela reste pour le lecteur à vérifier). □

Exemples. (1) Si $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbf{R}$, alors

$$|x_1y_1 + \cdots + x_ny_n| \leqslant \sqrt{x_1^2 + \cdots + x_n^2}\cdot\sqrt{y_1^2 + \cdots + y_n^2}.$$

(2) Si f et g sont deux fonctions continues $[0,1] \to \mathbf{R}$, alors

$$\left|\int_{t=0}^{1} f(t)g(t)\,dt\right| \leqslant \sqrt{\int_{t=0}^{1}(f(t))^2\,dt}\cdot\sqrt{\int_{t=0}^{1}(g(t))^2\,dt},$$

et l'égalité a lieu si et seulement si une des fonctions est un multiple de l'autre.

Théorème (Inégalité triangulaire)**.** *Si $x, y \in E$, alors*

$$\|x + y\| \leqslant \|x\| + \|y\|\,.$$

Démonstration. D'après l'inégalité de Cauchy–Schwarz,

$$\langle x,\, y\rangle \leqslant |\langle x,\, y\rangle| \leqslant \|x\|\,\|y\|\,.$$

Donc,

$$\begin{aligned}\|x + y\|^2 &= \langle x + y,\, x + y\rangle = \langle x,\, x\rangle + \langle x,\, y\rangle + \langle y,\, x\rangle + \langle y,\, y\rangle \\ &= \|x\|^2 + 2\,\langle x,\, y\rangle + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

D'où, $\|x + y\| \leqslant \|x\| + \|y\|$. □

X.4. L'orthogonal d'une partie

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Définition. Si $A \subset E$, l'*orthogonal* de A, noté $A^\perp$, est

$$A^\perp \overset{\text{déf}}{=} \{\, x \in E \mid x \perp A \,\} = \{\, x \in E \mid \forall y \in A,\ x \perp y \,\}.$$

Si $u \in E$, l'*orthogonal* de u, noté $u^\perp$, est

$$u^\perp \overset{\text{déf}}{=} \{\, x \in E \mid x \perp u \,\} = \{u\}^\perp.$$

Observations :

(1) $E^\perp = \{\vec{0}\}$,

(2) $\{\vec{0}\}^\perp = E$,

(3) $A^\perp \perp A$ pour tout $A \subset E$,

(4) $A \subset (A^\perp)^\perp$ pour tout $A \subset E$,

(5) $A^\perp$ est un sous-espace vectoriel de E pour tout $A \subset E$,

(6) $\mathrm{Vect}(A) \subset (A^\perp)^\perp$ pour tout $A \subset E$.

Exemple. Soit $u = (1, 2, 3)$ dans $\mathbf{R}^3$ muni du produit scalaire canonique. Alors

$$u^\perp = \{\, (x, y, z) \in \mathbf{R}^3 \mid x + 2y + 3z = 0 \,\}.$$

Proposition. *Si $A \subset B \subset E$, alors $B^\perp \subset A^\perp$.*

Corollaire. *Pour tout $A \subset E$, $((A^\perp)^\perp)^\perp = A^\perp$.*

Démonstration. D'un côté, $A \subset A^{\perp\perp}$, et donc $(A^{\perp\perp})^\perp \subset A^\perp$. De l'autre côté, $A^\perp \subset (A^\perp)^{\perp\perp}$. □

Proposition. *Si A et B sont deux parties de E, alors*

(1) $A^\perp \cap B^\perp = (A \cup B)^\perp = (A + B)^\perp$,

(2) $A^\perp \cup B^\perp \subset (A \cap B)^\perp$.

Théorème. *Si F est un sous-espace vectoriel de E et $\dim F < \infty$, alors*

(1) $E = F \oplus F^\perp$,

(2) $(F^\perp)^\perp = F$.

Corollaire. *Si* $\dim E = n < \infty$, *alors il existe* n *sous-espaces vectoriel* $V_1, \ldots, V_n$ *de* E *tels que* :

(1) $E = V_1 \oplus \cdots \oplus V_n$,

(2) $\dim V_i = 1$ *pour tout* i,

(3) $V_i \perp V_j$ *si* $i \neq j$.

X.5. Familles orthogonales et familles orthonormées

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Définition. Une famille $(e_1, \ldots, e_n)$ est dite *orthogonale* si et seulement si

$$\langle e_i, e_j \rangle = 0 \qquad \text{pour tous } i, j \in \{1, \ldots, n\} \text{ tels que } i \neq j.$$

Définition. Une famille $(e_1, \ldots, e_n)$ est dite *orthonormée* ou *orthonormale* si et seulement si pour tous $i, j \in \{1, \ldots, n\}$,

$$\langle e_i, e_j \rangle = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

Exemple. La base canonique de $\mathbf{R}^n$ est orthonormée par rapport au produit scalaire canonique.

Proposition. *Toute famille orthonormée est libre.*

Démonstration. Soient $(e_1, \ldots, e_n)$ une famille orthonormée et $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$. Supposons que $\alpha_1 e_1 + \cdots + \alpha_n e_n = \vec{0}$. Alors

$$\begin{aligned} 0 &= \|\alpha_1 e_1 + \cdots + \alpha_n e_n\|^2 = \langle \alpha_1 e_1 + \cdots + \alpha_n e_n, \alpha_1 e_1 + \cdots + \alpha_n e_n \rangle \\ &= \alpha_1^2 + \cdots + \alpha_n^2, \end{aligned}$$

et donc $\alpha_1 = \cdots = \alpha_n = 0$. □

Théorème. *Si* $\dim E < \infty$, *alors* E *admet un base orthonormée.*

Démonstration. Soient $V_1, \ldots, V_n$ des sous-espaces vectoriels de E tels que :

(1) $E = V_1 \oplus \cdots \oplus V_n$,

(2) $\dim V_i = 1$ pour tout i,

(3) $V_i \perp V_j$ si $i \neq j$

(voir la section X.4). Pour tout i de 1 à n, soit $v_i \in V_i$ tel que $\|v_i\| = 1$. Alors $(v_1, \ldots, v_n)$ est une base orthonormée de E. □

X.6. Matrice de Gram

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Définition. Soit $\mathcal{F} = (x_1, \ldots, x_n)$ une famille de vecteurs dans E. La *matrice de Gram* de $\mathcal{F}$ est la matrice

$$\begin{pmatrix} \langle x_1, x_1 \rangle & \cdots & \langle x_1, x_n \rangle \\ \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \cdots & \langle x_n, x_n \rangle \end{pmatrix}$$

En particulier, la matrice de Gram d'une famille est la matrice identité si et seulement si la famille est orthonormée.

Théorème. *Soit* G *la matrice de Gram d'une famille finie* $\mathcal{F} = (x_1, \ldots, x_n)$ *de vecteurs dans* E. *Alors*

(1) $\det G \geqslant 0$,

(2) $\det G > 0$ *si et seulement si* $\mathcal{F}$ *est libre.*

[...]

X.7. Projeté orthogonal d'un vecteur

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Lemme. *Soient* F *un sous-espace de* E, $x \in E$ *et* $y_1, y_2 \in F$ *tels que* $x - y_1 \perp F$ *et* $x - y_2 \perp F$. *Alors* $y_1 = y_2$.

Démonstration. Comme l'ensemble des vecteurs de E orthogonaux à F forme un sous-espace vectoriel de E (noté « $F^\perp$ »), on a :

$$y_1 - y_2 = (x - y_2) - (x - y_1) \perp F.$$

En particulier,

$$y_1 - y_2 \perp y_1 - y_2,$$

c'est-à-dire,

$$\langle y_1 - y_2,\, y_1 - y_2 \rangle = 0.$$

D'où, $y_1 - y_2 = \vec{0}$ et donc $y_1 = y_2$. □

Définition. Soit F un sous-espace vectoriel de E. Le *projeté orthogonal* de $x \in E$ sur F est l'unique élément $y \in F$ tel que $x - y \perp F$.

Remarque. Dans cette définition, il est entendu qu'un élément $y \in F$ tel que $x - y \perp F$ existe. Dans des espaces de dimension infinie, cela n'est pas toujours le cas, et il peut arriver qu'il n'y a pas de projeté orthogonal de x sur F.

Exercice. Soit $x \in E$.

(1) Donner une simple formule pour le projeté de x sur E.

(2) Donner une simple formule pour le projeté de x sur $\{\vec{0}\}$.

Proposition. *Soient F un sous-espace vectoriel de E et x un vecteur dans E. Le projeté orthogonal de x sur F existe si et seulement si $x \in F \oplus F^\perp$.*

En particulier, si $\dim F < \infty$, alors $E = F \oplus F^\perp$, et donc, pour tout $x \in E$, le projeté orthogonal de x sur F existe bien.

Notation. Soient F un sous-espace vectoriel de E et $x \in E$. On va noter « $\mathrm{proj}_F x$ » le projeté orthogonal de x sur F (s'il existe). Si $F = \mathrm{Vect}(u)$, alors au lieu de « proj_F » ou « $\mathrm{proj}_{\mathrm{Vect}(u)}$ » on pourra écrire « proj_u ».

Exemple. Soient u un vecteur non nul et x un vecteur quelconque dans E. Alors le projeté orthogonal de x sur $\mathrm{Vect}(u)$ est donné par la formule :

$$\mathrm{proj}_u x = \frac{\langle u,\, x\rangle}{\langle u,\, u\rangle} u = \left\langle \frac{u}{\|u\|},\, x \right\rangle \frac{u}{\|u\|}.$$

Proposition. *Soit F un sous-espace vectoriel de E. Alors :*

(1) *si $x, y \in E$ et qu'il existe les projetés orthogonaux de x et de y sur F, alors il existe le projeté orthogonal de $x + y$ sur F, et*

$$\mathrm{proj}_F(x + y) = \mathrm{proj}_F x + \mathrm{proj}_F y,$$

(2) *il existe le projeté orthogonal de $\vec{0}$ sur F, et*

$$\mathrm{proj}_F \vec{0} = \vec{0},$$

(3) *si $\alpha \in \mathbf{R}$, $x \in E$ et qu'il existe le projeté orthogonal de x sur F, alors il existe le projeté orthogonal de αx sur F, et*

$$\mathrm{proj}_F(\alpha x) = \alpha \,\mathrm{proj}_F x.$$

Théorème. *Soient F_1 et F_2 deux sous-espaces vectoriels de E orthogonaux l'un à l'autre ($F_1 \perp F_2$), x un vecteur dans E, et y_i le projeté orthogonal de x sur F_i pour $i = 1, 2$. Alors $y_1 + y_2$ est le projeté orthogonal de x sur $F_1 \oplus F_2$. En particulier, si $x \in F_1 \oplus F_2$, alors $y_1 + y_2 = x$.*

X.8. Orthonormalisation de Gram-Schmidt

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Théorème (Orthonormalisation de Gram-Schmidt)**.** *Soit $(v_1, \ldots, v_n)$ une famille libre dans E. Alors il existe une unique famille orthonormée $(e_1, \ldots, e_n)$ telle que* :

(1) *pour tout $k \in \{1, \ldots, n\}$,*

$$\mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(v_1, \ldots, v_k),$$

(2) *pour tout $k \in \{1, \ldots, n\}$,*

$$\langle e_k,\, v_k \rangle > 0.$$

Esquisse d'une demonstration. Posons

$$F_k = \mathrm{Vect}(v_1, \ldots, v_k) \qquad \text{pour } k \in \{0, \ldots, n\}.$$

Ainsi :

$$F_0 = \{\vec{0}\}, \qquad F_1 = \mathrm{Vect}(v_1), \qquad \ldots, \qquad F_n = \mathrm{Vect}(v_1, \ldots, v_n).$$

Notons que $\dim F_k = k$ pour tout $k \in \{0, \dots, n\}$.

Pour tout $k \in \{1, \dots, n\}$, soit U_k le *complèment orthogonal* de F_{k-1} dans F_k, c'est-à-dire, U_k est le sous-espace de F_k tel que :

$$U_k \perp F_{k-1} \quad \text{et} \quad U_k \oplus F_{k-1} = F_k.$$

Alors :

(1) $\dim U_i = 1$ pour tout i,

(2) $U_i \perp U_j$ si $i \neq j$,

(3)

$$U_1 = F_1, \qquad U_1 \oplus U_2 = F_2, \qquad \dots, \qquad U_1 \oplus \cdots \oplus U_n = F_n.$$

On peut montrer par récurrence qu'une famille orthonormée $(e_1, \dots, e_n)$ satisfait à la première condition donnée dans le théorème si et seulement si

$$e_k \in U_k \qquad \text{pour tout } k \in \{1, \dots, n\}.$$

Chaque espace U_k contient exactement deux vecteurs de norme 1, qui sont opposés l'un de l'autre. La deuxième condition donnée dans le théorème exclue au moins un des deux choix et ainsi assure l'unicité de e_k.

Il ne reste qu'à vérifier qu'un des deux vecteurs unitaires de U_k convient pour le rôle de e_k, c'est à dire, que v_k n'est pas orthogonal à U_k. Or, l'ensemble de vecteurs de F_k orthogonaux à U_k est F_{k-1}, alors que $v_k \in F_k \setminus F_{k-1}$. Donc, v_k n'est pas orthogonal à U_k. □

Supposons qu'une famille libre $(v_1, \dots, v_n)$ dans E est donnée, et qu'on cherche à « calculer » la famille orthonormée $(e_1, \dots, e_n)$ comme dans l'énoncé du théorème. Voici un algorithme de calcul de $(e_1, \dots, e_n)$ à partir de $(v_1, \dots, v_n)$, dit l'*algorithme de Gram-Schmidt*.

Posons

$$F_k = \operatorname{Vect}(v_1, \dots, v_k) \qquad \text{pour } k \in \{0, \dots, n\},$$

comme dans l'esquisse d'une démonstration du théorème.

Posons

$$\begin{aligned} u_1 &= \operatorname{proj}_{F_0^\perp} v_1 = v_1, \\ u_2 &= \operatorname{proj}_{F_1^\perp} v_2, \\ &\vdots \\ u_n &= \operatorname{proj}_{F_{n-1}^\perp} v_n. \end{aligned}$$

La famille $(u_1, \dots, u_n)$ est orthogonale (mais pas nécessairement orthonormale) et possède la propriété que, pout tout $k \in \{0, \dots, n\}$,

$$\operatorname{Vect}(u_1, \dots, u_k) = \operatorname{Vect}(v_1, \dots, v_k) = F_k.$$

On peut dire que la famille $(u_1, \dots, u_n)$ est obtenue en *orthogonalisant* la famille $(v_1, \dots, v_n)$.

Pour des applications pratique, on a besoin de formules plus explicites pour $u_1, \dots, u_n$, données en termes des opérations « basiques ». Pour en trouver, observons que

$$\operatorname{proj}_{u_1} x + \cdots + \operatorname{proj}_{u_k} x + \operatorname{proj}_{F_k^\perp} x = x,$$

et donc

$$\operatorname{proj}_{F_k^\perp} x = x - \operatorname{proj}_{u_1} x - \cdots - \operatorname{proj}_{u_k} x,$$

pour tout $k \in \{0, \dots, n\}$ et pour tout $x \in E$.

Ainsi :

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \operatorname{proj}_{u_1} v_2 \qquad = v_2 - \frac{\langle u_1,\, v_2\rangle}{\langle u_1,\, u_1\rangle} u_1, \\ &\vdots \\ u_n &= v_n - \sum_{i=1}^{n-1} \operatorname{proj}_{u_i} v_n \quad = v_n - \sum_{i=1}^{n-1} \frac{\langle u_i,\, v_n\rangle}{\langle u_i,\, u_i\rangle} u_i. \end{aligned} \tag{$*$}$$

La famille $(e_1, \dots, e_n)$ est obtenue de $(u_1, \dots, u_n)$ par la *normalisation* :

$$e_k = \frac{u_k}{\|u_k\|}, \quad k \in \{1, \dots, n\}.$$

La version de l'algorithme de Gram-Schmidt présentée ci-dessus est parfois dite la version *classique*. Lorsque cette version de l'algorithme est réalisé sur un ordinateur en utilisant l'arithmétique binaire arrondie (par exemple, le type `float` des langages C ou Python), des erreurs d'arrondi ont une tendance de s'accumuler pendant l'étape de l'orthogonalisation ($*$). Pour modérer cette accumulation des erreurs d'arrondi, on peut modifier la version classique en effectuant l'orthogonalisation par les formules suivantes :

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= \operatorname{proj}_{u_1^\perp} v_2, \\ u_3 &= \operatorname{proj}_{u_2^\perp}(\operatorname{proj}_{u_1^\perp} v_3), \\ &\vdots \\ u_n &= \operatorname{proj}_{u_{n-1}^\perp}(\cdots(\operatorname{proj}_{u_2^\perp}(\operatorname{proj}_{u_1^\perp} v_n))\cdots), \end{aligned} \tag{$**$}$$

où chaque expression de la forme « $\mathrm{proj}_{x^\perp} y$ » peut être évaluée selon la formule :

$$\mathrm{proj}_{x^\perp} y = y - \mathrm{proj}_x y = y - \frac{\langle x,\, y\rangle}{\langle x,\, x\rangle} x.$$

XI. Applications linéaires entre espaces vectoriels euclidiens[1h]

XI.1. Vecteurs singuliers et valeurs singulières d'une application linéaire

Dans cette section, E et F sont deux espaces vectoriels euclidiens de dimensions finies.

Définition. Soit $v \in E$ un vecteur unitaire ($\|v\| = 1$). Posons $V = \mathrm{Vect}(v)$. Le vecteur v est dit un *vecteur singulier à droite* pour une application linéaire $h\colon E \to F$ si et seulement si il existe un sous-espace vectoriel U dans F tel que :

$$h(V) \subset U \quad \text{et} \quad h(V^\perp) \subset U^\perp.$$

Définition. Soit $u \in F$ un vecteur unitaire ($\|u\| = 1$). Posons $U = \mathrm{Vect}(u)$. Le vecteur u est dit un *vecteur singulier à gauche* pour une application linéaire $h\colon E \to F$ si et seulement si il existe un sous-espace vectoriel V dans E tel que :

$$h(V) \subset U \quad \text{et} \quad h(V^\perp) \subset U^\perp.$$

Observons que, lorsque $h\colon E \to F$ est une application linéaire, V est un sous-espace vectoriel de E et U est un sous-espace vectoriel de F, on a les équivalence :

$$\begin{aligned} h(V) \subset U \quad &\Leftrightarrow \quad U^\perp \perp h(V) \qquad \text{et} \\ h(V^\perp) \subset U^\perp \quad &\Leftrightarrow \quad U \perp h(V^\perp). \end{aligned}$$

Remarque. Dans certains contextes où les vecteurs singuliers à droite ou à gauche apparaissent, la condition qu'ils soient unitaires (par définition) est sans importance : il suffit qu'il soient non nuls (ou même pas). Malheureusement, il n'y a pas de terme courant pour les « vecteurs singuliers non nuls, mais pas forcement unitaires ».

La « droite – gauche » terminologie provient de la façon comment l'application de fonctions est écrite. Si $h\colon E \to F$ est une application linéaire et $v \in E$, alors on peut appliquer h à v, et le résultat s'écrit « $h(v)$ », avec « v » *à droite* de « h ». Pour cela, les vecteurs de E définis comme *singuliers* sont dits *singuliers à droite*. Si $h\colon E \to F$ est une application linéaire et $u \in F$, alors on peut définir une application linéaire $u^*\colon F \to \mathbf{R}$ par la formule :

$$u^*(y) = \langle u, y\rangle \qquad \text{pour tout } y \in F,$$

et alors, quel que soit $v \in E$, on a :

$$u^*(h(v)) = \langle u, h(v)\rangle \in \mathbf{R}.$$

Dans « $u^*(h(v))$ », « u^* » est écrit *à gauche* de « h » (et « v » – à droite). Pour cela, les vecteurs de F définis comme *singuliers* sont dits *singuliers à gauche*.

Définition. Un nombre réel $\sigma \geqslant 0$ est dit une *valeur singulière* pour une application linéaire $h\colon E \to F$ si et seulement si il existe $v \in E$ et $u \in F$ tels que :

(1) $\|v\| = 1$ et v est un vecteur singulier à droite pour h,

(2) $\|u\| = 1$ et u est un vecteur singulier à gauche pour h,

(3) $h(v) = \sigma u$.

La proposition suivante est presque évidente, mais elle mérite d'être énoncée :

Proposition. *Soit $h\colon E \to F$ une application linéaire. Alors* :

(1) *tout vecteur unitaire de $\ker h$ est singulier à droite pour h,*

(2) *tout vecteur unitaire de $(\operatorname{img} h)^\perp$ est singulier à gauche pour h.*

Remarque. Vu que tout vecteur unitaire dans $\ker h$ est singulier à droite pour h, et que tout vecteur unitaire dans $(\operatorname{img} h)^\perp$ est singulier à gauche pour h, ces vecteurs singuliers n'ont pas beaucoup d'intérêt. Malheureusement, il n'y a pas de terminologie courante pour distinguer les vecteurs singuliers à droite qui sont dans $\ker h$ des autres, et les vecteurs singuliers à gauche qui sont dans $(\operatorname{img} h)^\perp$ des autres. On peut éventuellement dire que les vecteurs unitaires de $\ker h$ et de $(\operatorname{img} h)^\perp$ sont des vecteurs singuliers *triviaux*.

Proposition. *Si $v \in E$ est un vecteur singulier à droite pour une application linéaire $h\colon E \to F$, et que $v \notin \ker h$, alors $v \in (\ker h)^\perp$.*

Proposition. *Si $u \in F$ est un vecteur singulier à gauche pour une application linéaire $h\colon E \to F$, et que $u \notin (\operatorname{img} h)^\perp$, alors $u \in \operatorname{img} h$.*

Proposition. *Soit $h\colon E \to F$ une application linéaire. Si $(v_1, \dots, v_r)$ est une base orthonormée de $(\ker h)^\perp$ composée de vecteurs singuliers à droite pour h, alors $(h(v_1), \dots, h(v_r))$ est une base orthonormée de $\operatorname{img} h$ composée de vecteurs singuliers à gauche pour h.*

Théorème. *Quelle que soit une application linéaire $h\colon E \to F$, l'espace E admet une base orthonormée composée de vecteurs singuliers à droite pour h, et l'espace F admet une base orthonormée composée de vecteurs singuliers à gauche pour h.*

Ce théorème peut être démontré en passant par le lemme suivant :

Lemme. *Soit $h\colon E \to F$ une application linéaire. Soit $u \in E \setminus \{\vec{0}\}$ tel que*

$$\|h(x)\| \leqslant \|h(u)\| \qquad \textit{pour tout } x \in E \textit{ tel que } \|x\| \leqslant \|u\|.$$

Alors

$$h(x) \perp h(u) \qquad \textit{pour tout } x \in E \textit{ tel que } x \perp u.$$

La conclusion du lemme peut être écrite ainsi : $h(u^\perp) \subset h(u)^\perp$.

Les hypothèses de ce lemme peuvent être traduites ainsi : un vecteur non nul u maximise la valeur de $\|h(x)\|$ sur l'ensemble des vecteurs x de la même norme (qui peut être pensé comme une sphère de rayon $\|u\|$). Ainsi, ce lemme s'agit d'un extremum de la fonction qui à chaque vecteur x dans un ensemble donné associe la valeur $\|h(x)\|$. L'étude des extrema de fonctions fait partie du domaine de l'*analyse mathématique*, et la démonstration suivante de ce lemme est inspirée du calcul différentiel :

Démonstration du lemme. Sans perte de généralité, supposons que $\|u\| = 1$.

Posons

$$g(x) = \langle h(x), h(x)\rangle = \|h(x)\|^2 \quad \text{pour tout} \quad x \in E.$$

Ainsi on a une fonction $g\colon E \to \mathbf{R}$.

D'après les hypotheses, $g(x) \leqslant g(u)$ pour tout $x \in E$ tel que $\|x\| \leqslant 1$.

Soit $v \in E$ tel que $v \perp u$. Définissons la fonction $f\colon \mathbf{R} \to E$ ainsi :

$$f(t) = \frac{u + tv}{\|u + tv\|}.$$

Alors, $f(0) = u$ et $\|f(t)\| = 1 = \|u\|$ pour tout $t \in \mathbf{R}$. Donc, $g \circ f$ atteint un maximum global en 0. D'où, $(g \circ f)'(0) = 0$ si $(g \circ f)'(0)$ existe. On a :

$$(g \circ f)(t) = g(f(t)) = \|h(f(t))\|^2 = \frac{\|h(u+tv)\|^2}{\|u+tv\|^2}$$

$$= \frac{\langle h(u+tv),\, h(u+tv)\rangle}{\langle u+tv,\, u+tv\rangle}$$

$$= \frac{\langle h(u),\, h(u)\rangle + t\,\langle h(u),\, h(v)\rangle + t\,\langle h(v),\, h(u)\rangle + t^2\,\langle h(v),\, h(v)\rangle}{\langle u,\, u\rangle + t^2\,\langle v,\, v\rangle}$$

$$= \frac{\|h(u)\|^2 + 2t\,\langle h(u),\, h(v)\rangle + t^2\,\|h(v)\|^2}{1 + t^2\,\|v\|^2}.$$

Par un calcul usuel, on trouve que

$$(g \circ f)'(0) = 2\,\langle h(u),\, h(v)\rangle\,.$$

Donc, $\langle h(u),\, h(v)\rangle = 0$ et $h(v) \perp h(u)$. □

Démonstration du théorème. `[...]` □

`[...]`

XII. Endomorphismes d'espaces vectoriels euclidiens[1h]

XII.1. Projections orthogonales

Dans cette section, E est un espace vectoriel réel muni d'un produit scalaire.

Définition. Soit F un sous-espace vectoriel de E. Une *projection orthogonale* de E sur F est une application $p\colon E \to F$ qui à tout $x \in E$ associe un projeté orthogonale de x sur F : $x - p(x) \perp F$.

Proposition. *Soit F un sous-espace vectoriel de E. Une projection orthogonale p de E sur F existe si et seulement si $E = F \oplus F^\perp$, et dans ce cas l'application p est unique et linéaire.*

En particulier, si $\dim F < \infty$, alors $E = F \oplus F^\perp$, et donc la projection orthogonale de E sur F existe bien.

Notation. Soit F un sous-espace vectoriel de E. On va noter « proj_F » la projection orthogonale de E sur F (si elle existe). Si $F = \mathrm{Vect}(u)$, alors au lieu de « proj_F » ou « $\mathrm{proj}_{\mathrm{Vect}(u)}$ » on pourra écrire « proj_u ».

Exemple. Soit u un vecteur non nul. Alors la projection orthogonale sur $\mathrm{Vect}(u)$ est donnée par la formule :

$$\mathrm{proj}_u\, x = \frac{\langle u,\, x\rangle}{\langle u,\, u\rangle} u = \left\langle \frac{u}{\|u\|},\, x \right\rangle \frac{u}{\|u\|}.$$

Théorème. *Un endomorphisme $p\colon E \to E$ est une projection orthogonale (de E sur $\mathrm{img}\, p$) si et seulement si les deux conditions suivantes sont satisfaites :*

(1) $p|_{\mathrm{img}\, p} = \mathrm{id}_{\mathrm{img}\, p}$,

(2) $\ker p \perp \mathrm{img}\, p$.

Théorème. *Un endomorphisme $h\colon E \to E$ est une projection orthogonale (de E sur $\mathrm{img}\, h$) si et seulement si les deux conditions suivantes sont satisfaites :*

(1) $h^2 = h$ *(on dit que h est* idempotent*), et*

(2) $\langle h(x), y\rangle = \langle x, h(y)\rangle$ *pour tous* $x, y \in E$ (*on dit que* h *est* symétrique).

Proposition. *Soient* $h\colon E \to E$ *un endomorphisme de* E, *et* A *la matrice de* h *par rapport à une n'importe quelle base orthonormée de* E. *Alors* h *est une projection orthogonale si et seulement si*

(1) $A^2 = A$ (*on dit que* A *est* idempotente) *et*

(2) ${}^{\mathbf{t}}A = A$ (*A est symétrique*).

Dans les hypothèses de cette proposition, il est sous-entendu que E est de dimension finie.

Théorème. *Soient* F_1 *et* F_2 *deux sous-espaces vectoriels de* E *orthogonaux l'un à l'autre* ($F_1 \perp F_2$), p_1 *la projection orthogonale sur* F_1, *et* p_2 *la projection orthogonale sur* F_2. *Alors* $p_1 + p_2$ *est la projection orthogonale sur* $F_1 \oplus F_2$. *En particulier, si* $F_1 \oplus F_2 = E$ (*donc,* $F_2 = F_1^\perp$ *et* $F_1 = F_2^\perp$), *alors* $p_1 + p_2 = \mathrm{id}_E$.

XII.2. Endomorphismes orthogonaux et matrices orthogonales

Dans cette section, E est un espace vectoriel euclidien de dimension finie.

Définition. Un endomorphisme $h\colon E \to E$ est dit *orthogonal*[1] si et seulement si

$$\langle h(x), h(y)\rangle = \langle x, y\rangle \quad \text{pour tous} \quad x, y \in E.$$

Définition. Une matrice $A \in \mathbf{R}^{n\times n}$ est dite *orthogonale* si et seulement si

$${}^{\mathbf{t}}AA = I = A\,{}^{\mathbf{t}}A \quad (\text{c'est-à-dire, } {}^{\mathbf{t}}A = A^{-1}).$$

Observations :

(1) si A est orthogonale, alors $\det A \in \{\pm 1\}$,

(2) $A \in \mathbf{R}^{n\times n}$ est orthogonale si et seulement si les colonnes de A forment une base orthonormée de $\mathbf{R}^n$ (par rapport au produit scalaire canonique), et si et seulement si les lignes de A forment une base orthonormée de $\mathbf{R}^n$.

[1] C'est bien le terme qui est utilisé, mais il serait probablement plus logique de l'appeler un endomorphisme *orthonormé*.

Proposition. *Soit* $\mathcal{B}$ *une base orthonormée de* E. *Alors un endomorphisme* $h\colon E \to E$ *est orthogonal si et seulement si la matrice de* h *par rapport à* $\mathcal{B}$ *est orthogonale.*

Ainsi, si la matrice d'un endomorphisme h par rapport à une base orthonormée est orthogonale, alors h est orthogonal, et si h est orthogonal, alors les matrices de h par rapport à toutes les bases orthonormées sont orthogonales.

Exercice. Démontrer la dernière proposition.

Proposition. *Soit* $\mathcal{B}$ *une base orthonormée de* E. *Une base* $\mathcal{C}$ *de* E *est orthonormée si et seulement si la matrice de passage de* $\mathcal{B}$ *à* $\mathcal{C}$ *est orthogonale.*

Exercice. Démontrer cette proposition.

XII.3. Endomorphismes symétriques

Dans cette section, E est un espace vectoriel euclidien de dimension finie.

Définition. Un endomorphisme $h\colon E \to E$ est dit *symétrique* si et seulement si

$$\langle h(x), y\rangle = \langle x, h(y)\rangle \quad \text{pour tous} \quad x, y \in E.$$

Autrement dit, $h\colon E \to E$ est symétrique si et seulement si

$$\langle y, h(x)\rangle = \langle x, h(y)\rangle \quad \text{pour tous} \quad x, y \in E.$$

Exercice. Montrer que toute projection orthogonale est un endomorphisme symétrique.

Proposition. *Soit* $\mathcal{B}$ *une base orthonormée de* E. *Alors un endomorphisme* $h\colon E \to E$ *est symétrique si et seulement si la matrice de* h *par rapport à* $\mathcal{B}$ *est symétrique.*

Ainsi, si la matrice d'un endomorphisme h par rapport à une base orthonormée est symétrique, alors h est symétrique, et si h est symétrique, alors les matrices de h par rapport à toutes les bases orthonormées sont symétriques.

Exercice. Démontrer la dernière proposition.

Proposition. *Soit* h *un endomorphisme symétrique de* E. *Soit* F *un sous-espace de* E *stable par* h. *Alors* $F^\perp$, *lui aussi, est stable par* h.

Exercice. Démontrer cette proposition.

Théorème. *Quel que soit un endomorphisme symétrique $h\colon E \to E$, l'espace E admet une base orthonormée composée de vecteurs propres de h.*

Ce théorème peut être démontré en passant par le lemme suivant :

Lemme. *Soit h un endomorphisme symétrique de E. Soit $u \in E \setminus \{\vec{0}\}$ tel que*

(1) $\langle x, h(x)\rangle \leqslant \langle u, h(u)\rangle$ *pour tout $x \in E$ tel que $\|x\| = \|u\|$.*

Alors :

(2) $\langle u, h(x)\rangle = \langle x, h(u)\rangle = 0$ *pour tout $x \in E$ tel que $x \perp u$,*

(3) $u^\perp$ *est stable par h,*

(4) *u est un vecteur propre de h.*

Les hypothèses de ce lemme peuvent être traduites ainsi : un vecteur non nul u maximise la valeur de $\langle x, h(x)\rangle$ sur l'ensemble des vecteurs x de la même norme (qui peut être pensé comme une sphère de rayon $\|u\|$).

Démonstration du lemme. Sans perte de généralité, supposons que $\|u\| = 1$.

Posons

$$g(x) = \langle x, h(x)\rangle \quad \text{pour tout} \quad x \in E.$$

Ainsi on a une fonction $g\colon E \to \mathbf{R}$.

Soit $v \in E$ tel que $v \perp u$. Définissons la fonction $f\colon \mathbf{R} \to E$ ainsi :

$$f(t) = \frac{u + tv}{\|u + tv\|}.$$

Alors, $f(0) = u$ et $\|f(t)\| = 1 = \|u\|$ pour tout $t \in \mathbf{R}$. Donc, $g \circ f$ atteint un maximum global en 0. D'où, $(g \circ f)'(0) = 0$ si $(g \circ f)'(0)$ existe. On a :

$$(g \circ f)(t) = g(f(t)) = \langle f(t), h(f(t))\rangle = \frac{\langle u + tv, h(u + tv)\rangle}{\|u + tv\|^2}$$

$$= \frac{\langle u + tv, h(u + tv)\rangle}{\langle u + tv, u + tv\rangle}$$

$$= \frac{\langle u, h(u)\rangle + t\langle u, h(v)\rangle + t\langle v, h(u)\rangle + t^2\langle v, h(v)\rangle}{\langle u, u\rangle + t^2\langle v, v\rangle}$$

$$= \frac{\langle u, h(u)\rangle + t(\langle u, h(v)\rangle + \langle v, h(u)\rangle) + t^2\langle v, h(v)\rangle}{1 + t^2\|v\|^2}.$$

Par un calcul usuel, on trouve que

$$(g \circ f)'(0) = \langle u, h(v)\rangle + \langle v, h(u)\rangle = 2\langle u, h(v)\rangle = 2\langle v, h(u)\rangle.$$

Donc, $\langle u, h(v)\rangle = \langle v, h(u)\rangle = 0$.

La première partie de la conclusion est démontrée. Les deux autres s'en déduisent facilement. □

Démonstration du théorème. [...] □

Corollaire (Diagonalisation orthogonale d'un endomorphisme symétrique)**.** *Quel que soit un endomorphisme symétrique $h\colon E \to E$, il existe une base orthonormée $\mathcal{B}$ de E telle que la matrice de h par rapport à $\mathcal{B}$ est diagonale.*

Corollaire (Diagonalisation orthogonale d'une matrice symétrique)**.** *Quelle que soit une matrice réelle symétrique A, il existe une matrice réelle diagonale D et une matrice réelle orthogonale P telles que*

$$A = PDP^{-1} = PD({}^{\mathrm{t}}P).$$

Le dernier théorème et ses corollaires représentent un cas relativement simple de la *théorie spectrale* de l'*analyse fonctionnelle*.

XII.4. Décomposition polaire d'un endomorphisme

Dans cette section, E est un espace vectoriel euclidien de dimension finie.

[...]

Bibliographie

[1] Paul Halmos, *Finite-dimensional vector spaces [Espaces vectoriels de dimension finie]* (en anglais).

[2] Serge Lang, *Introduction to linear algebra [Introduction à l'algèbre linéaire]* (en anglais).

[3] Serge Lang, *Linear algebra [Algèbre linéaire]* (en anglais).

[4] Mikhail Postnikov, *Leçons de géométrie. Semestre II : algèbre linéaire et géométrie différentielle* (traduit du russe).