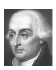
Algèbre linéaire

Christophe Antonini[1], Olivier Teytaud[2], Pierre Borgnat[3], Annie Chateau[4], and Edouard Lebeau[5]

[1]Enseignant en CPGE, Institut Stanislas, Cannes
[2]Chargé de rechercher INRIA, Université d'Orsay, Orsay
[3]Chargé de recherche CNRS, ENS Lyon, Lyon
[4]Maitre de conférence, Université Montpellier-2, Montpellier
[5]Enseignant en CPGE, Lycée Henri Poincaré, Nancy

26 juillet 2023



Chapitre d'introduction à l'algèbre linéaire.

1 Algèbre linéaire

Classiquement, après avoir étudié groupes, anneaux et corps, on se penche en général sur les espaces vectoriels. L'exemple le plus basique d'espace vectoriel est bien sûr $\mathbb{R}^n$, mais des cas plus abstraits sont importants : les espaces vectoriels de polynômes, de matrices de même type de morphismes linéaires. Une erreur classique est d'oublier que les combinaisons linéaires portent sur un nombre fini de vecteurs ; sinon, il convient de se placer dans un espace de Banach, où l'on peut donner sens à des sommes infinies.

Après quelques généralités (1.1), on se penchera sur les applications linéaires (1.2), les sommes de sous-espaces vectoriels (1.3), les quotients (1.4), les espaces affines (1.5), les bases d'espaces vectoriels (1.5), un peu de dualité (1.7), la transposition (1.7.3), les $\mathbb{K}$-algèbres (1.8), avant de se pencher sur quelques compléments et exercices (1.9).

1.1 Généralités

Définition 0.1 **Espace vectoriel**
Soit $\mathbb{K}$ un corps. $(E,+,.)$ est un **$\mathbb{K}$-espace vectoriel** (ou un **espace vectoriel sur** $\mathbb{K}$) si

- $(E,+)$ est un groupe abélien
- $.$ est une application de $\mathbb{K}\times E$ dans E
- $\forall(\lambda,\mu,x,y)\ (\lambda+\mu).x=\lambda.x+\mu.x \wedge \lambda.(x+y)=\lambda.x+\lambda.y \wedge (\lambda.\mu).x=\lambda.(\mu.x)\wedge 1.x=x$

Les éléments de E sont appelés **vecteurs**, les éléments de $\mathbb{K}$ sont appelés **opérateurs** ou **scalaires**. Le neutre pour $+$ est noté 0. « . » est appelé **produit externe**.

Des exemples classiques sont : les vecteurs dans le plan ou l'espace usuel, le corps $\mathbb{K}$ lui-même avec pour produit externe le produit usuel, les polynômes sur $\mathbb{K}$, éventuellement à plusieurs indéterminées et/ou de degré borné ; comme on le verra un peu plus loin avec les espaces produits, $\mathbb{K}^n$ muni des lois que l'on verra est un $\mathbb{K}$-espace vectoriel ; l'ensemble $\mathbb{R}^{\mathbb{N}}$ des suites réelles (resp. $\mathbb{C}^{\mathbb{N}}$ des suites complexes) aussi.

$\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel, et plus généralement tout espace vectoriel sur $\mathbb{C}$ est aussi un espace vectoriel sur $\mathbb{R}$.

Les propriétés suivantes sont importantes :

•$0.x = 0$

•$\lambda.0 = 0$

•$\lambda.x = 0 \rightarrow \lambda = 0 \vee x = 0$

•$(-\lambda).x = \lambda.(-x) = -(\lambda.x)$

•$\lambda.(x - y) = \lambda.x - \lambda.y$

•$(\lambda - \mu)x = \lambda.x - \mu.x$

DÉFINITION 0.2 **Convexité dans un espace vectoriel**
On appelle **segment** d'extrémités x et y dans un $\mathbb{R}$-espace vectoriel l'ensemble des $t.x + (1-t).y$ pour $t \in [0,1]$. (la définition s'étend au cas d'un $\mathbb{C}$-espace vectoriel en utilisant t réel dans $[0,1]$).

Une partie A d'un $\mathbb{K}$-espace vectoriel (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$) est dite **convexe** si tout segment d'extrémités dans A est inclus dans A.

Étant donnée A une partie convexe d'un $\mathbb{K}$-espace vectoriel une application f de A dans $\mathbb{R}$ est dite **convexe** si étant donnés x et y dans A et t dans $[0,1]$ on a $f(t.x+(1-t).y) \leq t.f(x).(1-t).f(y)$. L'application $-f$ est alors dite **concave**.

Propriétés •L'intersection de deux convexes est un convexe.

•Une boule ouverte est convexe.

•Une boule fermée est convexe.

•Un segment est convexe.

•Une application f est convexe si et seulement si l'ensemble des $(x, f(x))$ est convexe dans le produit $A \times \mathbb{R}$.

•Une application convexe sur un intervalle $]a, b[$ est continue

•$x \mapsto e^x$, $x \mapsto x^n$ avec n pair sont convexes.

•$x \mapsto \ln(x)$ est concave.

Application 0.1 La notion de partie convexe est nécessaire pour définir les fonctions convexes, dont on verra une multitude d'applications en partie **??** (fonctions convexes). Cela servira aussi par exemple pour les résultats **??** (théorème de Cauchy), et surtout pour les théorème de Hahn-Banach (**??**) (aux multiples applications !). Une utilisation amusante de la convexité stricte sera donnée avec le résultat **??** concernant les billards. L'inégalité de Jensen **??** fournit aussi nombre d'applications.

On peut noter l'existence de la notion de convexité locale, qui permet de définir les espaces de Fréchet, dans lesquelles on peut généraliser notamment le théorème de l'application ouverte (cf th. **??**).

PROPOSITION 0.1 **Connexité dans un $\mathbb{R}$-espace vectoriel normé**
Une partie ouverte d'un $\mathbb{R}$-espace vectoriel normé est connexe si et seulement si elle est connexe par arcs si et seulement si entre tout x et tout y de cette partie il existe une ligne brisée.

Démonstration *Il est clair que l'existence d'une ligne brisée implique l'existence d'un chemin (donc la connexité par arcs) qui implique à son tour la connexité. Il suffit donc de montrer que la connexité implique l'existence d'une ligne brisée.*

•*On se donne U une telle partie, connexe, supposée non vide (sinon le problème est trivial)*

•*On se donne x_0 dans U*

•*On note V l'ensemble des x que l'on peut joindre à x_0 par une ligne brisée.*

•*V est non vide (il contient x_0)*

•*V est ouvert (si x est dans V, alors $B(x,\epsilon) \subset U$ pour ϵ assez petit, et donc $B(x,\epsilon) \subset V$ - car tout point de cette boule peut être relié à x par un segment, donc à x_0 par une ligné brisée).*

•*V est fermé ; en effet soit x dans U adhérent à V, alors il existe une boule ouverte centrée sur x intersectant V, donc x est dans V - car toute boule est convexe.*

•*V, fermé, ouvert, non vide d'un connexe U, est égal à U.*

La figure 1 illustre cette démonstration.

PROPOSITION 0.2 **Distance dans un connexe ouvert d'un $\mathbb{R}$-espace vectoriel normé**

Soit U un ouvert connexe d'un $\mathbb{R}$-espace vectoriel normé . Si on note $d(x, y)$ l'inf des longueurs des lignes brisées joignant x à y, alors d est une distance et définit la même topologie que la norme.

Démonstration *Pas difficile du tout !*

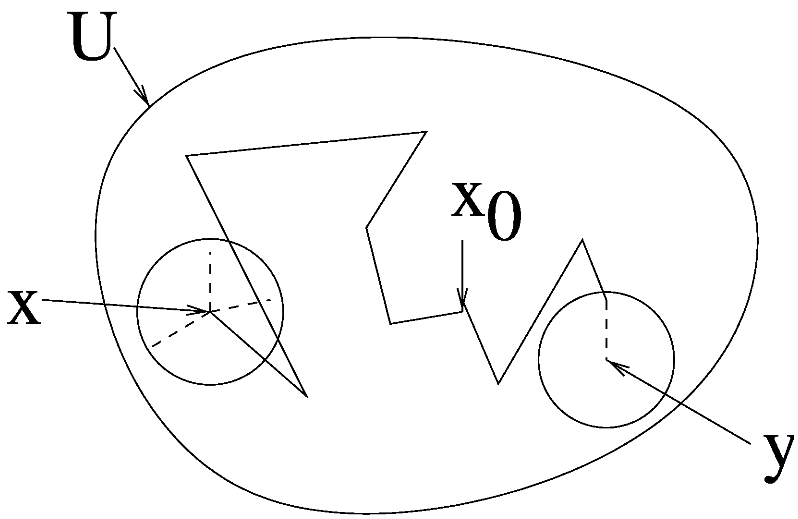


FIGURE 1 – Illustration de la proposition 0.1.
Commentaire : V, ensemble des points de U connectables à x par une ligne brisée, est ouvert (car tout y dans V est entouré d'une boule connectable à x en ajoutant juste un segment à y) et fermé (tout point z de l'adhérence de V est centre d'une boule intersectant V en un certain z', et peut donc être connecté à x via une ligne brisée contenant le segment $[zz']$ puis la ligne brisée de z' à x. Ouvert et fermé, non vide, V est égal à la composante connexe dans laquelle il se trouve, donc il est égal à U.

Définition 0.3 Espace vectoriel produit
Le produit de n espaces vectoriels sur un même corps $\mathbb{K}$, muni de l'addition terme à terme et avec pour multiplication $\lambda.(x_1, ..., x_n) = (\lambda.x_1, ..., \lambda.x_n)$ est un espace vectoriel sur $\mathbb{K}$. On l'appelle **espace vectoriel produit**.

Plus généralement, si $(E_i)_{i\in I}$ est une famille indexée sur I quelconque d'espaces vectoriels sur $\mathbb{K}$, alors on peut définir l'**espace vectoriel produit** $\prod_{i\in I} E_i$ avec là encore la somme terme à terme des composantes, et avec pour multiplication par λ la multiplication de chaque terme par λ.

On donne sans démonstration le théorème facile suivant :

Théorème 0.3
L'ensemble des fonctions de A dans E, noté E^A, avec E $\mathbb{K}$-espace vectoriel, est un $\mathbb{K}$-espace vectoriel. La somme de deux fonctions est la fonction qui à un élément associe la somme des deux images, et le produit d'un scalaire par une fonction est la fonction qui à un vecteur associe le produit du scalaire par l'image de ce vecteur. L'élément neutre est la fonction constante nulle.

Définition 0.4 Sous-espaces vectoriels
Une partie d'un espace vectoriel est un **sous-espace vectoriel** de cet espace lorsqu'elle est non vide, stable par addition et stable par multiplication par un scalaire.

On note qu'on peut se contenter de vérifier que la partie est non vide et que si λ et μ sont des scalaires et si x et y lui appartiennent, alors $\lambda.x + \mu.y$ appartient aussi à l'ensemble.

On peut également se contenter de montrer que 0 appartient à cette partie (souvent le plus facile en pratique) et que si x, y lui appartiennent et si λ est un scalaire, alors $x + \lambda.y$ appartient aussi à l'ensemble.

Proposition 0.4
Un sous-espace vectoriel est un espace vectoriel, et un espace vectoriel inclus dans un espace vectoriel (et muni des lois induites bien sûr) est un sous-espace vectoriel.

Bien que simple, la proposition ci-dessus doit être mentionnée. En effet, il n'en est pas de même de toutes les structures algébriques usuelles. Par exemple, un anneau inclus dans un anneau n'est pas toujours un sous-anneau ; il faut pour cela ajouter une condition, qui est le fait que l'élément neutre soit le même.

Les polynômes de degré plus petit que n sont un sous-espace de l'ensemble des polynômes. L'espace des fonctions bornées de A dans E est un sous-espace vectoriel de l'espace E^A des fonctions de A dans $\mathbb{R}$. L'ensemble des suites bornées de $\mathbb{K}$ est un sous-espace de l'espace des suites de $\mathbb{K}$. L'ensemble des suites de $\mathbb{K}$ nulles sauf pour un nombre fini de points (suites dites presque nulles) est aussi un sous-espace vectoriel de l'espace des suites de $\mathbb{K}$.

PROPOSITION 0.5
Une intersection quelconque de sous-espaces vectoriels est un sous-espace vectoriel.

Cela permet d'introduire la définition suivante :

DÉFINITION 0.5 **Sous-espace vectoriel engendré**
On appelle **sous-espace vectoriel** engendré par une partie A l'intersection de tous les sous-espaces vectoriels contenant A. On la note $Vect\ (A)$.

C'est le plus petit (pour l'inclusion) sous-espace vectoriel de E contenant A.

1.2 Applications linéaires

DÉFINITION 0.6 **Application linéaire**
Une application f d'un $\mathbb{K}$-espace vectoriel vers un autre est une **application linéaire** si $\forall(\lambda, x, y) \in \mathbb{K} \times E^2$:

$\bullet f(x + y) = f(x) + f(y)$

$\bullet f(\lambda.x) = \lambda.f(x)$

Une application linéaire est aussi appelée **morphisme d'espaces vectoriels**, ou **morphisme algébrique**. C'est en particulier un morphisme de groupes. Une application linéaire bijective est appelée **isomorphisme**, une application linéaire de E dans E est appelée **endomorphisme**. Un endomorphisme qui est aussi un isomorphisme est appelé **automorphisme**. L'inverse d'un isomorphisme est un isomorphisme, appelé **isomorphisme réciproque**.

On note $\mathcal{L}(E, F)$ l'ensemble des morphismes de E dans F ; c'est un sous-espace vectoriel de F^E. On note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E. On note $Isom(E, F)$ l'ensemble des isomorphismes de E dans F. On note $Aut(E)$ l'ensemble des automorphismes de E ; il est noté $GL(E)$ une fois qu'on l'a muni de la composition. $GL(E)$ est un groupe, appelé **groupe linéaire**.

La notation $E \simeq F$ signifie qu'il existe un isomorphisme de E dans F.

L'image réciproque d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel. On note $Ker\ f$ et on appelle **noyau** de f l'image réciproque de $\{0\}$, c'est un sous-espace vectoriel. Une application linéaire est injective si et seulement si son noyau est $\{0\}$. On notera que le noyau d'une application linéaire est le noyau du morphisme de groupes correspondant.

L'image directe d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel. On note $Im\ f$ et on appelle **image** de f l'ensemble $f(E) = \{f(x); x \in E\}$. Évidemment, f est surjective si et seulement si $Im\ f = F$.

On note Id la fonction identité de E ; c'est une application linéaire et un isomorphisme. On note $Inv\ f$ **l' ensembles des invariants de** f ; $Inv\ f = Ker(f - Id)$. On note $Opp\ f$ **l'ensemble des vecteurs changés en leur opposé** ; $Opp\ f = Ker(f + Id)$.

Si $\lambda \neq 0$, on appelle **homothétie de rapport** λ **d'un espace vectoriel** E l'application $x \mapsto \lambda.x$.

Enfin, on appelle **forme linéaire sur un espace vectoriel** E une application linéaire définie sur E et à valeurs dans $\mathbb{K}$.

Propriétés $\bullet$On peut se contenter pour vérifier la linéarité d'une application de s'assurer que $f(\lambda.x + \mu.y) = \lambda.f(x) + \mu.f(y)$

•L'image de 0 par une application linéaire est 0.

•L'identité est un automorphisme, ainsi que toute homothétie de rapport $\lambda \neq 0$ (son application réciproque étant l'homothétie de rapport λ^{-1}).

•L'image d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel. On note $Im\ f$ l'image de f, c'est un sous-espace vectoriel.

•La composée de deux applications linéaires est une application linéaire.

•L'application qui à f associe $g \circ f$ (resp. $f \circ g$) est un morphisme d'espaces vectoriels.

•$Ker\ f \subset Ker\ (g \circ f)$

•$Im(g \circ f) \subset Img$

•$g \circ f = 0$ si et seulement si $Im\ f \subset Ker\ g$

•L'application qui à un polynôme associe son polynôme dérivé est un endomorphisme surjectif, dont le noyau est l'ensemble des polynômes constants.

Définition 0.7 **n-ième itérée de f**
On appelle **n-ième itérée de f** la fonction f^n. f est dit **nilpotent** si il existe n tel que $f^n = 0$. Le plus petit n convenable est appelé **indice de nilpotence** de f. On convient que l'indice de nilpotence d'une fonction non nilpotente est $+\infty$.

si f est un endomorphisme nilpotent, alors $f - Id$ et $f + Id$ sont des automorphismes.

Par exemple, si $f^n = 0$, alors $(f - Id)^{-1} = -(Id + f + f^2 + ... + f^{n-1})$: pour le prouver, il suffit de composer cette somme par $f - Id$, de manière d'ailleurs analogique à la démonstration pour $x \neq 1$, $x \in \mathbb{R}$ de $\sum_{i=0}^{n-1} x^i = (1 - x^n)/(1 - x)$.

Définition 0.8 **Définitions dans le cas d'un espace vectoriel produit**
On appelle **k-ième projection canonique** d'un espace vectoriel produit $E_1 \times ... \times E_n$ l'application qui à x dans le produit associe sa composante dans E_k.

On appelle **k-ième injection canonique** d'un espace vectoriel produit $E_1 \times ... \times E_n$ l'application qui à x dans E_k associe $(0, ..., 0, x, 0, ...0)$.

On notera que si f_i est linéaire de E_i dans F_i, alors $(x_1, ..., x_n) \to (f_1(x_1), ..., f_n(x_n))$ est une application linéaire ; son noyau est le produit des noyaux.

1.3 Somme de sous-espaces vectoriels

On verra ici la somme et la somme directe au sens algébrique. Ces notions seront complétées plus tard par des notions d'orthogonalité dans les espaces munis de formes quadratiques, typiquement dans les espaces de Hilbert ; en particulier, cela permettra l'analyse de Fourier, où l'on étudie des fonctions par le biais de leurs écritures dans une base ayant de bonnes propriétés.

1.3.1 Généralités

Définition 0.9 **Somme de sous-espaces vectoriels**
On se donne $F_1, ..., F_n$ des sous-espaces vectoriels de l'espace vectoriel E. L'application ϕ qui à $(x_1, ..., x_n)$ associe $(x_1 + x_2 + ... + x_n)$ est une application linéaire sur l'espace produit $F_1 \times ... \times F_n$. L'image de ϕ est appelée **somme** des sous-espaces vectoriels $F_1, ..., F_n$, et est notée $F_1 + ... + F_n$ ou $\sum_{i \in [1,n]} F_i$.

Propriétés •$\sum_{i \in [1,n]} F_i$ est un sous-espace vectoriel de E.
•La somme des F_i contient tous les F_i.
•L'espace vectoriel engendré par l'union des F_i est leur somme.

Définition 0.10 **Somme directe de sous-espaces vectoriels**
On dit que la somme de $F_1, ..., F_n$ est une **somme directe** lorsque la fonction ϕ est injective. Au lieu de noter $\sum F_i$ on peut alors noter $\bigoplus F_i$

Propriétés •Cela revient à dire que tout vecteur de l'espace vectoriel somme $(\sum_{i \in [1,n]} F_i)$ s'écrit comme somme d'un élément de F_1, d'un élément de F_2, ..., d'un élément de F_n, cette décomposition étant unique.
•La somme est directe si et seulement si pour tout j $F_j \cap \sum_{i \neq j} F_i = \{0\}$.
•On peut encore simplifier ce critère ; la somme est directe si et seulement si pour tout j $F_j \cap \sum_{i=1}^{j-1} F_i = \{0\}$.

1.3.2 Espaces supplémentaires

Définition 0.11 **supplémentaires**
Deux sous-espaces vectoriels d'un espace E sont dits **supplémentaires** lorsque leur somme est directe et égale à E.

⊡ **Généralités** Propriétés •Deux espaces sont supplémentaires si et seulement si leur intersection est $\{0\}$ et si leur somme est E.
•Deux espaces sont supplémentaires si l'application ϕ (voir la définition d'une somme de sous-espaces vectoriels) est bijective.
•Deux sous-espaces vectoriels qui sont supplémentaires d'un même sous-espace vectoriel sont isomorphes [1].

Exemple 0.2 Dans $\mathbb{K}^n$, l'espace vectoriel engendré par un vecteur $(a_1, ..., a_n)$ est supplémentaire de l'espace vectoriel des $(x_1, ..., x_n)$ tels que $a_1.x_1 + ... + a_n.x_n = 0$.

1. Les définitions nécessaires à l'écriture élégante de la preuve de ce point arrivent un tout petit peu plus bas ; la preuve se fait en considérant la projection sur le supplémentaire en question par rapport à l'un des deux sous-espaces ; un espace supplémentaire du noyau est isomorphe à l'image, d'où le résultat.

THÉORÈME 0.6
Deux applications linéaires ayant les mêmes restrictions à deux espaces vectoriels supplémentaires sont égales.

Étant données deux applications linéaires f_1 et f_2 définies respectivement de F_1 dans G et de F_2 dans G, avec F_1 et F_2 deux supplémentaires de E, il existe une et une seule application linéaire dont les restrictions à F_1 et F_2 sont f_1 et f_2.

PROPOSITION 0.7
Étant donné un sous-espace vectoriel H, f une application linéaire, alors $Ker\ f_{|H} = Ker\ f \cap H$

THÉORÈME 0.8 **Théorème noyau-image**
$Im\ f$ est isomorphe à tout supplémentaire de $Ker\ f$.

Applications aux applications linéaires. Projections, symétries.

Démonstration *Il suffit de considérer la restriction de f à un supplémentaire de $Ker\ f$, et de montrer l'injectivité et la surjectivité.*

Il est à noter qu'en dimension infinie, l'existence d'un supplémentaire à tout sous-espace vectoriel n'est pas évidente (c'est une conséquence de l'axiome du choix).

DÉFINITION 0.12 **Projection**
On a vu que dans le cas où F et G étaient des espaces supplémentaires de E, pour tout x il existait un unique $(f, g) \in F \times G$ tel que $x = f + g$. On appelle **projection sur F parallèlement à G** l'application qui à x associe f.

DÉFINITION 0.13 **idempotent**
On dit qu'un endomorphisme f est **idempotent** lorsque $f \circ f = f$. Un endomorphisme idempotent est aussi appelé **projecteur**.

Propriétés Étant donné p projecteur, on a :
•$E = Ker\ p \oplus Im\ p$
•$Inv\ p = Im\ p$
On note qu'une projection est un projecteur. La réciproque suit :

PROPOSITION 0.9

Un projecteur p est en fait la projection sur $Im\ p$ parallèlement à $Ker\ p$.

PROPOSITION 0.10
$Id - f$ est un projecteur si et seulement si f est un projecteur.

Démonstration *Il suffit de développer $(Id - f)^2$.*

DÉFINITION 0.14 **Projecteurs associés**
Deux projecteurs sont dits **projecteurs associés** lorsque leur somme est l'identité.

PROPOSITION 0.11
Si f et g sont deux projecteurs associés, alors $Im\ f = Ker\ g$ et $Im\ g = Ker\ f$.

DÉFINITION 0.15 **Symétrie**
A et B étant supplémentaires, on appelle **symétrie** par rapport à A parallèlement à B l'endomorphisme s tel que $s_{|A} = Id$ et $s_{|B} = -Id$.

On a vu plus haut qu'un endomorphisme pouvait être défini par ses restrictions sur deux espaces supplémentaires.

DÉFINITION 0.16 **Involution**
Un endomorphisme f est une **involution** lorsque $f \circ f = Id$.

DÉFINITION 0.17 **associés**
Une symétrie s et un projecteur p sont dits **associés** lorsque $s = 2.p - Id$.

C'est le cas lorsque s et p se font par rapport au même espace et parallèlement au même espace.

Propriétes •Une symétrie est une involution.

•Étant donnée s symétrie, $E = Inv\ s \oplus Opp\ s$.

•s est la symétrie par rapport à $Inv\ s$ et parallèlement à $Opp\ s$.

•Toute symétrie est associée à un unique projecteur, et tout projecteur est associée à une unique symétrie.

DÉFINITION 0.18 **Hyperplan**
On appelle **hyperplan** de E tout sous-espace vectoriel de E possèdant comme supplémentaire une droite vectorielle.

⊡ **Dilatations, transvections** Notons que le supplémentaire n'est pas unique. En fait, on peut aussi définir de manière équivalente la notion d'hyperplan par le fait qu'un sous-espace vectoriel H est un hyperplan si et seulement si *toute* droite vectorielle passant par un vecteur appartenant à son complémentaire est un supplémentaire de H.

DÉFINITION 0.19 **dilatation d'hyperplan H, de direction D et de rapport λ**
Soit H un hyperplan d'un espace vectoriel E, et D une droite supplémentaire de H. On appelle **dilatation d'hyperplan H, de direction D et de rapport λ** l'application linéaire dont la

restriction à H est l'identité et dont la restriction à D est l'homothétie de rapport λ.

Soit H un hyperplan d'un espace vectoriel E, et h une forme linéaire de noyau H. On appelle **transvection d'hyperplan** H une application de E dans E de la forme

$$x \mapsto x + h(x).u$$

pour un certain u dans H.

⚠ *Attention 0.3* Ne pas se méprendre sur la notion de transvection d'hyperplan H ; il existe plusieurs transvections différentes d'hyperplan H.

Rappelons que la proposition **??** signale que $GL(E)$ est engendré par les transvections et les dilatations, et que $SL(E)$ est engendré par les transvections.

1.4 Espace vectoriel quotient

1.4.1 Généralités

DÉFINITION 0.20 **espace vectoriel quotient**
Soit F un sous-espace vectoriel de E. Alors la relation définie par

$$x\mathcal{R}y \iff x - y \in F$$

est une relation d'équivalence compatible avec l'addition et le produit externe. L'ensemble quotient est un espace vectoriel pour les lois induites ; il est appelé **espace vectoriel quotient** et est noté $E_{/F}$.

Propriétés •La surjection canonique qui à x associe $\overline{x}$ est linéaire. Son noyau est F.
•Si F et G sont supplémentaires, alors G est isomorphe à $E_{/F}$.

1.4.2 Application aux applications linéaires

THÉORÈME 0.12
Étant donné une application linéaire f de E dans F, alors f s'écrit de manière unique $f = i \circ f' \circ s$, avec s la surjection canonique de E dans $E_{/Ker\ f}$, i l'injection canonique de $Im\ f$ dans F, et f' un isomorphisme de $E_{/Ker\ f}$ dans Imf.

Démonstration *Facile !*

1.5 Translations – espaces affines

Pour plus d'informations on pourra consulter la partie **??**.

DÉFINITION 0.21 **Translations**

Étant donné $a \in X$, on appelle **translation** de vecteur a l'application d'un espace vectoriel X dans lui-même qui à x associe $x + a$. On note $\mathcal{T}(E)$ l'ensemble des translations de E.

On appelle **sous-espace affine** de E l'image d'un sous-espace vectoriel de E par une translation de E.

On appelle **direction d'un sous-espace affine** A l'ensemble des $x - y$ pour x et y dans A.

On dit qu'un sous-espace affine A est **parallèle** à un sous-espace affine B si et seulement si la direction de A est incluse dans la direction de B.

On dit que deux sous-espaces affines sont **parallèles** s'ils ont même direction.

On dit que deux sous-espaces affines sont **strictement parallèles** s'ils ont même direction et sont distincts.

Propriétés •$(\mathcal{T}(E), \circ)$ est un groupe commutatif. $(\mathcal{T}(E), \circ) \simeq (E, +)$.

•Un sous-espace affine de E contient 0 si et seulement si c'est un sous-espace vectoriel (en le munissant des lois induites).

•Si un espace affine A est de la forme $a + F$ avec $a \in A$ et F sous-espace vectoriel, alors il est aussi de la forme $x + F$ pour tout x de A.

•Le sous-espace affine A est égal à $a + F$, avec a quelconque dans A, et F la direction de A.

•Étant donnés A et B deux espaces affines, $a \in A$ et $b \in B$, de direction respectives F et G, alors $A \cap B \neq \emptyset \iff b - a \in F + G$

•Étant donnés deux espaces affines d'intersection non vide, l'intersection est un espace affine de direction l'intersection de leurs directions.

•Si A est parallèle à B alors $A \cap B = \emptyset$ ou $A \subset B$.

1.6 Familles libres. Familles génératrices. Bases

1.6.1 Généralités

DÉFINITION 0.22 **combinaison linéaire**

Un élément x de l'espace vectoriel E est dit **combinaison linéaire** d'une famille $(x_i)_{i \in I}$ d'éléments de E si il existe un entier n, un n-uplet $(\lambda_1, ..., \lambda_n)$ d'éléments de $\mathbb{K}$, et un n-uplet $(i_1, ..., i_n)$ d'éléments de I tels que $x = \sum_{1 \leq j \leq n} \lambda_j x_{i_j}$.

Un élément x est dit **combinaison linéaire** d'un sous-ensemble A de E si x est combinaison linéaire d'une famille d'éléments de A.

Une famille d'éléments d'un sous-espace vectoriel F est dite **famille génératrice** de F si l'espace vectoriel engendré par cette famille est F.

Une famille $(x_1, \ldots, x_n)$ d'éléments de E est dite **famille libre** si

$$\forall \lambda \in \mathbb{K}^n, \sum_{j \in [[1,n]]} \lambda_j x_j = 0 \Rightarrow (\forall j, \lambda_j = 0).$$

Une famille infinie est libre si toute sous-famille finie est libre.

Une famille est dite **famille liée** quand elle n'est pas libre.

Propriétés •L'image d'une famille génératrice de F par une application linéaire f est une famille génératrice de $f(F)$.

•L'ensemble des combinaisons linéaires des éléments de A est l'espace vectoriel engendré par A.

•Toute famille contenant le vecteur nul est liée.

•Les éléments d'une famille libre sont deux à deux distincts.

•Une famille est liée si et seulement si un de ses éléments est combinaison linéaire des autres.

•La famille (v) est libre si et seulement si $v \neq 0$.

•La famille (v, w) est libre si et seulement si $v \neq 0$ et $\nexists \lambda \in \mathbb{K} : w = \lambda.v$.

•Une famille est liée si et seulement si un de ses éléments appartient à l'espace vectoriel engendré par les autres.

•Toute sous-famille d'une famille libre est libre.

•Toute sur-famille d'une famille liée est liée.

THÉORÈME 0.13
On a les deux propriétés suivantes des familles libres ou liées :

•L'image par une application linéaire d'une famille liée est une famille liée.

•L'image d'une famille libre par une application injective est une famille libre.

Le théorème suivant, facile à démontrer, est fondamental pour la suite.

THÉORÈME 0.14

Étant donnée une famille libre et un élément appartenant à l'espace vectoriel engendré par cette famille, alors cet élément s'écrit de manière unique comme combinaison linéaire d'éléments de la famille.

DÉFINITION 0.23 **Base**
Une base d'un espace vectoriel E est une famille libre et génératrice.

PROPOSITION 0.15
Une famille est une base si et seulement si c'est une famille libre maximale (au sens de l'inclusion).

Démonstration *Si elle n'est pas maximale, on peut ajouter un élément x tout en la laissant libre ; on en déduit que ce n'est pas une famille génératrice car x n'est pas combinaison linéaire de cette famille. Si elle n'est pas génératrice, alors on peut ajouter un élément x n'appartenant pas à l'espace engendré ; cette nouvelle famille, contenant x, est aussi libre, donc la précédente n'était pas maximale.*

PROPOSITION 0.16
Une famille est une base si et seulement si c'est une famille génératrice minimale.

Démonstration *Même principe que la preuve ci-dessus.*

Définition 0.24 **coordonnées**
Étant donnée une base (e_i) d'un espace vectoriel F et un élément x de F, il existe une unique famille (λ_i) de support fini telle que $x = \sum \lambda_i e_i$. Les λ_i sont appelés **coordonnées** du vecteur x.

Notons que les coordonnées ne sont pas (du tout) invariantes par changement de base.

1.6.2 Conséquences pour les applications linéaires

Théorème 0.17
Étant donnée $(f_i)_{i \in I}$ une base de F et $(g_i)_{i \in I}$ une famille de G, il existe une unique application linéaire φ de F dans G telle que $\varphi(f_i) = g_i$ pour tout $i \in I$.

- La famille (g_i) est libre si et seulement si φ est injective.
- La famille (g_i) est génératrice si et seulement si φ est surjective.
- La famille (g_i) est une base de G si et seulement si φ est bijective.

1.6.3 Applications aux sous-espaces vectoriels

Proposition 0.18
L'ensemble $\mathcal{L}(E)$ des endomorphismes de E muni de l'addition, de la composition, et de la multiplication par un scalaire, est une $\mathbb{K}$-algèbre (cf. la définition 0.28).

Attention, cette algèbre n'est pas commutative (en général, i.e. dès que E n'est pas isomorphe à $\mathbb{K}$).

1.7 Dualité

Il existe de nombreuses formes de dualité (voir l'index), très différentes. La dualité dont il est question ici est quelquefois qualifiée de dualité *algébrique*. De même, on parle de dual algébrique ou de bidual algébrique pour éviter la confusion avec le dual ou le bidual topologiques. On sera ici très succinct, mais il convient de lire aussi la section 1.10.2 pour le très important cas particulier de la dimension finie, en particulier au niveau de l'isomorphisme $E \simeq E^{**}$.

1.7.1 Définitions et premières propriétés. Le dual et le bidual

Définition 0.25 **forme linéaire**
On appelle **forme linéaire** sur un $\mathbb{K}$-espace vectoriel E une application linéaire de E dans $\mathbb{K}$.

On appelle **dual** d'un $\mathbb{K}$-espace vectoriel E l'ensemble E^* des formes linéaires sur cet espace vectoriel.

On appelle **bidual** d'un $\mathbb{K}$-espace vectoriel le dual de son dual. On le note E^{**}.

On appelle **application linéaire canonique** de E dans E^{**} l'injection qui à x associe ϕ_x : $u \mapsto u(x)$ (on vérifiera facilement que c'est une application linéaire).

Propriétés •Une forme linéaire non nulle est surjective.

•Un sous-espace vectoriel est un hyperplan si et seulement si c'est le noyau d'une forme linéaire non nulle.

•Deux formes linéaires non nulles sont liées (dans le dual) si et seulement si elles ont même noyau.

•Étant données u et v des formes linéaires sur E, il existe $\lambda \in \mathbb{K}$ tel que $u = \lambda v$ si et seulement si $Ker\ u \subset Ker\ v$.

1.7.2 Orthogonalité

L'orthogonalité permet d'écrire un sous-espace vectoriel comme l'orthogonal d'un sous-espace vectoriel dans le dual. En clair, on peut donc écrire un sous-espace vectoriel comme l'intersection d'hyperplans. On verra aussi avec les topologies sur le dual l'intérêt de la dualité ; la convergence faible-* par exemple, très utile, n'existerait pas sans le dual.

Définition 0.26 **orthogonaux**
Étant donné x dans E et u dans E^*, on dit que x et u sont **orthogonaux** si et seulement si $u(x) = 0$

Étant donnée une partie non vide A de E, on appelle **orthogonal** de A dans E^* et on note $A^\perp$ l'ensemble des u orthogonaux à tous les éléments de A.

Étant donnée une partie non vide A de E^* on appelle **orthogonal** de A dans E et on note A^o l'ensemble des x orthogonaux à tous les éléments de A.

Propriétés On utilise le terme orthogonal sans plus de précision lorsque la propriété vaut à la fois dans le cas de l'orthogonal d'une partie de E dans le dual et dans le cas de l'orthogonal d'une partie du dual E^* dans E.

Il n'y a pas ici de bidualité ; l'orthogonal d'une partie du dual E^* est à considérer dans E.

•L'orthogonal d'une partie est l'orthogonal de l'espace engendré par cette partie.

•L'orthogonal d'une partie est un sous-espace vectoriel.

•Toute partie est incluse dans l'orthogonal de son orthogonale.

•$A \subset B$ alors l'orthogonal de B est inclus dans l'orthogonal de A.

•L'intersection des orthogonaux est l'orthogonal de l'union.

⚠ *Attention 0.4* Attention, ne pas confondre l'orthogonal de l'orthogonal de A qu'est ${A^\perp}^o \subset E$ et l'orthogonal de l'orthogonal de A qu'est ${A^\perp}^\perp \subset E^{**}$. Voir à ce sujet, dans le cas de la dimension finie, la section 1.10.2.

On verra un peu plus d'orthogonalité dans le cas de la dimension finie avec la section 1.10.2.

1.7.3 Transposition

Définition 0.27 **transposée de f**
Étant donnée f une application linéaire de E dans F on appelle **transposée de f** l'application de F^* dans E^* qui à v associe $v \circ f$. C'est une application linéaire, et on la note ${}^t f$.

•L'application qui à f associe ${}^t f$ est une application linéaire.

•$^t(g \circ f) =^t f \circ^t g$
•$^tId_E = Id_{E^*}$
•Si f est un isomorphisme, ^tf est un isomorphisme, et $^t(f^{-1}) = (^tf)^{-1}$
•$Ker\ ^tf = (Im\ f)^\perp$ et $Im\ ^tf = (Ker\ f)^\perp$
•f est surjective si et seulement si ^tf est injective.
•f est injective si et seulement si ^tf est surjective.

On verra des propriétés supplémentaires en dimension finie, avec le théorème 0.34.

1.8 $\mathbb{K}$-algèbres

Définition 0.28 **$\mathbb{K}$-algèbre**
$(A, +, \times, .)$ est une **$\mathbb{K}$-algèbre** si
•$(A, +, .)$ est un $\mathbb{K}$-espace vectoriel
•$(A, +, \times)$ est un anneau
•$(\lambda.a) \times b = a \times (\lambda.b) = \lambda.(a \times b)$
La $\mathbb{K}$-algèbre est en outre **commutative** lorsque $\times$ est commutative.

L'ensembles des suites à valeurs dans $\mathbb{K}$ est un $\mathbb{K}$-algèbre commutative. L'espace vectoriel des applications de A dans $\mathbb{K}$ est une $\mathbb{K}$ algèbre commutative.

Définition 0.29 **Morphisme de $\mathbb{K}$-algèbres**
Un **morphisme d'algèbre** est une application qui est à la fois un morphisme d'anneaux sur les anneaux sous-jacents et un morphisme d'espaces vectoriels sur les espaces vectoriels sous-jacents.

1.9 Zoologie d'algèbre linéaire

Tout d'abord signalons que :

- L'union de deux sous-espaces vectoriels (à ne pas confondre avec la somme d'espaces vectoriels) est un sous-espace vectoriel si et seulement si l'un des deux est inclus dans l'autre.
- Étant donnés F et G des espaces vectoriels et f et g des vecteurs alors $f + F \subset g + G$ si et seulement si $F \subset G \wedge f - g \in G$.

Quelques points quant à des endomorphismes particuliers (projecteurs, homothéties :

- Étant donnés deux projecteurs, la somme est un projecteur si et seulement si les composées de ces projecteurs (dans les deux sens, la composition n'étant pas commutative) sont nulles. Dans ce cas le noyau de la somme est l'intersection des noyaux, et l'image de la somme est la somme directe des images.
- Étant donné $k \neq 0$, on appelle **homothétie vectorielle de rapport** k l'application qui à x associe $k.x$; un endomorphisme f est une homothétie vectorielle si et seulement si pour tout x la famille $(x, f(x))$ est liée.

Prouvons le second point.
Démonstration *Le sens $\Rightarrow$ est immédiat, on prouve seulement la réciproque ($\Leftarrow$).*
On procède par étapes :

— *tout d'abord montrer que pour tout x il existe $k_x \in \mathbb{K}$ tel que $f(x) = k_x.x$.*
— *ensuite montrer que k_x est constant sur toute droite vectorielle.*
— *développer $f(x+y)$ avec (x, y) une famille libre. On en déduit $k_x = k_y$, car les deux égalent k_{x+y}.*

Examinons maintenant une famille de fonctions de $\mathbb{R}$ dans $\mathbb{R}$ pour avoir un exemple de famille libre infinie et de cardinal non dénombrable : la famille des fonctions f_a de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe 1 si $x > a$ et 0 si $x \leq a$ est libre.

Démonstration *On suppose qu'une telle fonction f_a est combinaison linéaire d'un nombre fini d'autres fonctions de cette forme, et on constate que la fonction doit être continue en a, c'est-à-dire là où justement elle ne l'est pas.*

Pour s'habituer aux manipulations sur les noyaux et les images, on peut montrer : (avec encore f un endomorphisme)

$$Ker\ f = Ker\ f^2 \iff Im\ f \cap Ker\ f = \{0\}$$
$$Im\ f = Im\ f^2 \iff Im\ f + Ker\ f = E$$
$$n \mapsto Ker\ f^n \text{ est croissante pour } \subset$$
$$n \mapsto Im\ f^n \text{ est décroissante pour } \subset$$
$$\exists n / Ker\ f^n = Ker\ f^{n+1} \to \forall k > n\ Ker\ f^k = Ker\ f^n$$
$$\exists n / Im\ f^n = Im\ f^{n+1} \to \forall k > n\ Im\ f^k = Im\ f^n$$

1.10 Algèbre linéaire en dimension finie

On présentera quelques généralités sur l'algèbre linéaire en dimension finie en section 1.10.1 et sur la dualité en dimension finie en section 1.10.2.

On consacrera ensuite plusieurs sections aux matrices : le calcul matriciel en section 1.10.3, les opérations sur les lignes et les colonnes en section 1.10.3 ; enfin, deux dernières sections seront consacrées aux matrices avec des exercices et de la zoologie (sections 1.10.3 et 1.10.4 respectivement).

On passera ensuite à des sections plus originales ; on reviendra un peu sur la dualité en dimension finie avec un peu de zoologie (section 1.10.5).

Les espaces vectoriels ont bien sûr une interprétation physique immédiate, en particulier pour la mécanique. Ils sont aussi à la base de structures algébriques comme les espaces de Hilbert, qui auront beaucoup d'applications en analyse fonctionnelle (par exemple pour l'analyse de Fourier). Ils servent aussi à caractériser des solutions d'équations différentielles (l'ensemble des solutions étant en général un espace affine, de dimension évaluable a priori selon les caractéristiques de l'équation). L'analyse matricielle sera à la base de l'optimisation via la matrice hessienne ou les matrices représentant les contraintes en programmation linéaire.

1.10.1 Généralités

DÉFINITION 0.30 **de dimension finie**
Un espace vectoriel est dit **de dimension finie** lorsqu'il admet une base finie. Dans le cas contraire il est dit **de dimension infinie**.

Dans un espace fini le cardinal d'une base est appelé **dimension** de l'espace (pour la cohérence de la définition il faudra voir que toutes les bases ont alors même cardinal). Un sous-espace vectoriel d'un espace vectoriel E est dit de **codimension finie** si la dimension de l'espace quotient est finie. On appelle alors **codimension** de cet espace la dimension de l'espace quotient. Sinon il est dit de **codimension infinie**. La notion de base a été définie dans la partie 1.6.

Dans la suite, E désigne un $\mathbb{K}$-espace vectoriel de dimension finie.

> THÉORÈME 0.19 **Théorème de la base incomplète**
>
> Si I est une famille libre et K une famille génératrice finie, avec $I \subset K$, alors il existe J avec $I \subset J \subset K$ tel que J soit une base.

Application 0.5 On verra des applications avec (entre autres) les théorèmes 0.24 (existence de supplémentaire) et 0.36 (à propos d'orthogonalité dans le dual), la proposition 0.48 (rang et matrices extraites).

Démonstration *On considère J libre maximale au sens du cardinal vérifiant $I \subset J \subset K$. Si toute famille plus grande incluse dans K est liée, alors tout élément de K est combinaison linéaire d'éléments de J ; tout élément de E s'écrit comme combinaison d'éléments de K, qui eux-mêmes s'écrivent comme combinaisons linéaires d'éléments de J. Donc la famille J est génératrice.*

> LEMME 0.20 **Lemme de Steinitz**
>
> Si E est non réduit à $\{0\}$, E admettant une famille génératrice I de cardinal n, toute famille de $n+1$ vecteurs (ou plus) est liée.

Démonstration *•Le cas $n = 1$ est immédiat. On procède alors par récurrence.*

•Supposons la propriété vraie pour 1, 2, ...,$n-1$; supposons E admettant une famille génératrice $b_1, ..., b_n$ de cardinal n. Notons H le sous-espace vectoriel de E engendré par $b_1, ..., b_{n-1}$.

•Donnons-nous une famille $e_1, ..., e_{n+1}$ de E. Notre but est de montrer que cette famille est liée.

•Soit, pour $i \in [\![1, n+1]\!]$, $h_i \in H$ et $\lambda_i \in \mathbb{K}$ tel que $e_i = h_i + \lambda_i b_n$.

•Si tous les λ_i sont nuls, alors l'hypothèse de récurrence appliquée à H permet de conclure (tous les e_i sont alors dans H, et donc ils sont liés).

•Supposons alors (sans perte de généralité) $\lambda_1 \neq 0$; dans ce cas, $b_n = \frac{1}{\lambda_1}(e_1 - h_1)$.

•On peut alors écrire :

$$e_i - h_i = \quad \frac{\lambda_i}{\lambda_1}(e_1 - h_1) e'_i = e_i - \frac{\lambda_i}{\lambda_1} e_1 = \underbrace{h_i}_{\in H} - \underbrace{\frac{\lambda_i}{\lambda_1} h_1}_{\in H}{}_{\in H}$$

Les e'_i pour $i \in [[2, n+1]]$ appartiennent donc à H ; ils sont n, et appartiennent à un espace généré par $n-1$ points, et donc par hypothèse de récurrence ils sont liés. On a donc une combinaison linéaire nulle des e'_i pour $i \in [[2, n+1]]$, à coefficients non tous nuls. On a donc une combinaison linéaire nulle des e_i pour $i \in [[1, n+1]]$, dont au moins un coefficient est non-nul ; ceci conclut la preuve.

> THÉORÈME 0.21
> Dans un espace de dimension finie, toutes les bases ont le même cardinal.

Démonstration *Conséquence immédiate du lemme 0.20.*

> THÉORÈME 0.22
> Deux espaces vectoriels de dimension finie sur le même corps sont isomorphes si et seulement si ils ont même dimension.

Démonstration *S'ils sont isomorphes, l'image d'une base de l'un par un isomorphisme est base de l'autre, donc ils ont même dimension. S'ils ont même dimension, alors avec* $(e_1, ..., e_n)$ *une base de l'un, et* $(f_1, ..., f_n)$ *une base de l'autre, l'application* $\sum \lambda_i e_i \mapsto \sum \lambda_i f_i$ *pour* $(\lambda_1, ..., \lambda_n) \in \mathbb{K}^n$ *est un isomorphisme, comme on le vérifie facilement (la fonction est bien définie car* $(e_1, ..., e_n)$ *est une base, linéarité évidente, injectivité immédiate car* $(f_1, ..., f_n)$ *est libre, surjectivité immédiate car* $(f_1, ..., f_n)$ *est génératrice).*

Théorème 0.23

F, sous-espace vectoriel de E, est égal à E si et seulement si $dim\ E = dim\ F$.

Démonstration *Il est clair que si F et E sont égaux, alors ils ont même dimension. Réciproquement, s'ils ont même dimension, alors une base de F est une famille libre de même cardinal qu'une base de E, donc c'est une base de E (voir le lemme de Steinitz ci-dessus).*

Théorème 0.24
Soit E espace vectoriel non nul. Tout sous-espace vectoriel de E admet un supplémentaire, et si $E = F \oplus G$, alors $dim\ E = dim\ F + dim\ G$, et une base de E s'obtient en réunissant une base de F et une base de G.

Démonstration *Conséquence du théorème de la base incomplète.*
Maintenant quelques théorèmes faciles, sans démonstration, mais fort pratiques néanmoins :

Théorème 0.25
Étant donnés F et G des sous-espaces vectoriels de E, on a $dim\ (F + G) + dim\ (F \cap G) = dim\ F + dim\ G$.

On consultera avec profit la section **??** quant à la définition d'une topologie sur les sous-espaces vectoriels d'un espace vectoriel .

Théorème 0.26 **Caractérisations de supplémentaires**
F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si la somme de leurs dimensions est la dimension de E.

F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si leur somme est égale à E.

F et G sont supplémentaires dans E si et seulement si leur somme est égale à E et si la somme de leurs dimensions est égale à la dimension de E.

Théorème 0.27

Étant donné F sous-espace vectoriel de E, la dimension de E/F est égale à la dimension de E moins la dimension de F. D'ailleurs, E/F est isomorphe à tout supplémentaire de F dans E.

Théorème 0.28

Si les E_i sont de dimension finie, alors $dim\ \prod_{i=1}^n E_i = \sum_{i=1}^n dim\ E_i$.

Théorème 0.29

Si E et F sont de dimension finie, alors $\mathcal{L}(E,F)$ est de dimension finie et $dim\ \mathcal{L}(E,F) = dim\ E.dim\ F$.

Démonstration *On considère une base de E et une base de F, notées respectivement (e_i) et (f_i) ; alors les $\varphi_{i,j}$ définies par $\varphi_{i,j}(e_i) = f_j$ et $\varphi_{i,j}(e_l) = 0$ pour $l \neq i$ forment une base de $\mathcal{L}(E,F)$.*

Définition 0.31 **rang**

On appelle **rang** d'une famille finie de vecteurs la dimension de l'espace vectoriel que cette famille engendre.

On appelle **rang** d'une application linéaire la dimension de son image lorsque celle-ci est finie. C'est en fait le rang de l'image d'une base lorsque le domaine est un espace vectoriel de dimension finie

Théorème 0.30 **Théorème du rang**
Si $f \in \mathcal{L}(E,F)$ et E de dimension finie, alors $Im\ f$ et $Ker\ f$ sont de dimension finie, et $dim\ E = dim\ Im\ f + dim\ Ker\ f$.

Démonstration *On considère un supplémentaire du noyau, et on montre qu'il est isomorphe à l'image de f (voir théorème 0.8).*

Corollaire 0.31

Soit f une application d'un espace vectoriel E vers un espace vectoriel F, avec E et F de même dimension finie, alors f est un isomorphisme si et seulement si l'une des propriétés suivantes est vérifiée :
- f a un noyau réduit au singleton $\{0\}$.
- f est injective.
- f est surjective.
- f est bijective.
- Le rang de f est égal à la dimension de E.

Proposition 0.32
Quelques propriétés du rang : •Le rang d'une famille finie de vecteurs est invariant par permutations

•Multiplier un vecteur d'une famille par un scalaire non nul ne change pas le rang de la famille
•On ne change pas le rang d'une famille de vecteurs en additionnant à un vecteur une combinaison linéaire des autres vecteurs
•Le rang d'une famille de vecteurs est le même si l'on supprime un vecteur nul

1.10.2 Dualité en dimension finie

La dualité existe dans plusieurs cadres : géométrie, algèbre, topologie. On s'intéresse ici à la dualité au sens algébrique. Le lecteur est invité à différencier les différentes dualités par un rapide tour d'horizon par l'entrée « dualité » de l'index. De même, il convient de ne pas confondre l'orthogonal au sens géométrique, et l'orthogonal dans le dual.

Dualité simple E est un espace vectoriel normé de dimension finie n.

DÉFINITION 0.32 **formes coordonnées associées**

Étant donnée une base $(e_i)_{i\in[\![1,n]\!]}$ de E, on appelle **formes coordonnées associées** les n formes linéaires e_j^* définies par $e_j^*(e_i)=\delta_{i,j}$.

On note que $x=\sum_{j=1}^n e_j^*(x).e_j$.

THÉORÈME 0.33
Quelques propriétés élémentaires de la dualité *en dimension finie* :
•$dim\ E^*=dim\ E$.
•La famille des (e_j^*) est une base de E^* ; on l'appelle **la base duale** de la base (e_i).
•Tout f appartenant à E^* s'écrit $f=\sum_i f(e_i).e_i^*$.
•Pour toute base (f_i) de E^*, il existe une base de E dont la base duale est la base des (f_i).

Application 0.6 Le dernier point servira par exemple pour relier la dimension d'un sous-espace vectoriel et celle de son orthogonal dans le dual (théorème 0.37).

Démonstration *La dimension de E^* est égale à la dimension de E, car dim $\mathcal{L}(E,\mathbb{K})=dim\ E.dim\ \mathbb{K}=dim\ E$ (cf théorème 0.29) ; il suffit donc de montrer que la famille des (e_i^*) est libre. On se donne une combinaison linéaire nulle, dont un coefficient au moins est non nul, des e_j^* ; en considérant l'image de e_i par cette combinaison linéaire nulle, on voit que le coefficient de e_i^* est nul pour tout i.*

Donnons-nous ensuite une base $f_1,\ldots,f_n$ de E^*. On se donne une application ϕ qui à x associe $(f_1(x),...,f_n(x))$. On montre sans difficultés que ϕ est un isomorphisme de E dans $\mathbb{K}^n$. Donc il existe une base (e_i) telle que $\phi(e_i)=(\delta_{1,i},...,\delta_{i,i},...\delta_{n,i})$; cette base convient.

On donne sans démonstration le théorème aisé suivant (utilisant la transposition, définie en section 1.7.3) :

THÉORÈME 0.34
Soient E et F des $\mathbb{K}$-espaces vectoriels de dimensions finies (non nécessairement égales). Alors

l'application transposition qui à $f \in \mathcal{L}(E,F)$ associe ${}^tf \in \mathcal{L}(F^*,E^*)$ est un isomorphisme. En outre le rang de ${}^tf$ est égal au rang de f et

$$Ker\ {}^tf = (Im\ f)^\perp$$

Bidual E est un espace vectoriel de dimension finie n.

THÉORÈME 0.35 **Bidual**

L'application

$$x \mapsto (\phi \mapsto \phi(x))$$

c'est-à-dire $x \mapsto x^{**}$ avec $x^{**} = \phi \mapsto \phi(x)$, de E dans E^{**} qui à x associe la fonction qui à ϕ dans E^* associe $\phi(x)$ est un isomorphisme ; on l'appelle **isomorphisme canonique** de E dans E^{**}. On peut donc identifier, via cet isomorphisme, E et E^{**}.

Démonstration *Les dimensions de E et E^{**} étant égales (on est toujours dans le cadre de la dimension finie) il suffit de voir que cette fonction est un morphisme injectif. Supposons x d'image nulle ; alors quel que soit f, $f(x) = 0$. Si x est non nul, alors on peut l'appeler e_1 et le compléter en une base e_i ; alors on constate que $e_i^*(x) = 0$ pour tout i, ce qui est contradictoire.*

Orthogonalité Ne confondons pas l'orthogonal au sens géométrique (dans un espace muni d'une forme quadratique, cf section **??**, dont les exemples les plus intuitifs sont les espaces de Hilbert, en particulier euclidiens ou hermitiens) et l'orthogonal au sens de la dualité.

E est un espace vectoriel normé de dimension finie n.

THÉORÈME 0.36
Quel que soit F sous-espace vectoriel de E, on a :
- $dim\ E = dim\ F + dim\ F^\perp$
- ${F^\perp}^o = F$

Démonstration *Si $F = \{0\}$ alors le résultat est clair. Sinon, on considère une base $(f_i)_{i\in[\![1,r]\!]}$ de F, et on la complète en une base $(f_i)_{i\in[\![1,n]\!]}$ de E. On considère la base duale (f_i^*) ; il est clair alors que les $(f_i)_{i>r}$ forment une famille libre génératrice de $F^\perp$, d'où la première assertion. On constate bien que ${F^\perp}^o$ est inclus dans F ; la dimension permet de conclure.*

THÉORÈME 0.37
Quel que soit F sous-espace vectoriel de E^*, on a :
- $dim\ E = dim\ F + dim\ F^o$
- ${F^o}^\perp = F$

Démonstration *La méthode est la même que précédemment ; il faut juste se souvenir que toute base du dual est la base duale d'une certaine base (cf théorème 0.33).*

Donnons enfin sans démonstration le théorème aisé suivant :

Théorème 0.38
Avec ϕ l'isomorphisme canonique de E dans E^{**}, pour tout F sous-espace vectoriel de E, on a $F^{\perp^{\perp}} = \phi(F^{\perp^{o}}) = \phi(F)$.

1.10.3 Calcul matriciel

Le calcul matriciel sert partout. Un grand nombre de problèmes industriels se ramènent à des problèmes linéaires ou quadratiques, donc de représentation matricielle. On parle en général de problème « linéaire » lorsqu'il s'agit de chercher des minima de fonctions de la forme $x \mapsto \sum_i \lambda_i x_i$ pour x tel que $Ax \geq c$ avec A une matrice et c un vecteur. On parle de problèmes quadratiques en général pour la recherche de minima de fonctions de la forme $x \mapsto Q(x)$, avec Q une forme quadratique, pour x tel que $Ax \geq c$ (Notons que parfois on parle de problèmes quadratiques avec des contraintes $B(x) \leq c$ avec B une forme quadratique définie positive). Les problèmes non-linéaires sont en général la recherche de minima de $x \mapsto f(x)$ parmi les x tels que $\forall i \in I, g_i(x) \geq 0$, lorsque le problème ne peut se ramener ni au linéaire, ni au quadratique. De manière amusante, le non-linéaire n'est donc pas simplement ce qui n'est pas linéaire ; le quadratique n'est ni linéaire ni non-linéaire. Même les problèmes non-linéaires se ramènent fréquemment soit à une résolution successive d'un grand nombre de problèmes linéaires, soit à une résolution successive d'un grand nombre de problèmes quadratiques. Les problèmes non-linéaires que l'on ne résout pas par retour à des problèmes quadratiques ou linéaires utilisent néanmoins des matrices et en particulier la matrice dite matrice hessienne : avec f une application suffisamment dérivable, la matrice hessienne de f est :

$$M_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Les propriétés de cette matrice sont en particulier liées aux extrêma de f ; voir le chapitre **??**. On appelle **hessien** le déterminant de la matrice hessienne. **Attention :** Ne pas dire que la hessienne est définie positive dès que le hessien est positif ! La hessienne, matrice symétrique par le théorème de Schwartz, est définie positive si pour tout vecteur x non nul, $x'Mx > 0$ avec x' le transposé de x (ou, de manière équivalente, si toutes ses valeurs propres sont positives strictement, ou, de manière encore équivalente, si toute sous-matrice principale a un déterminant positif).

On parlera ici de matrices finies, i.e. avec un nombre fini de lignes et de colonnes. On peut aussi considérer des matrices avec une quantité dénombrables de lignes et de colonnes. Cela sera en particulier utile pour l'analyse du recuit simulé ou des Monte-Carlo par chaînes de Markov. Ces techniques, stochastiques, permettent l'optimisation dans des espaces non-continus (pour le recuit simulé, mais aussi les algorithmes dits évolutionnaires ou les algorithmes dits génétiques) ou le calcul d'intégrales dans des espaces délicats (pour Monte-Carlo par chaînes de Markov). La généralisation à des quantités non dénombrables est usuellement qualifiée de « noyau de transition », i.e. une application $k(x, y)$ (pour x et y dans un même domaine non-dénombrable) telle que $\int k(x, .) = 1$ pour tout x.

Les représentations canoniques de matrices sont capitales pour représenter des problèmes de manière simple. Il est indispensable de bien maitriser ce que signifie que deux matrices sont équivalentes, ou le fait qu'elles sont semblables, et surtout ne pas croire que c'est pareil.

Définition 0.33 **Définitions de base sur les matrices**
On appelle **matrice de type** (n,p) **sur un corps** $\mathbb{K}$ toute application de $[\![1,n]\!] \times [\![1,p]\!]$ (intervalles de $\mathbb{N}$) dans $\mathbb{K}$. On la représente généralement comme suit :

$$\begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,p} \\ m_{2,1} & m_{2,2} & \dots & m_{2,p} \\ \vdots & \vdots & \ddots & \vdots m_{n,1} \\ m_{n,2} & \dots & m_{n,p} & \end{pmatrix}$$

On note $M_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n,p) sur le corps $\mathbb{K}$.

On appelle **matrice ligne** une matrice de type $(1,p)$, et **matrice colonne** une matrice de type $(n,1)$.

On appelle **matrice extraite** d'une matrice de type (n,p) la restriction de cette matrice à $I \times J$, avec $I \subset [\![1,n]\!]$ et $J \subset [\![1,p]\!]$.

On appelle **i-ième vecteur-ligne** de la matrice M de type (n,p) la restriction de cette matrice à $\{i\} \times [\![1,p]\!]$. On peut identifier un vecteur-ligne à un élément de $\mathbb{K}^p$.

On appelle **j-ième vecteur-colonne** de la matrice M de type (n,p) la restriction de cette matrice à $[\![1,n]\!] \times \{j\}$. On peut identifier un vecteur-colonne à un élément de $\mathbb{K}^n$.

On appelle **matrice associée** à un morphisme f de l'espace vectoriel E de dimension p dans l'espace vectoriel F de dimension n et aux bases $B = (e_i)$ et $B' = (f_i)$ de E et F respectivement la matrice M de type (n,p) telle que $M_{i,j} = f_i^*(e_j)$. On la note $Mat_{B,B'}(f)$.

Inversement, on appelle **application linéaire canoniquement associée à la matrice** M le morphisme de $\mathbb{K}^p$ dans $\mathbb{K}^n$ dont la matrice associée est M.

Bases sur les matrices *Application 0.7* L'introduction de ce chapitre fournit une liste d'applications des matrices. On peut citer aussi les applications de différentes formes spécialisées de matrices à des domaines directement opérationnels :

— matrices de Hadamard, par exemple pour la discrépance (cf e.g. [1]) ;

— matrices stochastiques pour l'étude de chaînes de Markov à espaces d'états finis ou dénombrables (voir [2]) ;

— déterminants pour l'analyse de plans d'expérience ;

et on en trouvera d'autres en zoologie 1.10.4.

Commençons par deux propositions aisées :

PROPOSITION 0.39
Une matrice de type n, p admet $C_n^{n'}.C_p^{p'}$ matrices extraites de type (n', p').

PROPOSITION 0.40
E et F étant de dimension respectives p et n sur le corps $\mathbb{K}$, on a

$$\mathcal{L}(E, F) \simeq \mathcal{M}_{n,p}(\mathbb{K})$$

via l'isomorphisme $f \mapsto Mat_{B,B'}(f)$.

Avec $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, et B et B' les bases canoniques, on a alors un isomorphisme canonique entre $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ et $\mathcal{M}_{n,p}(\mathbb{K})$.

PROPOSITION 0.41 **Matrice canonique d'une application linéaire en dimension finie**
Soit f une application linéaire entre E, espace vectoriel de dimension p, et F, espace vectoriel de dimension n ; soit r le rang de f. Alors il existe une base B de E et une base B' de F telles que

$$Mat_{B,B'}(f) = M$$

$$\text{avec} \quad M_{i,j} = 1 \text{ si } i = j \leq r \quad M_{i,j} = 0 \text{ sinon}$$

On appelle cette matrice **matrice canonique de** f.

Démonstration *Considérer une base B_1 d'un supplémentaire du noyau de f ; considérer $B_1' = f(B_1)$; considérer B_2 une base du noyau de f, et compléter B_1' en une base B'. Il reste à considérer $B = B_1 \cup B_2$.*

Remarque 0.1 La matrice d'une forme linéaire est une matrice-ligne.

DÉFINITION 0.34 **produit matriciel**
On appelle **matrice produit** des matrices A et B de types respectifs (n, q) et (q, p) la matrice C de type (n, p) telle que

$$C_{i,j} = \sum_{k=1}^{q} A_{i,k}.B_{k,j}$$

On note $C = A \times B$ ou $C = A.B$.

Enfin des propriétés faciles, mais capitales :

Proposition 0.42
•Le produit matriciel défini ci-dessus est associatif et bilinéaire.

•Avec $A = Mat_{B,B'}(f)$ et $B = Mat_{B',B''}(g)$, $B \times A = Mat_{B,B''}(g \circ f)$

•Avec $x \in E$, X le vecteur des coordonnées de x dans la base B, alors les coordonnées de $y = f(x)$ dans la base B' sont celles de Y, avec $Y = M \times X$, et $M = Mat_{B,B'}(f)$.

Espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$ Considérons E et F, $\mathbb{K}$-espaces-vectoriels de dimensions respectives p et n.

Définition 0.35 **matrices élémentaires de type** (n,p)
On appelle **matrices élémentaires de type** (n,p) les matrices $E_{i,j}$ avec $E_{i,j}(a,b) = \delta_{a,i}.\delta_{b,j}$; c'est-à-dire les matrices de type (n,p) ne comportant qu'un 1 (en position (i,j)) et des 0 partout ailleurs.

Proposition 0.43
$\mathcal{M}_{n,p}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel pour l'addition terme à terme et pour la multiplication terme à terme par un scalaire.

L'application de $\mathcal{L}(E,F)$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ qui à une application f associe $Mat_{B,B'}(f)$ est un isomorphisme.

Les matrices élémentaires forment une base de cet espace vectoriel qui est donc de dimension $n.p$.

On retrouve ainsi le fait que $dim\ \mathcal{L}(E,F) = dim\ E.dim\ F$.

Transposition de matrices On définit ici la transposition de matrices et son lien avec la transposition (plus abstraite) de morphismes d'espaces vectoriels (cf section 1.7.3).

Définition 0.36 **Transposée d'une matrice**
Étant donnée une matrice M on appelle **matrice transposée** de M la matrice N avec $N_{i,j} = M_{j,i}$. Si M est de type (n,p), alors N est de type (p,n).

On note $N = {}^tM$.

Proposition 0.44
•L'application $M \mapsto^t M$ est un isomorphisme entre $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{M}_{p,n}(\mathbb{K})$ en tant que $\mathbb{K}$-espaces vectoriels .

•${}^t(A \times B) = {}^tB \times {}^tA$

•$Mat_{B'^*,B^*}({}^tf) = {}^tMat_{B,B'}(f)$

Définition 0.37 **matrice carrée**
On appelle **matrice carrée** une matrice de type (n, n) pour un certain n. On note $\mathcal{M}_n(\mathbb{K}) = \mathcal{M}_{n,n}(\mathbb{K})$.

On appelle **matrice d'un endormophisme** f associée à une base B (finie) la matrice $Mat_{B,B}(f)$; on la note aussi $Mat_B(f)$.

On appelle **diagonale** d'une matrice carrée M de type (n, n) le vecteur $(M_{1,1}, ..., M_{i,i}, ..., M_{n,n})$.

On appelle **trace** d'une matrice carrée M la somme $\sum_{i=1}^n M_{i,i}$. On la note $tr(M)$. L'application $M \to tr(M)$ est une application linéaire.

On appelle **matrice unité d'ordre** n la matrice M avec $M_{i,j} = \delta_{i,j}$. C'est la matrice dans toute base de l'endomorphisme identité.

On appelle **matrice scalaire** une matrice égale à $\lambda.I$ avec λ un scalaire et I une matrice unité. C'est la matrice dans toute base de l'homothétie de rapport λ.

On appelle **matrice diagonale associée à un n-uplet** m la matrice M de type (n, n) définie par $M_{i,i} = m_i$ et $M_{i,j} = 0$ si $i \neq j$. On note $M = diag(m)$.

Une matrice est dite **symétrique** si elle est égale à sa transposée.

Une matrice M est dite **antisymétrique** si elle est égale à l'opposée de sa transposée, c'est-à-dire si ${}^tM = -M$.

Une matrice carrée est dite **triangulaire supérieure** si $j < i \to M_{i,j} = 0$

On note $\mathcal{T}_n^s$ l'ensemble des matrices carrées triangulaires supérieures d'ordre n.

Une matrice carrée est dite **triangulaire inférieure** si $j > i \to M_{i,j} = 0$

On note $\mathcal{T}_n^i$ l'ensemble des matrices carrées triangulaires inférieures d'ordre n.

Proposition 0.45
Quelques propriétés des matrices carrées :

•$\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre, isomorphe à la $\mathbb{K}$-algèbre des endomorphismes de $\mathbb{K}^n$ (ou de E, pour tout espace vectoriel E de dimension n).

•Si M est inversible, alors sa transposée aussi et $({}^tM)^{-1} = {}^t(M^{-1})$.

•$tr(AB) = tr(BA)$ (que A et B soient carrées ou non, pourvu que le produit soit carré)

•Une matrice est scalaire si et seulement si c'est la matrice associée à un endomorphisme de la forme $x \mapsto \lambda.x$, i.e. à une homothétie.

•L'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$.

•L'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$ et l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$; si $\mathbb{K}$ n'est pas de caractéristique 2, ils sont supplémentaires ; ils sont alors de dimensions respectives $\frac{n.(n+1)}{2}$ et $\frac{n.(n-1)}{2}$, engendrés respectivement par les $E_{i,j}+E_{j,i}$ pour $i \leq j$ et par les $E_{i,j}-E_{j,i}$ pour $i<j$. Toute matrice carrée M s'écrit $A+S$, avec A antisymétrique et S symétrique, avec $A=\frac{1}{2}(M-{}^tM)$ et $S=\frac{1}{2}(M+{}^tM)$

•Le produit de deux matrices symétriques est symétrique si et seulement si les deux matrices commutent.

•L'ensemble des matrices triangulaires supérieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$.

•L'ensemble des matrices triangulaires inférieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$.

•L'ensemble des matrices triangulaires inférieures est isomorphe à l'ensemble des matrices triangulaires supérieures. La transposition réalise un isomorphisme entre ces espaces vectoriels.

•Les éléments inversibles de l'ensemble des matrices triangulaires supérieures (resp. inférieures) sont les matrices triangulaires dont tous les coefficients diagonaux sont non nuls ; ils forment un sous-groupe multiplicatif de $\mathcal{M}_n(\mathbb{K})$.

•Étant donnée une matrice A de $\mathcal{M}_n(\mathbb{K})$, on appelle **commutant de** A le sous-ensemble de $\mathcal{M}_n(\mathbb{K})$ des matrices commutant avec A.

•Étant donnée A une matrice de $\mathcal{M}_n(\mathbb{K})$, on définit $A^0=I$ avec I la matrice unité d'ordre n, et $A^{i+1}=A.A^i$ (au sens du produit matriciel). On peut ainsi définir étant donné un polynôme l'image de A par ce polynôme, en utilisant la multiplication matricielle et la multiplication par un scalaire.

•Étant donnée A une matrice de $\mathcal{M}_n(\mathbb{K})$, l'application de $\mathbb{K}[X] \to \mathcal{M}_n(\mathbb{K})$ qui à P associe $P(A)$ est un morphisme d'algèbres. Son image est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$; on la note $\mathbb{K}(A)$, elle est commutative.

•$\mathbb{K}(A)$ est de dimension finie, au plus n.

Le cas des matrices carrées : la $\mathbb{K}$-algèbre $\mathcal{M}_n(\mathbb{K})$

Démonstration *Seul le dernier point pose difficulté. Il nécessitera le théorème de Cayley-Hamilton, qui montre que $A^n \in Vect(A^0,...,A^{n-1})$. La suite de la preuve est claire.*

Changement de bases Soit E espace vectoriel de dimension n, et (e_i) et (f_j) des bases de E.

DÉFINITION 0.38 **matrice de passage**

On appelle **matrice de passage** de la base (e_i) à la base (f_j) la matrice P de type (n,n) définie par $P_{i,j}=e_i^*(f_j)$; on la note $P_{(e_i),(f_j)}$. Il s'agit en fait de la matrice $Mat_{(f_j),(e_i)}(Id)$.

PROPOSITION 0.46

Quelques propriétés des matrices de passage :

•Le produit de la matrice de passage de B à B' par la matrice de passage de B' à B'' est la matrice de passage de B à B''.

•$P_{B,B'}^{-1}=P_{B',B}$

•Étant donné X le vecteur des coordonnées de x dans une base B, les coordonnées de x dans la base B' sont données par X' avec $X'=P_{B',B}X$.

•$Mat_{B',C'}(f)=P_{C',C}.Mat_{B,C}(f).P_{B,B'}$

•Dans le cas d'un endomorphisme $Mat_{B'}(f)=P_{B',B}.Mat_B(f).P_{B,B'}$

⚠ *Attention 0.8* La matrice de passage de B à B' donne les coordonnées dans B en fonction des coordonnées dans B' et pas le contraire... La terminologie vaut ce qu'elle vaut ! On peut le retenir en considérant que la matrice de passage de la base B à la base B' est la matrice dans B de l'endomorphisme dont l'image de B est B'.

Groupe linéaire et groupe spécial linéaire Voir **??**.

Groupe orthogonal réel et groupe spécial orthogonal réel Voir **??**. Ne pas confondre avec groupe linéaire et groupe spécial linéaire.

DÉFINITION 0.39 **Matrice associée à une famille finie de vecteurs**
Étant donnée une famille de vecteurs $(x_1, ..., x_p)$ d'un espace vectoriel de dimension finie n et une base $(v_1, ..., v_n)$ de E, on appelle **matrice associée à la famille des x_i et à la base des v_i** la matrice M définie par $M_{i,j} = v_i^*(x_j)$.

On appelle **rang d'une matrice** M le rang de la famille de ses vecteurs colonnes. On le note $rg(M)$.

Une matrice N extraite de M, avec N carrée inversible d'ordre le rang de M, est appelée **matrice principale** de M.

Rang d'une matrice *Application 0.9* Les matrices principales permettent de caractériser commodément les matrices définies positives, qui servent beaucoup en optimisation ; voir la section **??**.

La preuve des propriétés qui suivent est simple et permet de jongler un peu entre matrices et morphismes :

PROPOSITION 0.47
Quelques propriétés des rangs de matrices : •Le rang d'une matrice associée à un système de vecteurs et à une base est indépendant de cette base.

•$rg(Mat_{B,B'}(f)) = rg(f)$

•$rg({}^tM) = rg(M) \leq min(n,p)$, avec M de type (n,p).

•$rg(M) = n \iff f$ surjective

•$rg(M) = p \iff f$ injective

•$rg(MM') \leq min(rg(M), rg(M'))$

PROPOSITION 0.48
Le rang d'une matrice M non nulle est égal à n maximal tel qu'il existe une matrice extraite inversible de type (n,n) de M.

Démonstration *Il est clair que le rang de M est supérieur au rang de toute matrice extraite inversible (considérer une combinaison linéaire nulle des vecteurs correspondants de M).*

Étant donné r le rang de M il existe une famille libre de r vecteurs parmi la famille des vecteurs colonnes de M (par le théorème de la base incomplète).

On peut alors considérer la transposée de cette matrice : son rang est le même, et donc on peut se restreindre au même nombre de colonnes.

Matrices équivalentes, matrices semblables allons introduire et étudier les matrices semblables et les matrices équivalentes ; il faut bien noter que deux matrices sont équivalentes si elles représentent le même morphisme dans des bases différentes, alors que deux matrices sont semblables si elles représentent le même endomorphisme dans des bases différentes. Pour l'endomorphisme, on n'a qu'une base à changer ; pour un morphisme, on peut changer la base de départ et la base d'arrivée. Beaucoup de matrices de même type sont équivalentes : il suffit qu'elles aient même rang. La relation « être semblable » est beaucoup plus rare.

On va parler ici de matrices inversibles, mais on n'a pas de préoccupation vis à vis d'une éventuelle forme quadratique équipant l'espace vectoriel .

DÉFINITION 0.40 **Matrices équivalentes**
Deux matrices A et B de même type (n, p) sont dites **équivalentes** si il existe P et Q des matrices inversibles de types respectifs (p, p) et (n, n) telles que

$$B = Q.A.P$$

PROPOSITION 0.49
Quelques propriétés immédiates :
•Il s'agit d'une relation d'équivalence.
•Deux matrices de même type sont équivalentes si et seulement si elles représentent un même morphisme dans des bases différentes.
•Deux matrices de même type sont équivalentes si et seulement si elles ont le même rang.

⊡ **Matrices équivalentes** ⚠ *Attention 0.10* Dans la deuxième propriété, il faut bien voir qu'il faut changer éventuellement à la fois la base de départ et la base d'arrivée. On pourra faire le lien avec la proposition 0.41 quant aux matrices canoniques.

DÉFINITION 0.41 **Matrices semblables**
Deux matrices carrées A et B de même type sont dites semblables s'il existe une matrice inversible P telle que

$$A = P.B.P^{-1}$$

PROPOSITION 0.50

•Il s'agit d'une relation d'équivalence
•Deux matrices sont semblables si elles représentent un même endomorphisme dans deux bases différentes
•Deux matrices semblables ont même rang
•Deux matrices semblables ont même trace

⊡ **Matrices semblables**

Démonstration

Facile pour les trois premiers points ; pour le quatrième il suffit de se rappeller que la trace de AB est égale à la trace de BA, cf le troisième point de la proposition 0.45. C'est grâce à ce quatrième point que l'on peut définir la trace d'un endomorphisme, indépendamment du choix d'une base.

⚠ *Attention 0.11*

Cette fois-ci, contrairement au cas des matrices équivalentes, deux matrices de même rang ne sont pas nécessairement semblables.

⚠ *Attention 0.12*

La deuxième caractérisation fait appel à un endomorphisme et pas une application linéaire quelconque ; la base de départ est la même que celle d'arrivée.

Cofacteurs et inversion On va introduire différentes définitions dont un aboutissement sera la détermination analytique de l'inverse d'une matrice. Quoiqu'informatiquement inefficace (beaucoup trop lente), cette inversion est utile pour les propriétés analytiques qu'elle entraine.

DÉFINITION 0.42 **Mineur et cofacteur**

Le déterminant de la matrice M à laquelle on ôte la j-ième colonne et la i-ième ligne est appelé **mineur** (i,j) **de** M. On le note généralement $\Delta_{i,j}$.

Le déterminant de la matrice M à laquelle on ôte la j-ième colonne pour la remplacer par le i-ième vecteur de la base est appelé **cofacteur** (i,j) de M. On le note généralement $\gamma_{i,j}(M)$.

La matrice γ ainsi définie est appelée **comatrice** de M. On la note généralement $com(M)$.

La matrice ${}^t\gamma$ est appelée **matrice complémentaire** de M. On la note généralement $\tilde{M}$.

Pour y voir plus clair, le mineur (i,j) de M est :

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & M_{i-1,j+1} & \dots & M_{i-1,n} \\
M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & M_{i+1,j+1} & \dots & M_{i+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

et le cofacteur (i, j) de M est :

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
M_{i,1} & M_{i,2} & M_{i,3} & \dots & M_{i,j-1} & 1 & M_{i,j+1} & \dots & M_{i,n} \\
M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & 0 & M_{i+1,j+1} & \dots & M_{i+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

qui est d'ailleurs égal à

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\
M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & 0 & M_{i+1,j+1} & \dots & M_{i+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

Proposition 0.51

$$\gamma_{i,j} = (-1)^{i+j}.\Delta_{i,j}$$

Démonstration *Nous donnons ici un rapide plan de preuve. Considérer les permutations σ telles que $\sigma(i) = j$. Considérer alors leurs restrictions à $I = [[1, n]] \setminus \{i\}$ et $J = [[1, n]] \setminus \{j\}$ (on restreint à la fois le domaine et le codomaine). Considérer alors pour chaque σ ainsi restreint $\sigma' \circ \sigma \circ \sigma''$ avec σ' la seule bijection croissante de J vers $[[1, n-1]]$ et σ'' la seule bijection croissante de $[[1, n-1]]$ vers I. On a alors $\epsilon(\sigma) = \epsilon(\sigma' \circ \sigma \circ \sigma'') \times (-1)^{i+j}$.*

Proposition 0.52
La comatrice de la transposée de M est la transposée de la comatrice de M.

Il est nécessaire pour la suite d'avoir lu la partie **??**.

Théorème 0.53

$$\tilde{M}.M = det\ M.Id$$
$$\text{et} \quad M.\tilde{M} = det\ M.Id \text{en particulier, si } M \text{ est inversible, } \ M^{-1} = \frac{1}{det\ M}\tilde{M}$$

Démonstration *Considérons le terme (i,j) de la matrice $\tilde{M}.M$. Il s'agit de $\sum_{k=1}^{n} \gamma_{k,i}.M_{k,j}$.*

Considérons le terme (i,j) de la matrice det $M.Id$.

Il s'agit de $\delta_{i,j}.det\ M$.

On cherche donc à montrer que $\delta_{i,j}.det\ M = \sum_{k=1}^{n} \gamma_{k,i}.M_{k,j}$.

On distingue deux cas :

•$i \neq j$

On considère alors la matrice M, sur laquelle on remplace la colonne i par la colonne j (on ne les permute pas, on supprime la colonne i, et on copie la colonne j à la place).

Le déterminant de la matrice obtenue est nul, puisqu'on a deux fois la même colonne. On développe par rapport à la colonne i. On obtient

$$\sum_{k=1}^{n} M_{k,j}.\gamma_{k,i} = 0$$

et donc le résultat dans ce cas est bien montré.

•$i = j$

Dans ce cas on considère le développement de M par rapport à la i-ième colonne et on obtient

$$det\ M = \sum_{k=1}^{n} \gamma_{k,i}.M_{k,i} = \sum_{k=1}^{n} \gamma_{k,i}.M_{k,j}$$

d'où le résultat souhaité.

Le second résultat se déduit du premier en considérant la transposée de chacun des deux produits.

Matrices par blocs Les matrices par blocs sont fort pratiques pour les produits, les déterminants, les inversions. En particulier, les ordinateurs fonctionnent en général avec des caches, c'est-à-dire des mémoires très rapides intermédiaires entre le processeur et la mémoire vive : ces « caches » ont une taille très limitée, et dimensionner les blocs d'un calcul par blocs de manière à ce que le cache ne soit pas débordé (ce qui entraine des communications trop fréquentes avec la beaucoup plus lente mémoire vive) change tout à la performance d'un algorithme. On peut en fait déterminer la taille du cache en examinant la vitesse d'un calcul matriciel en fonction de la taille des blocs. Notons que des découpages par bandes existent aussi.

Les méthodes expliquées ci-dessous dans le cas de quatre blocs, se généralisent naturellement au cas d'un nombre quelconque de blocs.

Produit par blocs Étant données les matrices A, B, C, D, A', B', C' et D', avec

$$largeur(A) = largeur(C) = hauteur(A') = hauteur(B')$$

$$largeur(B) = largeur(D) = hauteur(C') = hauteur(D')$$

on considère les matrices M et M' définies comme suit :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

Alors

$$M.M' = \begin{pmatrix} A.A' + B.C' & A.B' + B.D' \\ C.A' + D.C' & C.B' + D.D' \end{pmatrix}$$

Inverse par blocs Si $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, avec A et B inversibles, alors M est inversible et $M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}.C.B^{-1} \\ 0 & B^{-1} \end{pmatrix}$.

Déterminant par blocs Si $M = \begin{pmatrix} A & C0 & B \end{pmatrix}$, alors $det\ M = det\ A.det\ B$.

Opérations sur les lignes et les colonnes On trouvera des manipulations proches de celles décrites ici dans la partie sur les déterminants **??**. La proposition **??** illustre aussi des opérations sur les lignes et les colonnes.

Définition 0.43 **système d'équations linéaires**
On appelle **système d'équations linéaires** une équation de la forme $MX = Y$, où M (matrice) et Y (vecteur) sont donnés et où X est l'inconnue.

Les **opérations sur les lignes et les colonnes** d'une matrice ou d'un système linéaire sont par définition :

•l'addition d'une ligne (resp. colonne) i à une ligne (resp. colonne) $j \neq i$

•la multiplication d'une ligne (resp. colonne) i par un scalaire $\lambda \neq 0$

•la permutation de deux lignes (resp. colonnes) i et $j \neq i$

Ces opérations seront notées respectivement :

•$L_i \leftarrow L_i + L_j$ (resp. $C_i \leftarrow C_i + C_j$)

•$L_i \leftarrow \lambda.L_i$ (resp. $C_i \leftarrow \lambda.C_i$)

•$L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$)

pourra éventuellement ajouter à une ligne (resp. une colonne) une autre ligne (resp. colonne) multipliée par un scalaire λ ; cela se notera $L_i \leftarrow L_i + \lambda L_j$ (resp. $C_i \leftarrow C_i + \lambda C_j$) : il s'agit d'une composition des deux premiers types d'opérations.

Exemple 0.13 • Examinons un système sur lequel on « bricole » sur les lignes ou colonnes. Avant :

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

Après $L_1 \leftarrow L_1 + 2\ L_2$:

$$\begin{pmatrix} 1 & 4 & -1 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}$$

•Sur une matrice maintenant. Avant :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Après $C_1 \leftrightarrow C_2$

$$\begin{pmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{pmatrix}$$

Proposition 0.54
Les opérations sur les lignes correspondent à des multiplications à gauche par des matrices inversibles ; les opérations sur les colonnes correspondent à des multiplications à droite par des matrices inversibles. L'inverse d'une opération sur les lignes (resp. les colonnes) est une opération sur les lignes (resp. les colonnes).

Démonstration *$L_i \leftarrow L_i + L_j$ correspond à la multiplication à gauche par $I + E_{i,j}$;*

$L_i \leftarrow \lambda.L_i$ correspond à la multiplication à gauche par $diag(1, ..., 1, l, 1, ..., 1)$, i.e. par la matrice avec des 1 sur la diagonale, sauf λ en place (i, i) ;

$L_i \leftrightarrow L_j$ correspond à la multiplication à gauche par $I + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$.

Pour les opérations sur les colonnes, on se ramène au même type d'opérations sur les lignes par transposition.

Proposition 0.55
•Le déterminant d'une matrice n'est pas modifié en ajoutant à une ligne une combinaison linéaire des autres lignes.

•Multiplier une ligne par λ multiplie le déterminant par λ.

•Permuter deux lignes multiplie le déterminant par -1.

Les mêmes résultats sont valables pour les colonnes.

Démonstration *Cela découle immédiatement des propriétés du déterminant, et du fait que le déterminant d'une matrice est égal au déterminant de sa transposée. Pour des rappels, on peut consulter la partie* **??**.

Théorème 0.56 **Formules de Cramer**
Considérons le système d'équations linéaire $MX = Y$, avec M de type (n, n) :

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,n} \end{pmatrix}, \quad Y = {}^t(y_1, ..., y_n) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

On suppose en outre que M est <u>inversible</u>.

Alors X est solution, avec

$$x_i = \frac{\begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}}{det\ M}$$

Démonstration *La solution X est clairement unique, car par inversibilité de M, $X = M^{-1}Y$.*
On a alors $Y = \sum_{k=1}^{n} X_k C_k$ avec C_k la k-ième colonne de M.

Donc, quel que soit i, on a $M^{(i)} = \sum_{k=1}^{n} X_k {M'_k}^{(i)}$

avec

$$M_k{}^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}$$

et

$${M'_k}^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & M_{1,k} & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & M_{2,k} & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & M_{n,k} & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}$$

On en déduit donc, en supprimant de la somme $\sum_{k=1}^{n} X_k {M'_k}^{(i)}$ les termes nuls (déterminants avec deux colonnes identiques), $X_i {M'_i}^{(i)} = M^{(i)}$, or ${M'_i}^{(i)} = det\ M$, d'où le résultat voulu.

Théorème 0.57 **Méthode du pivot de Gauss**

(déterminants avec deux colonnes identiques)
La **méthode de Gauss** consiste à :

1) permuter les lignes pour avoir un coefficient non nul en haut à gauche de la matrice ; ce coefficient est appelé **pivot**

2) soustraire la première ligne multipliée par un coefficient adéquat à chacune des autres lignes de manière à avoir des zéros sur toute la première colonne en dehors du premier coefficient

3) Procéder de même sur la matrice extraite, simplement dépourvue de sa première ligne et sa première colonne.

Le point 1) pourra toujours être réalisé si on trouve toujours un coefficient non nul à échanger ; pour peu que la matrice soit inversible, cette condition sera toujours vérifiée. Si elle ne l'est pas, on peut travailler sur la matrice extraite par suppression de la première colonne.

En réitérant cette méthode, on arrive à obtenir une matrice triangulaire supérieure. En fait la matrice obtenue est de la forme illustrée sur la figure 2, du moins après permutation des colonnes.

La matrice ainsi obtenue est donc beaucoup plus maniable : le calcul du déterminant (si la matrice est carrée) se résume à la multiplication des éléments diagonaux, la résolution de systèmes

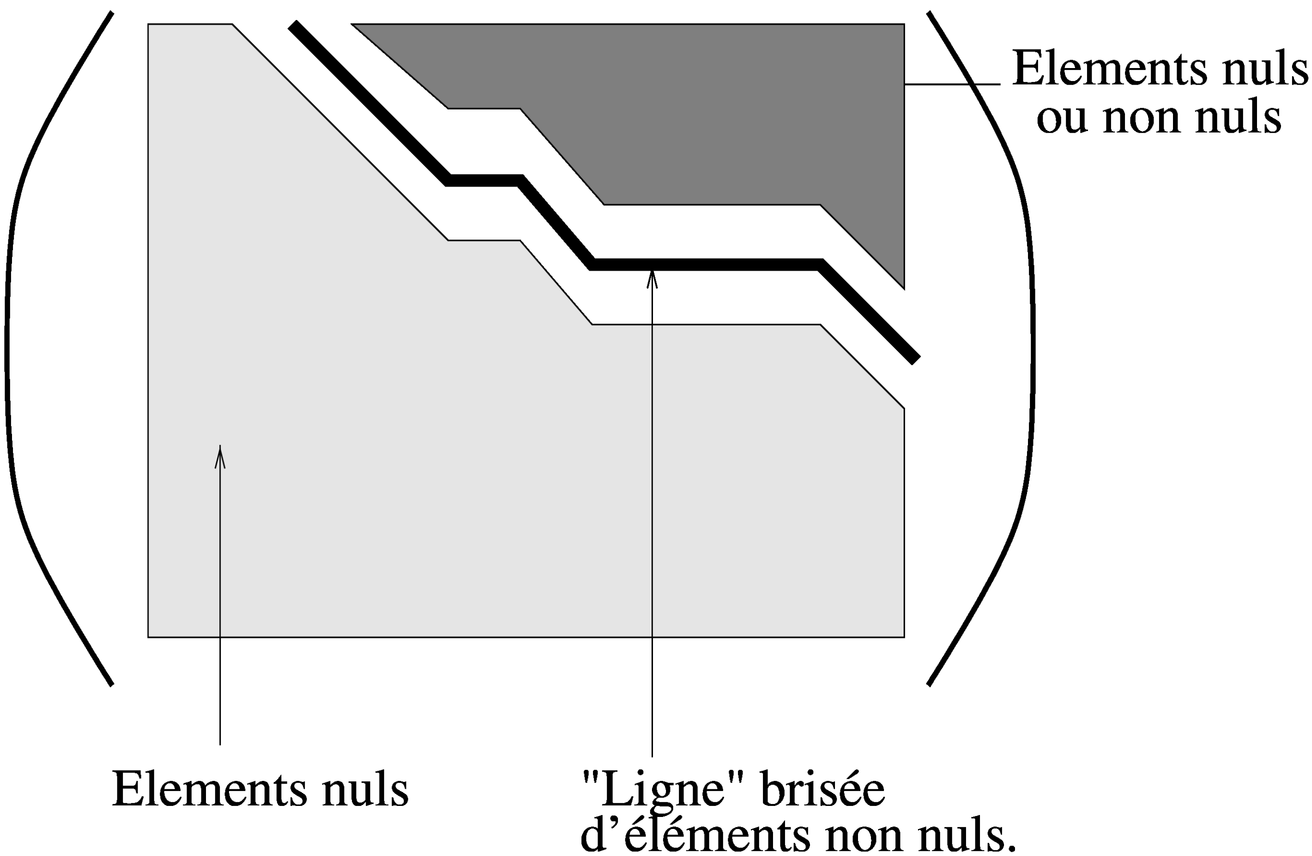


FIGURE 2 – Matrice obtenue après pivot de Gauss.
Commentaire : La « ligne brisée » évoquée comporte soit des sauts à droite, soit des sauts en diagonale en bas à droite. Le premier élément de la ligne brisée se trouve quelque part sur la première ligne (éventuellement en haut à gauche).

linéaires est plus aisée (il convient de noter pour cela que les opérations sur les colonnes sont de simples permutations sur les inconnues), le calcul du rang est immédiat (nombre d'éléments avant le dernier élément non nul de la dernière colonne).

Intuition Afin de minimiser les pertes de précision d'un calcul numérique, il est préférable de choisir un pivot grand en valeur absolue.

La méthode du **pivot total** consiste à chercher le pivot non pas seulement sur la colonne en cours, mais de permuter éventuellement les colonnes pour avoir un pivot plus grand. Par opposition au pivot total, la méthode ci-dessus est dite **pivot partiel**.

THÉORÈME 0.58 **Décomposition** $A = LU$

Étant donnée une matrice A supposée inversible, on définit $a_k = |(A_{i,j})_{i,j \leq k}|$, déterminant de la matrice obtenue en se restreignant aux k premières lignes et k premières colonnes.

On appelle **décomposition** $A = LU$ un produit du type $A = LU$ avec L matrice triangulaire inférieure ne comportant que des 1 sur la diagonale, U matrice triangulaire supérieure.

Alors, il existe une décomposition $A = LU$ si et seulement si les a_k sont non nuls pour tout k dans $[\![1, n]\!]$.

Démonstration *Il suffit d'utiliser la méthode de Gauss, en considérant les matrices correspondant aux opérations sur les lignes et les colonnes. C'est-à-dire que l'on obtient un produit $\prod_{i=1}^{n} M_{n-i} A = U$, avec M_i la matrice correspondant à l'opération sur les lignes et les colonnes effectué à la i-ième étape. Mais l'inverse d'une opération sur les lignes ou les colonnes est une opération sur les lignes ou les colonnes ; donc on peut aussi écrire $A = \prod_{i=1}^{n} N_i U$, avec N_i l'inverse de M_i. Le produit des N_i est bien de la forme désirée, triangulaire supérieure à diagonale unité, comme on s'en convaincra en considérant la stabilité de l'ensemble des matrices triangulaires supérieures à diagonale unité par multiplication par les matrices des opérations sur les lignes et les colonnes utilisées dans le pivot de Gauss.*

La condition sur les déterminants ne couvre pas toutes les matrices A inversibles possibles. Cela va être résolu avec la proposition qui suit :

PROPOSITION 0.59

Si A est inversible, il existe une matrice de permutation P telle que $PA = LU$.

Démonstration *Si a_k est nul, il existe nécessairement une permutation de lignes qui arrange ça, sinon A ne pourrait pas être de rang plein - il suffit donc de multiplier les différentes permutations de lignes nécessaires pour obtenir une matrice vérifiant les conditions demandées.*

THÉORÈME 0.60 **Décomposition de Cholesky**

On appelle **décomposition de Cholesky** ou **décomposition** $A = {}^tRR$, un produit de la forme $A = {}^t RR$, avec R triangulaire inférieure inversible.

A admet une décomposition de Cholesky si et seulement si A est symétrique définie positive.

Démonstration *On pourra consulter le livre [3] pour cette preuve.*

Exercices sur les matrices *Exemple 0.14* Une forme linéaire f sur $\mathcal{M}_n(\mathbb{K})$ telle que $f(XY) = f(YX)$ est proportionnelle à l'application $M \mapsto tr(M)$.

Démonstration *Il suffit de considérer X et Y des matrices élémentaires, et de généraliser à toute matrice par linéarité de f.*

Exemple 0.15 Une matrice de $\mathcal{M}_n(\mathbb{K})$ commute avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$ si et seulement si elle est scalaire, i.e. de la forme λId avec Id la matrice identité et $\lambda \in \mathbb{K}$.

Démonstration *Écrire la condition de commutativité avec les matrices élémentaires suffit.*

1.10.4 Zoologie sur les matrices et leurs déterminants

On réutilise la notion de matrice circulante vue en section **??** :

PROPOSITION 0.61
L'ensemble des matrices circulantes de type (n, n) est une K-algèbre engendrée par la matrice M_0 suivante :

$$M_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Démonstration *Il convient de vérifier tout d'abord qu'il s'agit bien d'une K-algèbre. Pour ensuite voir qu'elle est engendrée par M_0, il suffit de voir que la matrice circulante associée à $(x_1, ..., x_n)$ est la matrice*

$$x_1.{M_0}^0 + x_2.{M_0}^1 + ... + x^n.{M_0}^{n-1}$$

et le résultat est acquis.

1.10.5 Zoologie de la dualité en dimension finie

On va présenter ici les polynômes de Lagrange (avec leurs propriétés d'interpolation), et la réécriture de sous-espaces vectoriels en tant qu'intersection d'hyperplans.

Polynômes de Lagrange On considère l'espace $E = \mathbb{R}_n[X]$ des polynômes à une indéterminée et de degré au plus n. Soient a_0, ..., a_n des réels deux à deux distincts. On définit $n + 1$ formes linéaires sur E par $f_i(P) = P(a_i)$.

PROPOSITION 0.62

Les (f_i) forment une base de E^*.

Démonstration *Puisque la dimension de E est égale à la dimension de E^*, il suffit de voir que la famille est de rang $n+1$, ce qui est vérifié si et seulement si l'espace vectoriel dual est de dimension 0. Supposons qu'un certain polynôme P appartienne à cet orthogonal, alors il s'annule en a_0, ..., a_n ; donc il est nul, puisqu'il est de degré au plus n.*

PROPOSITION 0.63 Polynômes de Lagrange
(f_i) est la base duale des P_i, avec

$$P_i(x) = \Pi_{j\neq i}\frac{X - a_j}{a_i - a_j}$$

Les P_i sont appelés polynômes de Lagrange. L'intérêt de ces polynômes va être la facilité avec laquelle on va interpoler avec ces polynômes ; il suffit de connaître leurs valeurs en les a_i. Malheureusement, cette technique est une interpolation, c'est-à-dire que le polynôme obtenu par interpolation (corollaire ci-dessous) incorpore tous les bruits et converge en pratique très mal vers une vraie fonction connue seulement par ses valeurs en un nombre fini de points, dès que cette fonction n'est pas « vraiment » un polynôme. Ce constat est à la base de la théorie de la régularisation (e.g., régularisation de Tykhonov).

Démonstration *Facile, il suffit de vérifier que* $P_i(a_j) = \delta_{i,j}$.

COROLLAIRE 0.64
On en déduit notamment que tout polynôme de degré n s'écrit comme combinaison linéaire des P_i, les coefficients étant donnés par les f_i, c'est-à-dire que tout P de degré $\leq n$ s'écrit

$$P = \sum_{i=1}^{n} f_i(P)P_i = \sum_{i=1}^{n} P_i^*(P)P_i = \sum_{i=1}^{n} P(a_i)P_i$$

Un exemple d'utilisation Maple :

Exemple Maple

```
> interp([0,1,2,3],[exp(0),exp(1),exp(2),exp(3)],x);
```

$$\frac{1}{6}e^3x^3 - \frac{1}{2}e^3x^2 + \frac{1}{3}e^3x - \frac{1}{2}e^2x^3 +$$
$$2e^2x^2 - \frac{3}{2}e^2x + \frac{1}{2}ex^3 - \frac{5}{2}ex^2 + 3ex - \frac{1}{6}x^3 + x^2 - \frac{11}{6}x + 1$$

Il est intéressant de tracer ensuite la courbe exponentielle et les graphes des interpolations à différents ordres superposés.

PROPOSITION 0.65
Soit F un sous-espace vectoriel de E de dimension p, E étant un espace vectoriel de dimension finie n. Alors il existe $n - p$ formes linéaires linéairement indépendantes f_i telles que $F = \{x \ ; \ \forall i \ f_i(x) = 0\}$ c'est-à-dire que F s'exprime comme intersection de $n - p$ hyperplans. Il est en outre impossible de définir F comme intersection de moins de $n - p$ hyperplans.

Définition d'un sous-espace vectoriel par une famille d'équations

Démonstration *Pour voir qu'une telle famille existe, il suffit de considérer une base de F et sa base duale. Pour vérifier qu'on ne peut faire moins, il suffit de considérer que pour tout G sous-espace vectoriel de E et tout H hyperplan de E on a $dim\ (G \cap H) \geq dim\ G - 1$ (par la formule $dim\ G + dim\ H = dim(G + H) + dim(G \cap H)$).*

Cette proposition permet de voir certaines choses clairement : ainsi, tout sous-espace vectoriel de $\mathbb{R}^n$ peut être écrit sous formes d'un systèmes d'équations linéaires ; toute droite de $\mathbb{R}^3$ est une intersection de deux hyperplans.

Références

[1] B. Chazelle, *The Discrepancy Method*, Cambridge University Press, 2000.

[2] H. Brézis, *Analyse fonctionnelle*, Masson, 1983.

[3] P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod, 1998.