Chapitre 1

Espaces vectoriels de dimension finie

1.1 Premières définitions

On travaille sur un corps noté $\mathbb{K}$, que l'on pourra imaginer être, à votre guise, le corps des réels, celui des complexes, ou le corps $\mathbb{Z}/p\mathbb{Z}$, avec p premier.

Définition 1.1.1. Un espace vectoriel sur $\mathbb{K}$ est un ensemble, notons-le E, muni d'une opération interne (de $E \times E$ dans E) notée $+$ et d'une multiplication externe de $\mathbb{K}$, c'est-à-dire une application de $\mathbb{K} \times E$ dans E qui envoie (λ, u) sur $\lambda \cdot u$ (souvent noté λu), vérifiant les propriétés suivantes :

i. $(E, +)$ est un groupe abélien (élément neutre, noté souvent 0 et appelé vecteur nul, associativité, inverse, l'inverse de u sera noté $-u$, et commutativité),

ii. $1.u = u$, 1 étant le neutre du groupe multiplicatif $\mathbb{K}^*$,

iii. $(\lambda + \mu).u = \lambda.u + \mu.u$,

iv. $\lambda.(u + v) = \lambda.u + \mu.v$,

v. $(\lambda\mu).u = \lambda.(\mu.u)$.

Exemple 1.1.2. $(\mathbb{K}^n, +, \cdot)$, $(\mathbb{K}[X], +, \cdot)$, $(\mathbb{K}^X, +, \cdot)$, où $\mathbb{K}^X$ désigne l'ensemble des fonctions d'un ensemble X vers le corps $\mathbb{K}$. En particulier, si $X = \mathbb{N}$, on obtient l'ensemble des suites à valeurs dans $\mathbb{K}$.

L'ensemble des solutions d'un système linéaire à coefficients dans $\mathbb{K}$ (sans second membre) est un espace vectoriel. Idem pour l'ensemble des solutions d'une équations différentielle linéaire sans second membre.

Contre-exemple 1.1.3. Avec quoi il ne faut pas confondre les espaces vectoriels. Tout d'abord, les espaces vectoriels sont des objets droites et linéaires, très rigides, un cercle de $\mathbb{R}^2$ n'est pas un espace vectoriel pour une structure naturelle.

Plus subtil : il ne faut pas confondre espace vectoriel et espace affine. Un espace vectoriel possède un élément neutre qui est unique de par ses propriétés. Il est nul, mais la structure dans lequel il est plongé lui donne ces privilèges (un peu comme Donald Trump). Alors que dans un espace affine, aucun point n'a plus de privilège qu'un autre point.

Définition 1.1.4. On appelle combinaison linéaire d'une famille de vecteurs u_i, $i \in I$, toute somme finie $\sum_{i\in I} \lambda_i.u_i$, avec les $\lambda_i \in \mathbb{K}$, presque tous nuls.

Remarque 1.1.5. Notez bien cette propriété de transitivité qui dit qu'une combinaison linéaire de combinaison linéaires d'une famille fixée de vecteurs, reste une combinaison linéaire de cette famille. Cela peut se montrer avec force calculs en utilisant tous les axiomes de distributivité et pseudo-associativité mis à disposition par les espaces vectoriels.

Définition 1.1.6. Une famille de vecteurs $(u_i)_{i\in I}$ de E est dite libre si une combinaison linéaire $\sum_{i\in I} \lambda_i.u_i = 0$ implique tous les λ_i nuls.

Définition 1.1.7. Une famille de vecteurs $(u_i)_{i\in I}$ de E est dite génératrice (pour E) si tout vecteur de E peut s'écrire comme combinaison linéaire de la famille $(u_i)_{i\in I}$.

Remarque 1.1.8. Lorsqu'on retire un vecteur d'une famille libre, celle-ci reste libre. Lorsqu'on ajoute un vecteur à une famille génératrice, celle-ci reste génératrice.

Si l'on permute les éléments d'une famille libre, resp. génératrice, celle-ci reste libre, resp. génératrice.

Enfin, un élément non nul constitue une partie libre (exercice). La famille des éléments de E constitue une partie génératrice. Notez que la première assertion demande de travailler sur un corps, alors que la seconde ne l'exige pas.

On n'a pas encore défini la dimension, mais on peut définir la "dimension-finie".

Définition 1.1.9. Un espace vectoriel E est dit (provisoirement) de "dimension-finie" si E possède une famille génératrice finie.

Remarque 1.1.10. L'idée est ici de généraliser les propriétés efficaces des ensembles finis : on sait que si A et B sont deux ensembles finis tels que $A \subset B$, alors $|A| = |B|$ implique $A = B$, où $|.|$ désigne le cardinal. En général, les espaces vectoriels sont des ensembles infinis, mais la dimension (finie) va permettre de palier ce problème.

1.2 Le lemme d'échange

Lemme 1.2.1. *(Lemme d'échange)*
Soit E un espace de "dimension-finie", $(f_1, \cdots, f_p)$ une partie libre, et $(g_1, \cdots, g_q)$ une famille génératrice de E. Alors, il existe j tel que $(f_1, \cdots, f_{p-1}, g_j)$ est libre.

Démonstration. On va le prouver par l'absurde. Supposons au contraire que pour tout j de 1 à q, $(f_1, \cdots, f_{p-1}, g_j)$ est non libre (on dira qu'elle est liée).

Alors, pour tout j, on peut trouver des scalaires non tous nuls $\lambda_1, \cdots, \lambda_{p-1}$, μ_j tels que

$$\lambda_1 f_1 + \cdots + \lambda_{p-1} f_{p-1} + \mu_j g_j = 0.$$

Montrons que μ_j est non nul. En effet, s'il l'était, on aurait $\lambda_1 f_1 + \cdots + \lambda_{p-1} f_{p-1} = 0$, et comme $(f_1, \cdots, f_{p-1})$ est libre, les λ_i seraient nuls également, ce qui est absurde. Donc, $\mu_j \neq 0$. Mais comme $\mathbb{K}$ est un corps, μ_j est inversible. Il vient alors

$$g_j = -\mu_j^{-1}\lambda_1 f_1 - \cdots - \mu_j^{-1}\lambda_{p-1} f_{p-1}.$$

Conclusion, pour tout j, g_j est une combinaison linéaire des f_i, pour i de 1 à $p-1$.

Or, f_p est combinaison linéaire de la famille $(g_1, \cdots, g_q)$, puisque celle-ci est génératrice de E. Donc, f_p est combinaison linéaire des f_i, pour i de 1 à $p-1$. Ceci est impossible car on aurait une combinaison non triviale dans la famille $(f_1, \cdots, f_p)$ qui est censée être libre.

Cela prouve le lemme d'échange.

◇

Le lemme d'échange nous permet de montrer la proposition :

Proposition 1.2.2. *On suppose que E contient une partie libre $(f_1, \cdots, f_p)$, et une partie génératrice $(g_1, \cdots, g_q)$. Alors, $p \leq q$.*

Démonstration. On va appliquer plusieurs fois le lemme d'échange, ainsi que la stabilité par permutation.

On note j_1 un nombre (assuré par le lemme d'échange) tel que $(f_1, \cdots, f_{p-1}, g_{j_1})$ est libre. Par permutation, $(g_{j_1}, f_1, \cdots, f_{p-1})$ est encore libre.

On applique donc le lemme d'échange à la partie libre $(g_{j_1}, f_1, \cdots, f_{p-1})$ et la partie génératrice $(g_1, \cdots, g_q)$. On obtient un nombre j_2 tel que $(g_{j_1}, f_1, \cdots, f_{p-2}, g_{j_2})$ est libre. Par permutation $(g_{j_1}, g_{j_2}, f_1, \cdots, f_{p-2})$ est libre. Par récurrence, on obtient des nombres $j_1, j_2, \cdots, j_p$ tels que $(g_{j_1}, g_{j_2}, \cdots, g_{j_p})$ est libre.

Montrons que les j_k sont tous deux à deux distincts. En effet, si $j_k = j_{k'}$ pour $k \neq k'$, on aurait $g_{j_k} - g_{j_{k'}} = 0$, ce qui prouverait que $(g_{j_k}, g_{j_{k'}})$ n'est pas libre et donc, par adjonction, $(g_{j_1}, g_{j_2}, \cdots, g_{j_p})$ serait non libre. Contradiction.

Conclusion, les j_k sont tous distincts. On a donc extrait p nombres distincts (les j_k) parmi des nombres de 1 à q. Ceci n'est possible que si $p \leq q$.

◇

Il est temps de définir la notion de base.

Définition 1.2.3. Une base de E est une famille à la fois libre et génératrice.

Remarque 1.2.4. Une base $\underline{e} := (e_1, \cdots, e_n)$ de E permet d'identifier E à $\mathbb{K}^n$. En effet, si $\underline{e}$ est une base, alors, tout vecteur u de E peut se décomposer en $u = \sum_{i=1}^{n} \lambda_i e_i$ (car $\underline{e}$ est génératrice), et la famille $(\lambda_1, \cdots, \lambda_n)$ de $\mathbb{K}^n$ est unique.

L'espace $\mathbb{K}^n$ devient alors un espace de référence.

Théorème 1.2.5. *Soit E un espace de "dimension-finie". Alors, toutes les bases ont même cardinal.*

Démonstration. Si $(e_1, \cdots, e_p)$ et $(e'_1, \cdots, e'_{p'})$ sont deux bases de E, alors, en appliquant la proposition qui précède, on a d'une part $p \leq p'$ mais aussi par symétrie, $p' \leq p$. On en déduit que $p = p'$. ◇

Remarque 1.2.6. Attention, pour l'instant, on n'a pas prouvé que E possède des bases. Toute l'astuce est dans la convention que toute proposition sur l'ensemble vide est vraie ! Par exemple, si un élève n'a jamais rendu son devoir en confinement, vous pouvez écrire "le travail est fait sérieusement" sur son bulletin.

1.3 Théorème de la base incomplète

Le théorème de la base incomplète va nous permettre de donner la preuve de l'existence de bases, mais aussi de les construire. Le voici déjà sous sa forme complète (Haha !). Attention toutefois à ne pas confondre cette version avec le lemme d'échange qui lui ressemble étrangement.

Théorème 1.3.1. *Théorème de la base incomplète Soit $(f_1, \cdots, f_p)$ une famille libre d'un espace E engendré par une famille finie $(g_1, \cdots, g_q)$. Alors,*

i. soit $(f_1, \cdots, f_p)$ est une base de E,

ii. sinon, il existe j tel que $(f_1, \cdots, f_p, g_j)$ est une famille libre.

Démonstration. On suppose donc que la famille $(f_1, \cdots, f_p, g_j)$ est non libre pour tout j de 1 à q et on veut montrer qu'alors, $(f_1, \cdots, f_p)$ est une base de E.

Il n'est pas mauvais de reprendre l'argument déjà utilisé au moment de la preuve du lemme d'échange :

Pour tout j, on peut trouver des scalaires non tous nuls $\lambda_1, \cdots, \lambda_p$, μ_j tels que

$$\lambda_1 f_1 + \cdots + \lambda_p f_p + \mu_j g_j = 0.$$

Le scalaire μ_j est non nul. En effet, s'il l'était, on aurait $\lambda_1 f_1 + \cdots + \lambda_p f_p = 0$, et comme $(f_1, \cdots, f_p)$ est libre, les λ_i seraient nuls également, ce qui est absurde. Donc, $\mu_j \neq 0$. Mais comme $\mathbb{K}$ est un corps, μ_j est inversible.

Conclusion, pour tout j, g_j peut s'écrire comme combinaison linéaire de la famille $(f_1, \cdots, f_p)$. Comme tout vecteur de E peut s'écrire comme combinaison linéaire des g_j et comme tous les g_j peuvent s'écrire comme combinaison linéaire des f_i, il vient que tout vecteur de E peut s'écrire comme combinaison linéaire des f_i, ce qui prouve que la famille $(f_1, \cdots, f_p)$ est bien une base de E.

◇

Il faut surtout retenir la version courte :

Théorème 1.3.2. *Tout famille libre d'un espace vectoriel de "dimension-finie" peut être complétée en une base. En particulier, si E est non nul, il existe une base (non vide).*

Démonstration. Un vecteur non nul forme une partie libre et on le complète en une base par le théorème précédent. ◇

Remarque 1.3.3. Par extension, l'espace nul a pour base l'ensemble vide. A l'opposé, un espace vectoriel qui n'est pas de "dimension-finie" possède également une base, si l'on accepte le fameux "axiome du choix".

Définition 1.3.4. Si E admet une famille génératrice finie, on peut lui associer un (unique) nombre n égal au cardinal de toutes bases. Ce nombre est appelé dimension de E et noté $\dim E$. On pourra dire désormais que E est de dimension finie (n).

Corollaire 1.3.5. *Dans un espace de dimension n, une partie libre est une base si et seulement si elle est de cardinal n.*

Démonstration. Soit $(f_1, \cdots f_m)$ une partie libre de E, de dimension n. Si $(f_1, \cdots f_m)$ est une base, alors, par définition-théorème, $m = n$.

Réciproquement, supposons $m = n$. Comme $(f_1, \cdots f_m)$ est libre, elle peut se compléter en une base de E, mais comme $m = n$ et comme toutes les bases sont de cardinal n, il s'agit d'une complétion triviale (on n'a rien ajouté). Donc, $(f_1, \cdots f_m)$ est une base. ◇

On peut commencer un premier volet sur le dénombrement sur corps fini. C'est le début d'une belle aventure qui, on l'espère, en éclairera plus d'un sur l'algèbre linéaire, selon le principe général qui peut s'énoncer en "compter, c'est comprendre".

Proposition 1.3.6. *Soit E un espace vectoriel de dimension n sur un corps $\mathbb{K}$ fini de cardinal q. ALors, le cardinal de E est fini, égal à q^n.*

Démonstration. On a vu que E possédait au moins une base, et qu'une base permet d'identifier (bijectivement) chaque vecteur de E à un élément de $\mathbb{K}^n$. Donc, le cardinal de E est égal au cardinal de $\mathbb{K}^n$, c'est-à-dire q^n. ◇

1.4 Théorème de la base extraite

On a aussi des théorèmes analogues pour extraire une base à partir d'une famille génératrice.

Théorème 1.4.1. *Théorème de la base extraite*

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. et $(g_j)_{j\in J}$ une famille génératrice de E (qui peut éventuellement être infinie). Alors, on peut extraire une base de E, c'est-à-dire une partie J_0 de J à n éléments telle que $(g_j)_{j\in J_0}$ est une base de E.

Démonstration. Montrons tout d'abord que l'on peut extraire une famille finie $(g_j)_{j\in J_1}$ qui reste génératrice.

Par hypothèse, il existe une base $(e_i)_{1\leq i\leq n}$ de E. De plus, chaque élément s'écrit comme combinaison linéaire (finie !) des g_j. Comme les e_i sont en nombre fini et que chaque e_i se décompose selon une famille finie de g_j, on peut trouver une sous-famille J_1 de J telle que tous les e_i s'écrivent comme combinaison linéaire des g_j, $j \in J_1$.

Comme la famille $(e_i)_{1\leq i\leq n}$ est génératrice, tout vecteur de E s'écrit comme combinaison linéaire de e_i donc, comme combinaison linéaire des g_j, $j \in J_1$. On tient donc notre famille génératrice finie extraite. Quitte à retirer des vecteurs de la famille, on peut supposer que J_1 est minimale telle que $(g_j)_{j\in J_1}$ est génératrice.

Montrons alors que $(g_j)_{j\in J_1}$ est une base. Supposons que $(g_j)_{j\in J_1}$ ne soit pas une base. Alors, elle n'est pas libre. On peut trouver une combinaison linéaire $\sum_{j\in J_1} \lambda_j g_j$ non triviale des g_j, et donc, en inversant un coefficient, disons λ_i non nul, on voit que g_i peut s'écrire en fonction des autres. La famille $(g_j)_{j\in J_1}$ étant génératrice, on peut retirer g_i de la famille $(g_j)_{j\in J_1}$ et celle-ci reste génératrice. Ceci est en contradiction avec la minimalité de la famille.

◇

Corollaire 1.4.2. *Soit E un espace vectoriel de dimension n. Une famille génératrice $(g_j)_{j\in J}$ est une base si et seulement si elle est de cardinal n.*

Démonstration.
On utilise encore une fois le fait que l'on peut extraire de $(g_j)_{j\in J}$ une base, et que toutes les bases ont le même cardinal. Ceci implique que l'on n'a rien retiré à cette famille, et donc qu'elle est bien une base. ◇

1.5 Applications aux sous-espaces vectoriels

Définition 1.5.1. Si $(E, +, .)$ est un espace vectoriel, un sous-espace vectoriel de E est une partie F de E telle que $(F, +, .)$ soit également un espace vectoriel (pour les mêmes lois + et . que E).

On a tout intérêt à montrer qu'une partie est un sous-espace vectoriel à l'aide de cette proposition.

Proposition 1.5.2. *Soit F une partie d'un espace vectoriel E. F est un sous-espace vectoriel de E si, et seulement si, il vérifie les conditions suivantes :*

i. F est non vide,

ii. pour tout λ, μ de $\mathbb{K}$, u, v de F, $\lambda.u + \mu.v \in F$.

Démonstration. A faire en exercice. En gros, il faut se poser les questions suivantes :

i. pourquoi si u, v sont dans F, $u + v$ l'est également ?

ii. pourquoi si u est dans F, $-u$ l'est également ?

iii. Pourquoi le vecteur nul est dans F ?

Le reste est hérité des propriétés de E. ◇

Remarque 1.5.3. On a donc tout intérêt, si jamais vous voulez montrer que vous avez un espace vectoriel, à montrer qu'il s'agit d'un sous-espace vectoriel d'un espace standard.

Par exemple, si vous voulez montrer que l'ensemble des fonctions continues est un espace vectoriel de $\mathbb{R}$ dans $\mathbb{R}$ (pour + et . standard), vous allez prouver qu'il s'agit d'un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$ (des fonctions de $\mathbb{R}$ dans $\mathbb{R}$). Idem pour les fonctions dérivables...

Vous avez un autre exemple avec les suites linéaires sur $\mathbb{R}$ (il s'agit d'un sous-espace vectoriel de l'espace vectoriel des suites réelles).

Il est raisonnable de se demander si un sous-espace d'un espace de dimension finie est encore de dimension finie.

Lemme 1.5.4. *Soit E un sous-espace de dimension finie n. Alors, tout sous-espace F de E est de dimension finie m, avec $m \leq n$.*

Démonstration. On veut montrer tout d'abord que F est de dimension finie, et là, ce sera ... fini.

Si F est nul, tout va bien. Sinon, je peux trouver un vecteur non nul f_1 dans F. Si (f_1) est génératrice de F, c'est une base de F, sinon, on peut trouver un vecteur $f_2 \in F$ qui n'est pas combinaison linéaire de (f_1), et donc (f_1, f_2) est libre dans F. Par récurrence, on obtient une famille $(f_1, f_2, \cdots, f_k)$ libre dans F, et si celle-ci n'est pas génératrice de F,

on construit un vecteur f_{k+1} dans F qui n'est pas combinaison linéaire de $(f_1, f_2, \cdots, f_k)$, et donc, tel que $(f_1, f_2, \cdots, f_k, f_{k+1})$ est libre. Mais ce joli processus ne peut continuer à l'infini.

En effet, comme F est dans E de dimension finie n, et qu'une partie libre de F est forcément aussi une partie libre de E, alors toute partie libre de F est de cardinal inférieur à n. Donc, F possède une base finie (une partie libre maximale obtenue par ce procédé) et cette base est de cardinal m inférieur à n. ◇

Voici un exemple de comment la dimension joue son rôle dans le monde des sous-espaces vectoriels en dimension finie (inspiré par le rôle du cardinal dans les ensembles de cardinal fini).

Proposition 1.5.5. *Soit $F \subset F'$ deux sous-espaces vectoriel d'un $\mathbb{K}$-espace vectoriel E. Alors $\dim F = \dim F'$ si et seulement si $F = F'$.*

Démonstration. L'implication réciproque est claire. Montrons l'implication directe.

On choisit une base de F. Il s'agit d'une partie libre, et qui plus est, elle se trouve dans F'. Comme son cardinal est égal à la dimension de F (par définition) et donc, à la dimension de F' (par hypothèse), c'est une base de F', par le corollaire 1.3.5. ◇

Notation 1.5.6. Si $(u_i)_{i\in I}$ est une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E, on note $\langle u_i\rangle_{i\in I}$ le sous-espace vectoriel de E engendré par les u_i, c'est-à-dire le plus petit sous-espace vectoriel de E qui les contient tous. L'ensemble $\langle u_i\rangle_{i\in I}$ est donc l'ensemble des $\mathbb{K}$-combinaisons linéaires (forcément finies !) des u_i, $i \in I$.

1.6 Construction de base

Voici une petite variante constructive du théorème de la base incomplète. Elle nous montre quels sont les choix qui se posent à nous lorsqu'on veut construire des bases d'un espace vectoriel E de dimension finie n.

i. On choisit un vecteur libre e_1 (c'est-à-dire non nul) de E,

ii. on choisit un vecteur e_2 de E tel que (e_1, e_2) est libre, c'est-à-dire un vecteur de E qui n'est pas dans $\langle e_1\rangle$,

iii. on choisit un vecteur e_3 de E tel que (e_1, e_2, e_3) est libre, c'est-à-dire un vecteur de E qui n'est pas dans $\langle e_1, e_2\rangle$,

iv. par récurrence, on choisit, pour tout k, $1 \leq k \leq n$, un vecteur e_k de E tel que $(e_1, e_2, \cdots, e_k)$ est libre, c'est-à-dire un vecteur de E qui n'est pas dans $\langle e_1, \cdots, e_{k-1}\rangle$,

v. pour $k = n$, on a obtenu une famille libre à n éléments : c'est donc bien une base de E.

Réciproquement, toute base de E peut être construite ainsi, puisque si $(e_1, \cdots, e_n)$ est une base, alors $(e_1, \cdots, e_k)$ est libre pour tout k de 1 à n.

Ceci va nous permettre de compter les bases d'un espace vectoriel E de dimension n sur un corps fini $\mathbb{K}$.

Proposition 1.6.1. *Soit E un espace vectoriel de dimension n sur un corps $\mathbb{K}$ fini, de cardinal q. Alors, son nombre de bases est égal à $(q^n - 1)(q^n - q)\cdots(q^n - q^{n-1})$.*

Démonstration. Tout d'abord, E est de cardinal q^n par la proposition 1.3.6. Le choix de e_1 propose donc $q^n - 1$ possibilités (on retire le vecteur nul).

Quel est le nombre de choix pour e_k, c'est-à-dire, à la k-ième étape ?

On doit choisir le vecteur e_k dans E, mais hors du sous-espace $F_{k-1} := \langle e_1, \cdots, e_{k-1}\rangle$. Or, ce sous-espace est engendré par la famille $(e_1, \cdots, e_{k-1})$, et vu que par construction, cette famille est libre, c'est une base de F_{k-1} (libre et génératrice !).

Conclusion, $\dim F_{k-1} = k-1$ et donc son cardinal est $|F_{k-1}| = q^{k-1}$. Le choix de e_k correspond donc à $q^n - q^{k-1}$ possibilités.

Comme tous ces choix sont indépendants, on multiplie pour obtenir le nombre total de bases, pour obtenir comme attendu exactement $(q^n-1)(q^n-q^1)\cdots(q^n-q^{n-1})$ bases. ◇

1.7 Opérations sur les sous-espaces vectoriels

Dans un espace vectoriel, une définition n'arrive jamais seule, mais toujours par deux ! Ainsi, l'ensemble des sous-espaces vectoriels d'un espace vectoriel E fixé (de dimension quelconque) est muni de deux opérations : l'intersection et l'addition (mais attention, la réunion n'est pas une opération dans les sous-espaces vectoriels.)

Définition 1.7.1. Soit F, F' deux sous-espaces vectoriel de l'espace vectoriel E, alors $F + F'$ est l'ensemble des éléments de E qui s'écrivent comme somme d'un élément de F et d'un élément de F'.

On peut l'écrire ainsi :

$$F + F' := \{u + u',\ u \in F,\ u' \in F'\}$$

On définit bien des opérations puisque :

Proposition 1.7.2. *Soit F, F' deux sous-espaces vectoriel de l'espace vectoriel E, alors $F + F'$ et $F \cap F'$ sont des sous-espaces vectoriels de E.*

Démonstration. Comme F et F' sont des sous-espaces vectoriel, ils contiennent le vecteur nul donc l'intersection est non vide. De plus, ils sont tous deux stables par combinaison linéaire et donc l'intersection l'est aussi.

Notons que $F + F'$ contient le vecteur nul, donc est non vide. Il reste à montrer qu'il est stable par combinaison linéaire. On prend $u, v \in F$, $u', v' \in F'$ et on veut montrer $\lambda(u + u') + \mu(v + v') \in F + F'$ pour tout scalaire λ, μ. C'est clair car

$$\lambda(u + u') + \mu(v + v') = (\lambda u + \mu v) + (\lambda u' + \mu v').$$

◇

Notez que $F + F = F$, mais cela ne prouve pas pour autant que $1 + 1 = 2$.

Remarque 1.7.3. On peut caractériser utilement ces opérations.

L'intersection $F \cap F'$ est le plus grand sous-espace vectoriel de E à la fois inclus dans F et F'.

La somme de $F + F'$ est le plus petit sous-espace vectoriel de E contenant à la fois F et F'.

On peut voir des analogies entre $F \cap F'$ et $F + F'$ avec les pgcd et ppcm.

Remarque 1.7.4. On pourra remarquer que l'intersection est commutative, associative, elle a un élément neutre E et un élément absorbant $\{0\}$.

De façon duale, l'addition est commutative, associative, elle a un élément neutre $\{0\}$ et un élément absorbant E.

Il est facile de voir que si F_i est une famille finie de sous-espaces vectoriels, alors $\dim \sum_i F_i \leq \sum_i \dim F_i$. En effet, si on fixe une base pour chaque F_i et que l'on en prend la réunion, on obtient une famille forcément génératrice de $\sum_i F_i$. Donc, elle est de cardinal plus grand que $\dim \sum_i F_i$.

Il en résulte que pour deux sous-espaces F et F', $\dim F + \dim F' - \dim(F + F') \geq 0$. La formule de Grassmann vient préciser cette inégalité.

Théorème 1.7.5. *Formule de Grassmann Soit F, F' deux sous-espace vectoriels d'un espace vectoriel E de dimension finie. Alors,*

$$\dim F + \dim F' = \dim(F \cap F') + \dim(F + F')$$

Démonstration. On va en fait montrer $\dim(F + F') = \dim F + \dim F' - \dim(F \cap F')$ en fournissant une base de $F + F'$ de cardinal $\dim F + \dim F' - \dim(F \cap F')$.

Soit $m = \dim F$, $m' = \dim F'$. Soit $k = \dim F \cap F'$ et $(e_1, \cdots, e_k)$ une base de $F \cap F'$. La famille $(e_1, \cdots, e_k)$ est donc une partie libre de F et on peut la compléter en une base $(e_1, \cdots, e_k, f_{k+1}, \cdots, f_m)$ de F, et de même on obtient une base $(e_1, \cdots, e_k, f'_{k+1}, \cdots, f'_{m'})$ de F'.

Montrons que $(e_1, \cdots, e_k, f_{k+1}, \cdots, f_m, f'_{k+1}, \cdots, f'_{m'})$ est une base de $F + F'$. Comme elle possède le bon cardinal, cela concluera notre affaire.

Elle est clairement génératrice puisque tout élément de $F + F'$ se décompose[1] en somme d'un élément de F et d'un élément de F'.

Montrons qu'elle est libre, pour cela, on écrit

$$\lambda_1 e_1 + \cdots + \lambda_k e_k + \mu_{k+1} f_{k+1} + \cdots + \mu_m f_m + \mu'_{k+1} f'_{k+1} + \cdots + \mu'_{m'} f'_{m'} = 0.$$

Ceci implique

$$\mu'_{k+1} f'_{k+1} + \cdots + \mu'_{m'} f'_{m'} = -\lambda_1 e_1 - \cdots - \lambda_k e_k - \mu_{k+1} f_{k+1} - \cdots - \mu_m f_m.$$

Comme le membre de gauche est dans F' et le membre de droite dans F, il vient $\mu'_{k+1} f'_{k+1} + \cdots + \mu_{m'} f_{m'} \in F \cap F'$. Donc, il se décompose dans la base $(e_1, \cdots, e_k)$ de $F \cap F'$. Or, $(e_1, \cdots, e_k, f'_{k+1}, \cdots, f'_{m'})$ est libre (c'est une base de F'), donc la relation est forcément triviale. On en déduit que tous les μ'_i sont nuls.

Il vient alors $\lambda_1 e_1 + \cdots + \lambda_k e_k + \mu_{k+1} f_{k+1} + \cdots + \mu_m f_m = 0$. Donc tous les coefficients sont nuls puisque $(e_1, \cdots, e_k, f_{k+1}, \cdots, f_m)$ est libre (c'est une base de F). ◇

Remarque 1.7.6. On se sert souvent de la formule de Grassmann pour montrer que si $\dim F + \dim F' > \dim E$, alors $F \cap F'$ n'est pas réduit au vecteur nul. En effet, on a alors

$$\dim F \cap F' = \dim F + \dim F' - \dim(F + F') > \dim F + \dim F' - \dim E > 0.$$

1. Notez qu'en général, la réunion d'une famille génératrice de F et d'une famille génératrice de F' donne une famille génératrice de $F + F'$.

Remarque 1.7.7. On pourra admirer au passage l'autodualité de cette formule : si l'on remplace la dimension par la codimension et si on échange + et ∩, on retombe sur la même formule.

Remarque 1.7.8. La formule de Grassmann montre une analogie entre les espaces vectoriels et la théorie des ensembles. On voit naturellement la formule de Grassmann comme l'analogue de la formule des cardinaux

$$|A| + |B| = |A \cap B| + |A \cup B|\,.$$

On a vu dans sa preuve que derrière la formule de Grassmann se cache en effet cette formule ensembliste au niveau des bases.

Toutefois, on aurait tort de généraliser le résultat pour tout sous-espace en imaginant une formule du crible adaptée au espaces vectoriels. Prenons encore l'exemple de trois droits vectorielles F, G, H dans le plan E. De sorte que $\dim(F+G+H) = 2$, alors qu'une formule du crible donnerait

$$\dim F + \dim G + \dim H - \dim F \cap G - \dim G \cap H - \dim H \cap F + \dim F \cap G \cap H = 3$$

1.8 Sommes directe et supplémentaires

Dans le contexte précédent, on voudrait qu'un élément w de $F + F'$ s'écrive *de façon unique* sous la forme $w = u + u'$, avec $u \in F$, $u' \in F'$. On dira dans ce cas que F et F' sont en somme directe, et on écrira $F + F' = F \oplus F'$. Il est équivalent de dire que l'unique décomposition du vecteur nul 0 sur $F + F'$ est $0 = 0 + 0$.

Attention, $\oplus$ est une notation qui précise une propriété de la somme, mais ce n'est en aucun cas une opération.

De la même manière, on écrira $F_1 + F_2 + \cdots + F_k = F_1 \oplus F_2 \oplus \cdots \oplus F_k$ si tout élément w de $F_1 + F_2 + \cdots + F_k$ se décompose de façon unique sous la forme $w = \sum_i u_i$, $u_i \in F_i$.

Notons que l'on a l'associativité de $\oplus$ (si on a $(F_1 \oplus F_2) \oplus F_3$, alors on a $F_1 \oplus (F_2 \oplus F_3)$).

Proposition 1.8.1. *En dimension finie, deux sous-espaces F et F' sont en somme directe si et seulement si en concaténant une base de F et une base de F', on obtient une base de $F + F'$.*

Démonstration. On a déjà vu qu'en concaténant deux familles génératrices de F et de F', on obtenait une famille génératrice de $F + F'$.

On voit facilement que si un élément v de $F + F'$ se décompose de deux manières différentes, alors en concaténant deux bases de F et F', on obtient forcément une famille liée. Et la réciproque est tout aussi facile à voir. ◇

Tout ceci peut se voir en dimension quelconque, mais en dimension finie, on peut se servir de la formule de Grassmann pour obtenir :

Corollaire 1.8.2. *Les assertions suivantes sont équivalentes pour deux sous-espaces F et F' :*

i. F *et* F' *sont en somme directe*

ii. $\dim F + F' = \dim F + \dim F'$

iii. $F \cap F' = \{0\}$

Attention à toute généralisation abusive de ceci pour $k > 2$ sous-espaces. Un contre-exemple simple à bien avoir en tête : prenons, pour F_1, F_2, F_3 trois droites vectorielles dans le plan réel. Je vous laisse vérifier que $F_i \cap F_j = \{0\}$, dès que $i \neq j$, et pourtant la somme $F_1 + F_2 + F_3$ est loin d'être directe, sinon un triplet (u_1, u_2, u_3), où u_i est vecteurs directeurs de F_i, serait une famille libre de $\mathbb{R}^2$, absurde.

Voici comment modifier le corollaire précédent quand on a plusieurs sous-espaces.

Corollaire 1.8.3. *Les assertions suivantes sont équivalentes pour une famille de sous-espaces* F_i, $1 \leq i \leq k$ *:*

i. *les* F_i *sont en somme directe*

ii. $\sum_i \dim F_i = \dim \sum_i F_i$

iii. *pour tout* j, $F_j \cap \sum_{i \neq j} F_i = \{0\}$

Démonstration. Essayez de vous convaincre que

i. l'unicité de la décomposition d'un élément en somme d'éléments de F_i revient à dire que la concaténation de bases des F_i est une base de $\sum_i F_i$ (déjà-vu)

ii. l'unicité de la décomposition d'un élément en somme d'éléments de F_i revient à dire qu'un élément de F_j ne peut pas s'écrire dans $\sum_{i \neq j} F_i$.

◇

1.9 "Un" sous-espace supplémentaire

On insiste ici sur le "un". La tentation est trop grande quand on débute de dire "le" supplémentaire à cause de la référence à "le" complémentaire dans la théorie des ensembles. Il faut d'abord voir que le complémentaire d'un sous-espace n'est jamais un sous-espace puisqu'il ne saurait contenir le neutre !

Définition 1.9.1. Soit F un sous-espace vectoriel de E. On dit que F' est un supplémentaire de F dans E si $F \oplus F' = E$.

Proposition 1.9.2. *Soit* F *un sous-espace vectoriel de dimension* k *de* E*, avec* $\dim E = n$*. Les conditions suivantes sont équivalentes*

i. F' *est un supplémentaire de* F *dans* E*,*

ii. $F \cap F' = \{0\}$ *et* $\dim F' = n - k$*,*

iii. $F + F' = E$ *et* $\dim F' = n - k$*,*

iv. *il existe une base de* F *complétée par une base de* F' *en une base de* E

v. *toute base de* F *concaténée avec toute base de* F' *donne une base de* E*.*

Démonstration. Les trois premières assertions sont équivalentes par le corollaire 1.8.2.

S'il existe une base $(e_1, \cdots, e_k, e_{k+1}, \cdots, e_n)$ de E où $(e_1, \cdots, e_k)$ est une base de F et $(e_{k+1}, \cdots, e_n)$ est une base de F', alors tout élément de E se décompose dans cette base, et donc se décompose dans $F + F'$, ainsi $E = F + F'$, et on a une somme directe par le critère des dimensions (encore le corollaire 1.8.2. Donc 4 implique 1.

Supposons que F est F' sont supplémentaires dans E. Si on choisit une base $\underline{e}$ de E concaténée avec une base quelconque de F et une base quelconque de F', alors, tout u de E se décompose en un élément de F et un élément de F', ce qui fait que u se décompose dans $\underline{e}$. Ainsi, $\underline{e}$ est génératrice et donc c'est une base par cardinalité. On vient donc de montrer que 1 implique 5.

Maintenant que 5 implique 4 (clairement) donc implique 1. On a achevé la preuve.

◇

Chapitre 2

Applications linéaires

2.1 Définitions

Soit E et F deux espaces vectoriels sur un corps $\mathbb{K}$, de dimension respective n et m.

Définition 2.1.1. $f : E \to F$ est une application linéaire si :

i. $f(u+v) = f(u) + f(v)$, pour tous $u, v \in E$

ii. $f(\lambda u) = \lambda f(u)$, pour tous $\lambda \in \mathbb{K}$ et tous $u \in E$.

Définition 2.1.2. Une application linéaire de E dans F est également appelé morphisme de E dans F. Si elle est bijective, on dit que c'est un isomorphisme [1]. Il s'agit là de termes génériques pour désigner en algèbre une application compatible avec des structures (groupes, anneaux, corps, $\mathbb{K}$-algèbres) qui, dans le contexte, ne sont plus à spécifier.

Exemple 2.1.3 (Exemple fondamental).

Une application linéaire à bien connaître est l'application dite *application coordonnée*, dont on doit bien réaliser qu'elle dépend du choix d'une base.

On fixe $\underline{e} = (e_1, \cdots, e_n)$ une base de E, et on pose $\varphi_{\underline{e}} : \begin{array}{ccc} E & \to & \mathbb{K}^n \\ u & \mapsto & (u_1, \cdots, u_n) \end{array}$,

où $(u_1, \cdots, u_n)$ sont les coordonnées de u dans la base $\underline{e}$; autrement dit, $u = \sum_{i=1}^{n} u_i e_i$.

L'application φ_e est bien définie car la famille $\underline{e}$ est une base et on a donc l'existence et l'unicité du n-uplet (u_i) associé à u. De plus, il est facile de voir que φ_e est linéaire.

De plus, elle est bijective, d'inverse $(u_i)_{1 \leq i \leq n} \mapsto \sum_{i=1}^{n} u_i e_i$.

Remarque 2.1.4. Une grande spécificité des espaces vectoriels (particulièrement en dimension finie) est que l'on peut construire des bases, et donc, des isomorphismes en l'espace E avec une espace de référence $\mathbb{K}^n$ qui va servir d'outil de calcul. Cela est d'autant plus fort que les espaces vectoriels vont s'immiscer dans toutes les mathématiques grâce à leur grande capacité à approximer les courbes, surfaces, et plus généralement, ce que l'on appelle les variétés différentielles.

Proposition 2.1.5. *Soit $f : E \to F$ une application (pas forcément linéaire !). Alors, en fixant une base de E et une base de F, on obtient des isomorphismes $E \simeq \mathbb{K}^n$, et*

1. Montrez en exercice que l'application réciproque d'un morphisme bijectif est un morphisme.

$F \simeq \mathbb{K}^m$, qui permettent de définir une fonction $\tilde{f} : \mathbb{K}^n \to \mathbb{K}^m$, à travers ce diagramme commutatif :

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ \downarrow\simeq & & \downarrow\simeq \\ \mathbb{K}^n & \underset{\tilde{f}}{\to} & \mathbb{K}^m \end{array}$$

L'application $\tilde{f}$ permet de caractériser f avec des coordonnées.

Exemple 2.1.6. Pour $m = 2$ et $n = 3$, on donne $\tilde{f}(x_1, x_2, x_3) = (x_1 + x_2, x_1x_2)$ (qui, au passage, n'est pas linéaire !).

Conséquence : on en tire une nouvelle caractérisation de la linéarité de f.

Proposition 2.1.7. *Une application f est linéaire si et seulement si $\tilde{f}$ est linéaire. En particulier, si les coordonnées de $\tilde{f}$ sont des fonctions polynômes (en plusieurs variables) homogènes de degré 1, alors f est linéaire.*

Définition 2.1.8. Un polynôme à plusieurs variables est dit homogène de degré k si la somme des degrés de ses monômes est constante et vaut k.

Remarque 2.1.9. Sur un corps $\mathbb{K}$ infini un polynôme est totalement caractérisé par sa fonction polynôme associée et donc, le critère du degré est nécessaire et suffisant ; c'est donc une caractérisation des applications linéaires. Si cette caractérisation est souvent bien pratique, cela ne veut pas dire qu'il faut utiliser exclusivement cette propriété/définition pour prouver qu'une application est linéaire !

Exemple 2.1.10.

i. $\varphi : \begin{array}{ccc} \mathbb{R}[X] & \to & \mathbb{R} \\ P & \mapsto & \int_a^b P(t)\mathrm{d}t \end{array}$

φ est bien linéaire, puisque $\varphi(\lambda P) = \lambda\varphi(P)$, et $\varphi(P + Q) = \varphi(P) + \varphi(Q)$.

Ici, pas besoin des coordonnées, la définition est plus pratique !

ii. Tr : $\begin{array}{ccc} \mathcal{M}_n(\mathbb{K}) & \to & \mathbb{K} \\ (a_{i,j})_{i,j} & \mapsto & \sum_{i=1}^n a_{i,i} \end{array}$ Ici, pour montrer la linéarité de la trace (qui est caractérisée par des coordonnées), il suffit de dire que c'est une application homogène de degré 1.

Remarque 2.1.11. On l'aura bien compris avec la proposition 0.1.3 et l'exemple 0.1.8, les coordonnées permettent d'introduire un nouvel outil de calculs : les matrices (voir plus tard).

2.2 Opérations sur les applications linéaires et théorème fondamental

2.2.1 Opérations

— **Espace $\mathcal{L}(E, F)$**

Pour commencer, l'ensemble des fonctions linéaires de E dans F, noté $\mathcal{L}(E,F)$, est un espace vectoriel pour l'addition et le produit externe donnés par l'évaluation en x : $(\varphi+\psi)(x)=\varphi(x)+\psi(x)$ et $(\lambda\varphi)(x)=\lambda\varphi(x)$.
Notez au passage que cette structure d'espace vectoriel sur $\mathcal{L}(E,F)$ est construite pour que l'évaluation $\mathrm{ev}_x : \mathcal{L}(E,F)\to F$, $\varphi\mapsto\varphi(x)$ en un vecteur fixé x soit une application linéaire.
Pour démontrer que l'on a bien un espace vectoriel, on montre que c'est un sous-espace vectoriel de l'espace vectoriel des fonctions de E dans F, noté $\mathcal{F}(E,F)$. On laisse la preuve au lecteur.

— **Composée de deux applications linéaires**
Dans cette partie, on note également G un espace vectoriel.
On considère l'application de composition :

$$\begin{array}{ccccccc} \mathcal{L}(F,G) & \times & \mathcal{L}(E,F) & \longrightarrow & \mathcal{L}(E,G) \\ (\varphi & , & \psi) & \mapsto & \varphi\circ\psi \end{array}$$

Schématiquement, on peut le voir (attention au strabisme !) comme :

$$\begin{array}{ccccccc} E & \xrightarrow{\psi} & F & \xrightarrow{\varphi} & G \\ e & \mapsto & \psi(e) & \mapsto & \varphi(\psi(e)) =: \varphi\circ\psi(e) \end{array}$$

On vérifie tout de suite que la linéarité de φ et de ψ entraîne celle de $\varphi\circ\psi$.

On note à présent que, si $E=F=G$, la loi « $\circ$ » devient une opération interne ; auquel cas, on note $\mathcal{L}(E,E)=:\mathrm{End}(E)$.

Définition 2.2.1. Une algèbre $(\mathcal{A},+,\cdot,\times)$ sur un corps $\mathbb{K}$ est un ensemble qui vérifie les propriétés suivantes :
• $(\mathcal{A},+,\cdot)$ est un espace vectoriel sur $\mathbb{K}$;
• l'opération $\times$ est interne sur $\mathcal{A}$, et telle que :
• $(A,+,\times)$ est un anneau unitaire ;
• pour tout $k\in\mathbb{K}$, et pour tous $x,y\in A$, $k.(x\times y)=(k.x)\times y$.

Proposition 2.2.2. $(\mathrm{End}(E),+,\cdot,\circ)$ *est une algèbre.*

Démonstration. On vérifie en effet que :
— $(\mathrm{End}(E),+,\dot{})$ est un espace vectoriel ;
— $\circ$ est une loi interne sur $\mathrm{End}(E)$, compatible avec les autres lois, avec unité l'application identité de E, notée Id_E.

◇

Remarque 2.2.3. On identifie alors un neutre pour la loi $\circ$: $\begin{array}{rccc} \mathrm{Id}_E : & E & \to & E \\ & u & \mapsto & u \end{array}$

— **Annonce pour la suite**
Grâce à l'application $\tilde{f}:\mathbb{K}^n\to\mathbb{K}^m$, on aura $\mathcal{L}(E,F)\simeq\mathcal{M}_{m,n}(\mathbb{K})$. C'est-à-dire que l'on va pouvoir mettre un objet de calcul (les matrices) sur $\mathcal{L}(E,F)$.

2.2.2 Théorème fondamental

Voici le théorème qui servira tout le temps par la suite, en particulier lorsque l'on assimilera plus tard l'espace de applications linéaires de E dans F à un espace de matrices (outil de calcul).

Théorème 2.2.4. *Une application linéaire $\varphi \in \mathcal{L}(E,F)$ est entièrement déterminée, et de manière unique, par l'image d'une base de E.*

Autrement dit : pour $(e_i)_{i\in I}$ base de E et $(f_j)_{j\in J}$ famille de F, il existe un unique $\varphi \in \mathcal{L}(E,F)$ tel que $\varphi(e_i) = f_i$ pour tout $i \in I$ $(\star)$.

Démonstration.

— Unicité : supposons qu'il existe φ_1 et φ_2 qui vérifie $(\star)$.
Alors, pour tout $i \in I$, $\varphi_1(e_i) = f_i = \varphi_2(e_i)$.
On en déduit que φ_1 et φ_2 coïncident sur une base, donc, pour tout $x = \sum_{i=1}^{n} x_i e_i \in E$, par linéarité des applications,

$$\varphi_1(x) = \varphi_1\left(\sum_{i\in I} x_i e_i\right) = \sum_{i\in I} x_i \varphi_1(e_i) = \sum_{i\in I} x_i \varphi_2(e_i) = \varphi_2(x)$$

Ainsi, $\varphi_1 = \varphi_2$.

— Existence : soit $(f_j)_j$ une famille de F. Soit alors φ l'application définie comme suit : $\varphi : x = \sum_{i=1}^{n} x_i e_i \mapsto \sum_{i\in I} x_i f_i$.
Tout d'abord, il faut montrer que φ est bien définie. Comme $(e_i)_i$ est une base de E, c'est une famille libre, et donc, les $(x_i)_i$ sont définis de manière uniques ; cela montre que φ est bien définie.
Ensuite, notons que $\varphi : E \to F$ est linéaire ; c'est clair.
Enfin, il faut tout de même montrer que φ vérifie $\varphi(e_i) = f_i$, pour tout i.
Or, $\varphi(e_i) = \sum_{j\in I} x_j f_j = x_i f_i + \sum_{j\neq i} x_j f_j = 1.f_i + 0 = f_i$.

◇

2.3 Injectivité et surjectivité

Nous allons commencer par dire quelques mots sur l'injectivité et la surjectivité en matière de morphismes. Tout d'abord quelques rappels. Encore une fois injectivité et surjectivité viennent par deux, cette fois-ci pour se rejoindre dans la notion de "bijectivité".

Définition 2.3.1. Soit $f : E \to F$ une application.

i. f est une application injective si : $f(x) = f(y) \Longrightarrow x = y$ *(autrement dit, un élément de F a au plus un antécédent)* ;

ii. f est une application surjective si : pour tout $y \in F$, il existe un $x \in E$ tel que $y = f(x)$ *(autrement dit, un élément de F a au moins un antécédent)*

Proposition 2.3.2. *Si $f \in \mathcal{L}(E,F)$, c'est-à-dire si f est linéaire, alors f est injective si, et seulement si, $\{x \in E;\ f(x) = 0\} = \{0\}$.*

$\{x \in E;\ f(x) = 0\}$ est appelé le noyau de f, et il est noté $\mathrm{Ker}(f)$.

Démonstration. Montrons l'implication directe. Soit $x \in E$ tel que $f(x) = 0$. Or, comme f est linéaire, on sait aussi que $f(0) = 0$. Donc on a $f(x) = f(0)$; comme f est injective, on en déduit que $x = 0$.

Réciproquement, si $f(x) = f(y)$, alors $f(x) - f(y) = 0$, et comme f est linéaire, cela implique $f(x-y) = 0$. Par hypothèse, on en déduit alors que $x - y = 0$, et donc, $x = y$. ⋄

Remarque 2.3.3. Le noyau est donc en lien avec l'injectivité. La preuve montre quelque chose de plus constructif : l'ensemble des x tels que $f(x) = a$, où a est fixé dans F est soit l'ensemble vide, soit de la forme $x_0+\mathrm{Ker}(f) := \{x_0+u,\ u \in \mathrm{Ker}(f)\}$, où u_0 est une *solution particulière* de $f(x) = a$. Il s'agit là d'un résultat très général en mathématiques, que l'on retrouve dans les systèmes linéaires, les équations différentielles linéaires, ce qui nous amènera à introduire un peu plus tard les espaces affines. Mais c'est surtout la première annonce de ce que l'on appelle (dans les milieux autorisés) l'isomorphisme canonique $E/\mathrm{Ker}(f) \simeq \mathrm{Im}(f)$.

Proposition 2.3.4. *Soit $\varphi \in \mathcal{L}(E, F)$.*

i. Si φ est injective, et si $(e_i)_i$ est une famille libre dans E, alors la famille $(\varphi(e_i))_i$ est également libre dans F ;

ii. Si φ est surjective, et si $(e_i)_i$ est génératrice de E, alors la famille $(\varphi(e_i))_i$ est également génératrice de F.

iii. S'il existe une base $(e_i)_i$ de E telle que la famille $(\varphi(e_i))_i$ est libre dans F, alors φ est injective.

iv. S'il existe une base $(e_i)_i$ de E telle que la famille $(\varphi(e_i))_i$ est génératrice dans F, alors φ est surjective.

Démonstration. Encore une fois, le mot-clef est la linéarité !

i. Prenons une famille de scalaires $(\lambda_i)_i$ tels que $\sum_i \lambda_i \varphi(e_i) = 0$. Cela implique, par linéarité, $\varphi\left(\sum_i \lambda_i e_i\right) = 0$. Comme φ est supposée injective, on obtient $\sum_i \lambda_i e_i = 0 \Longrightarrow \lambda_i = 0$ pour tout i, puisque la famille $(e_i)_i$ est supposée libre.

ii. La seconde démonstration repose exactement sur le même schéma.

iii. On veut montrer que φ est injective, c'est-à-dire que $\varphi(x) = 0$ implique $x = 0$. On écrit $x = \sum_i x_i e_i$ dans la base, et $\varphi(x) = 0$ implique $\sum_i x_i \varphi(e_i) = 0$, par linéarité de φ. Or, $(\varphi(e_i))_i$ est libre et donc tous les x_i sont nuls, donc $x = 0$.

iv. Comme $(\varphi(e_i))_i$ est génératrice, tout y de F se décompose en $y = \sum_i y_i \varphi(e_i)$ et on voit facilement que $x := \sum_i y_i e_i$ est un antécédent de y.

⋄

Remarque 2.3.5. En termes de dimensions des espaces E et F, on voit alors ce qu'implique l'existence d'un morphisme injectif ou surjectif :

i. Si f est injective, alors $\dim(E) \leq \dim(F)$;

ii. Si f est surjective, alors $\dim(E) \geq \dim(F)$;

iii. Si f est bijective, alors $\dim(E) = \dim(F)$.

Corollaire 2.3.6. *$\varphi \in \mathcal{L}(E,F)$ est un isomorphisme si et seulement si φ envoie une base (fixée) $(e_i)_i$ sur une base $(f_i)_i$ si et seulement si pour toute base, son image par φ est une base.*

Notons que jusqu'à présent, à partir d'un morphisme φ, on a regardé l'image d'une famille. Réciproquement, si on se fixe une famille $(e_i)_i$ de E et une famille $(f_i)_i$ de F, existe-t-il un morphisme φ qui envoie e_i sur f_i pour tout i ?

La réponse est non en général : par exemple si $e_1 = -e_2$, la linéarité exige $f_1 = -f_2$ pour pouvoir espérer l'existence de $\varphi \in \mathcal{L}(E,F)$ tel que $\varphi(e_i) = f_i$, $i = 1, 2$. Vous l'aurez compris, l'obstruction réside dans la relation linéaire. Si l'on impose que la famille (e_i) est libre, alors, existence sera assurée. Ce théorème précise le théorème fondamental 2.2.4.

Théorème 2.3.7. *Soit $(e_i)_{i\in I}$ une famille libre de E, et $(f_i)_{i\in I}$ une famille quelconque de F. Alors, il existe $\varphi \in \mathcal{L}(E,F)$ tel que $\varphi(e_i) = f_i$. Si de plus $(e_i)_{i\in I}$ est une base, alors $\varphi \in \mathcal{L}(E,F)$ existe et est unique.*

Démonstration. D'après le théorème de la base incomplète, l'existence de φ peut se ramener au cas où $(e_i)_{i\in I}$ est une base (quitte à compléter les f_i de façon quelconque). Supposons donc que ce soit le cas. On définit donc pour tout x de E, l'image $\varphi(x) = \sum_{i\in I} x_i f_i$, avec $x = \sum_{i\in I} x_i e_i$. L'application φ est bien définie puisque les x_i sont uniquement déterminés par x. On vérifie que φ est bien linéaire et que $\varphi(e_i) = f_i$.

Montrons l'unicité. Soit φ' est une application linéaire telle que $\varphi'(e_i) = \varphi(e_i)$. Alors, comme (e_i) est une base, tout x peut se décomposer en $x = \sum_i x_i e_i$, et

$$\varphi'(x) = \sum_i x_i \varphi'(e_i) = \sum_i x_i f_i = \varphi(x).$$

D'où l'unicité.

◇

Remarque 2.3.8. Le lecteur très perspicace aura noté que l'existence de φ repose sur le fait que (e_i) est libre alors que l'unicité de φ repose sur le fait que (e_i) est génératrice. Encore un coup de la dualité (attention spoiler !).

Exemple 2.3.9. Si l'on suppose ici que $\mathbb{K}$ un corps fini de cardinal q, combien y a-t-il d'isomorphismes entre E et F ? Combien y a-t-il d'applications injectives ?

Réponse : si E et F n'ont pas la même dimension, il n'y a pas l'isomorphismes ! Si $\dim(E) = \dim(F)$, il y en a donc autant que de bases de F : $(q^n - 1)\cdots(q^n - q^{n-1})$. En effet, il suffit de fixer une base (e_i) de E et toute base (f_i) de F fournira un unique isomorphismes φ tel que $\varphi(e_i) = f_i$ pour tout i.

Pour ce qui est des morphismes injectifs, il faut que $m := \dim E \leq \dim F =: n$, sinon, il n'y en a pas. Si l'inégalité a lieu, alors, en fixant encore une fois une base $(e_i)_{1\leq i\leq m}$ de E, on a une bijection entre l'ensemble des applications injectives φ de $\mathcal{L}(E,F)$ et l'ensembles de parties libres à m éléments de F, qui à φ associe $(\varphi(e_i))$, et dont la réciproque est donnée par l'application qui envoie la partie libre $(f_i)_{1\leq i\leq m}$ de F vers l'unique application φ telle que $\varphi(e_i) = f_i$ (cette application est automatiquement injective puisque l'image d'une base est une partie libre). Le nombre cherché est donc $(q^n - 1)\cdots(q^n - q^{m-1})$.

C'est la dualité qui va nous permettre de calculer le nombre d'applications linéaires surjectives d'un espace vers un autre.

2.4 Formule du rang

Nous allons associer à un morphisme $\varphi \in \mathcal{L}(E, F)$ deux nombres permettant de mesurer respectivement son injectivité et sa surjectivité, il s'agit du noyau et de l'image de φ. Ils sont liés par la formule du rang. Noyaux et images sont encore les paires qui vont se renvoyer l'un à l'autre par la dualité.

Définition 2.4.1. Si f est une application linéaire d'un espace E vers un espace F, on note $\mathrm{Im}(f)$ l'image de E par f, que l'on appelle aussi image directe, $f(E)$, dans la théorie des ensembles. Il est clair que, de la même manière que f est injective si et seulement si $\mathrm{Ker}(f)$ est nul, f est surjective si et seulement si $\mathrm{Im}(f) = F$.

Proposition 2.4.2. *Soit $\varphi \in \mathcal{L}(E, F)$. Alors, $\mathrm{Im}(f)$ est un sous-espace vectoriel de F et $\mathrm{Ker}(f)$ est un sous-espace vectoriel de E.*

Démonstration.

i. C'est tout droit, en utilisant la linéarité de φ. Pour $\mathrm{Im}(\varphi)$ par exemple, on vérifie que :

   — $\mathrm{Im}(\varphi) \subset F$;
   — $\mathrm{Im}(\varphi)$ est non vide (il contient 0 !) ;
   — pour tous y, $y' \in F$, et pour tout $\lambda, \lambda' \in \mathbb{K}$, $\lambda y + \lambda' y' \in F$ (par linéarité !).

◇

Exemple 2.4.3. On veut montrer que l'ensemble des (x, y, z) de $\mathbb{R}^3$ tels que $x - y = 0$ et $2x + 3y - 5z = 0$ est un sous-espace vectoriel de $\mathbb{R}^3$. Le mieux est de dire que c'est le noyau de l'application linéaire φ de $\mathbb{R}^3$ dans $\mathbb{R}^2$ qui envoie (x, y, z) sur $(x-y, 2x+y-5z)$. C'est bien polynomial homogène de degré 1, donc, φ est bien une application linéaire.

Exemple 2.4.4. On veut un système de générateur de $\mathrm{Im}(\varphi)$ où φ est l'application linéaire $(x, y) \mapsto (x-2y, x+3y, -y)$. Comme, par construction, φ est surjective si on prend pour espace d'arrivée $\mathrm{Im}(\varphi)$, un système de générateur est donné par l'image d'une base de $\mathbb{R}^2$. Si on part de la base canonique de $\mathbb{R}^2$, un système de générateur de $\mathrm{Im}(\varphi)$ est donc $(1, 1, 0), (-2, 3, -1)$. Comme c'est une famille clairement libre, c'est une base de $\mathrm{Im}(\varphi)$.

2.4.1 Restriction

Si φ est une application linéaire de E dans E', et si F est un sous-espace vectoriel de E, alors l'application [2] $\varphi\,|_F$ de F dans E', qui envoie x de F sur $\varphi(x)$ est appelée restriction de φ à F.

Proposition 2.4.5. *Si φ est dans $\mathcal{L}(E, E')$, et si F est un sous-espace vectoriel de E, alors $\varphi_{|F}$ est dans $\mathcal{L}(F, E')$. On a*

i. $\mathrm{Ker}(\varphi\,|_F) = F \cap \mathrm{Ker}(\varphi)$. *En particulier $\varphi\,|_F$ est injective si et seulement si F est en somme directe avec* $\mathrm{Ker}(\varphi)$,

2. On ne la note pas φ parce que, en tout état de cause, il ne s'agit pas de la même application puisque son ensemble de départ est différent.

ii. l'application (linéaire) de $\mathcal{L}(E, E')$ dans $\mathcal{L}(F, E')$ qui envoie φ sur $\varphi\mid_F$ est surjective.

Démonstration.

i. La première égalité est claire et le point particulier vient du fait que 1) l'injectivité d'un morphisme est équivalente à l'annulation de son noyau, et 2) être en somme directe revient à dire que l'intersection est nulle.

ii. On veut montrer que si ψ est une application linéaire de F dans E', alors, il existe une application linéaire de E dans E' qui prolonge ψ, c'est-à-dire, dont la restriction à F est ψ.

On peut construire φ avec des bases. En effet : soit $(f_1, \cdots, f_k)$ une base de F qui se prolonge en une base $(f_1, \cdots, f_k, \cdots, f_n)$ de E. On pose $e'_i = \psi(f_i)$ pour i allant de 1 à k, et on prolonge la famille $(e'_i)_{1\leq i\leq k}$ en une famille $(e'_i)_{1\leq i\leq n}$ de E' (sans propriété spécifiée pour cette famille). Alors, on sait par le théorème fondamental qu'il existe une application linéaire φ de E dans F qui envoie $(f_1, \cdots, f_k, \cdots, f_n)$ sur $(e'_1, \cdots, e'_k, \cdots, e'_n)$. Il est clair que $\varphi\mid_F = \psi$ par unicité d'une application linéaire dont l'image d'une base est fixée (puisque $e'_i = \psi(f_i)$, $1 \leq i \leq k$).

◇

Le résultat qui suit est en fait un lemme, mais il a été tellement utilisé dans les cours d'algèbre linéaire, qu'il a rapidement accédé au grade de théorème, pour bons et loyaux services. On l'appelle parfois "théorème du rang", mais dans le cours, on l'évoquera sous le nom de formule du rang pour le distinguer d'un autre théorème du rang.

Définition 2.4.6. Le rang d'une application linéaire f est la dimension de son image. On le note $\mathrm{rg}(f)$.

Théorème 2.4.7. ***(Formule du rang)*** *Soit $\varphi \in \mathcal{L}(E, F)$. Alors on a la formule suivante :*

$$\dim \mathrm{Ker}(f) + \mathrm{rg}(f) = \dim(E)$$

Démonstration. On construit une base $(e_1, \cdots, e_k)$ de $\mathrm{Ker}(f)$, que l'on complète en une base $(e_1, \cdots, e_k, e_{k+1}, \cdots, e_n)$ de E. On considère le sous-espace $S := \langle e_{k+1}, \cdots, e_n\rangle$, de sorte que $\mathrm{Ker}(f) \oplus S = E$. D'après la proposition 2.4.5, $f\mid_S$ est injective.

On va montrer que le morphisme $\tilde{f}$ de S dans $\mathrm{Im}(f)$ qui envoie x sur $f(x)$ est un isomorphisme. On aura alors $n - \dim(\mathrm{Ker}(f) = \dim(S) = \dim \mathrm{Im}(f)$, ce qui fournira la formule du rang.

Il suffit de montrer la surjectivité (on a déjà vu l'injectivité). Soit donc y dans $\mathrm{Im}(f)$, il existe x dans E tel que $f(x) = y$. Soit $x = x_K + x_S$ la décomposition de x dans la somme directe $\mathrm{Ker}(f) \oplus S$, il vient que $y = f(x) = f(x_K + x_S) = f(x_K) + f(x_S) = f(x_S)$. D'où la surjectivité de $\tilde{f}$ puisque l'on a trouvé un antécédent dans S. ◇

Corollaire 2.4.8. *Si $\dim(E) = \dim(F)$, alors on en déduit que φ est injective si, et seulement si, elle est surjective.*

Démonstration. On raisonne par équivalence :

$$\begin{aligned} f \text{ est injective} &\Leftrightarrow \dim \mathrm{Ker}(f) = \{0\} \Leftrightarrow \dim \mathrm{Im}\, f = \dim(E) = \dim(F) \\ &\Leftrightarrow f \text{ est surjective} \end{aligned}$$

◇

Remarque 2.4.9. Avec ce corollaire, on voit ici bien que la dimension remplace la notion de cardinal, en tenant le même rôle ; on rappelle effectivement que pour les ensembles, si un ensemble A est inclus dans un ensemble B, et que les cardinaux sont les mêmes, alors les ensembles sont égaux.

Chapitre 3

Dualité

3.1 Définitions et premiers effets d'annonce

Dans ce chapitre, sauf mention du contraire, E désignera un espace vectoriel *de dimension finie*.

Le dual d'un $\mathbb{K}$-espace vectoriel E est l'espace $\mathcal{L}(E, \mathbb{K})$, c'est-à-dire des applications linéaires de E dans $\mathbb{K}$ qui sont aussi appelées "formes [1] linéaires". C'est donc un cas particulier d'applications linéaires. Pourquoi donc consacrer un chapitre particulier sur $\mathcal{L}(E, F)$ lorsque F est l'espace vectoriel de référence de dimension 1 ?

On en voit deux raisons.

Une première est d'ordre pratique, l'étude d'une application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^m$ se ramène naturellement à l'étude des application linéaires coordonnées qui sont des application linéaires de $\mathbb{K}^n$ dans $\mathbb{K}$. Les formes linéaires sont donc les "particules élémentaires" qui constituent les applications linéaires. Notons que si l'on avait mis $\mathbb{K}$ comme espace de départ dans $\mathcal{L}(E, F)$, on aurait obtenu l'espace $\mathcal{L}(\mathbb{K}, F)$ qui s'assimile canoniquement à l'espace F par le biais de l'isomorphisme $\varphi \mapsto \varphi(1)$ (donc, rien de nouveau sous le soleil !).

Une seconde est d'ordre théorique. La dualité est utilisée comme un "dictionnaire" qui transforme un mot M en un mot dual M^*. Si un théorème T est constitué d'une phrase, on obtient un second théorème gratuit T^* à l'aide de ce dictionnaire. La dualité devient une machine à multiplier par deux les résultats. Le rêve d'un chef d'entreprise ! Voici tout abord (Spoiler Alert !) un tableau qui va résumer le dictionnaire annoncé.

Il y aura ensuite une jolie application aux formes bilinéaires, mais ça, c'est une autre histoire, et chaque chose en son temps.

1. On appelle en général *forme* toute application de E dans $\mathbb{K}$.

Espace E	Espace dual E^*
base $(e_i)_{1\leq i\leq n}$	base duale $(e_i^*)_{1\leq i\leq n}$
matrice de chgmt de base P	matrice de chgmt de base ${}^tP^{-1}$
$x \in E$, $x = \sum_{i=1}^n e_i^*(x)e_i$	$\varphi \in E^*$, $\varphi = \sum_{i=1}^n \varphi(e_i)e_i^*$
sous-espace $F \subset E$	$F^\perp \subset E^*$
$\dim F$	$\mathrm{codim} F^\perp$
$F \subset G$	$G^\perp \subset F^\perp$
droite	hyperplan
hyperplan	droite
$+$	$\cap$
$\cap$	$+$
$f \in \mathcal{L}(E,F)$	${}^tf := ? \circ f \in \mathcal{L}(F^*, E^*)$
ι=inclusion de E dans F	${}^t\iota$=restriction de F^* à E^*
Ker (contrainte)	Im (liberté) $\mathrm{Im}({}^tf) = \mathrm{Ker}(f)^\perp$
Im (liberté)	Ker (contrainte) $\ker({}^tf) = \mathrm{Im}(f)^\perp$
injectif	surjectif
surjectif	injectif

On constate en regardant le dictionnaire qu'il est involutif, c'est-à-dire, que $(M^*)^* = M$. Cela provient du fait qu'en dimension finie, un espace E est canoniquement isomorphe à son bidual E^{**}.

3.2 Base duale

L'ensemble $E^* = \mathcal{L}(E, \mathbb{K})$ est doté d'une structure d'espace vectoriel (c'est un espace d'applications linéaires) de dimension $\dim E^* = \dim(E) \cdot \dim(\mathbb{K}) = n$.

Comme un espace vectoriel E n'arrive pas avec une base particulière, il est naturel de penser qu'il en est de même de l'espace dual E^*. Toutefois, si on munit E d'une base $(e_i)_{1\leq i\leq n}$, alors, on a une base naturelle de E^* appelée base duale de $(e_i)_{1\leq i\leq n}$ et notée $(e_i^*)_{1\leq i\leq n}$. En effet, on a vu que le théorème fondamental de l'algèbre linéaire impliquait qu'une base $(e_i)_{1\leq i\leq n}$ fournit un isomorphisme entre $\mathcal{L}(E,F)$ et F^n, qui envoie un morphisme sur l'image de (e_i). Si on pose $F = \mathbb{K}$, alors, on voit que E^* est isomorphe à $\mathbb{K}^n$ par $(\varphi \mapsto \varphi(e_i))$. La base duale n'est rien autre que la base correspondant à la base canonique de $\mathbb{K}^n$ via cet isomorphisme.

Proposition 3.2.1. *Soit E un espace vectoriel de dimension n et $(e_i)_{1\leq i\leq n}$ une base de E. Alors, il existe une unique base $(e_i^*)_{1\leq i\leq n}$ de E^* définie par $e_i^*(e_j) = \delta_{ij}$, pour tout j, où δ_{ij} désigne le symbole de Kronecker.*

Démonstration. Tout d'abord, par le théorème 2.2.4, l'application linéaire e_i^* est bien définie de façon unique, puisqu'elle est imposée par image d'une base.

Il reste à montrer qu'il s'agit bien d'une base, et comme on en a le bon nombre (n), il suffit de voir que c'est une famille libre. On suppose donc une relation linéaire $\sum_{i=1}^n \lambda_i e_i^* = 0$. Comme on a une égalité entre deux applications linéaires, on peut évaluer

l'égalité en e_j. On obtient alors pour tout j

$$0 = \sum_{i=1}^{n} \lambda_i e_i^*(e_j) = \sum_{i=1}^{n} \lambda_i \delta_{ij} = \lambda_j.$$

Ceci conclut notre affaire. ◇

Contre-exemple 3.2.2. Pour trouver un contre-exemple en dimension infinie, il suffit de regarder l'espace vectoriel des polynômes $\mathbb{K}[X]$ muni de sa base infinie $(e_i)_{i\in\mathbb{N}}$, avec $e_i := X^i$. L'évaluation en 1 définie par $\mathrm{ev}_1(P) = P(1)$ est une forme linéaire sur $\mathbb{K}[X]$. Supposons que $(e_i^*)_{i\in\mathbb{N}}$ soit une famille génératrice (elle est libre, d'après la preuve de la proposition). Alors, on aurait $\mathrm{ev}_1 = \sum_i \lambda_i e_i^*$, avec les λ_i presque tous nuls (par définition, une combinaison linéaire est finie !). Soit j tel que $\lambda_j = 0$. On a alors

$$1 = \mathrm{ev}_1(X^j) = \big(\sum_i \lambda_i e_i^*\big)(e_j) = \lambda_j = 0,$$

ce qui est absurde.

On va voir que la base duale est bien génératrice (pour la dimension finie) par un moyen alternatif bien utile. On peut le dire ainsi : les coordonnées de $x \in E$ dans la base (e_i) sont les $e_i^*(x)$, et les coordonnées de $\varphi \in E^*$ dans la base duale sont les $\varphi(e_i)$.

Proposition 3.2.3. *Soit E un espace vectoriel muni d'une base $(e_i)_{1\le i\le n}$. Soit $x \in E$ un vecteur de E et $\varphi \in E^*$ une forme linéaire sur E. Alors,*

$$x = \sum_{i=1}^{n} e_i^*(x) e_i, \;\; \varphi = \sum_{i=1}^{n} \varphi(e_i) e_i^*.$$

Démonstration. Comme (e_i) est une base, on peut écrire $x = \sum_i \lambda_i e_i$, et ce, de façon unique. On utilisant la linéarité de e_j^*, on obtient, pour tout j,

$$e_j^*(x) = \sum_i \lambda_i e_j^*(e_i) = \sum_i \lambda_i \delta_{ij} = \lambda_j.$$

Ceci assure la première égalité.

Pour montrer l'égalité $\varphi = \sum_{i=1}^n \varphi(e_i) e_i^*$, il suffit de montrer l'égalité appliquée à la base (e_j), par le théorème 2.2.4, puisqu'une application linéaire est entièrement déterminée par l'image d'une base fixée. Or,

$$\big(\sum_{i=1}^{n} \varphi(e_i) e_i^*\big)(e_j) = \sum_{i=1}^{n} \varphi(e_i) e_i^*(e_j) = \sum_{i=1}^{n} \varphi(e_i)\delta_{ij} = \varphi(e_j).$$

Ceci assure la seconde égalité. ◇

Remarque 3.2.4. On doit à la première égalité que les applications e_i^* sont appelées *applications coordonnées*, vu l'égalité $e_i^*(x) = x_i$, où x_i est la i-ème coordonnée de x.

3.3 Changement de base

Une question naturelle arrive : si l'on a une nouvelle base (f_j) de E que l'on sait écrire en fonction de l'ancienne base (e_i), quelles sont les relations entre les bases duales (e_i^*) et (f_j^*) ?

Proposition 3.3.1. *Soit (e_i) et (f_j) deux bases de E. On note $f_j = \sum_i a_{ij} e_i$, avec $a_{ij} \in \mathbb{K}$, $1 \leq i, j \leq n$, les f_j dans la base (e_i). Alors, $e_j^* = \sum_i a_{ji} f_i^*$.*

Démonstration. Encore une fois, on va utiliser le théorème fondamental de l'algèbre linéaire, qui dit qu'une application linéaire est entièrement déterminée par l'image d'une base. On va donc montrer que les deux applications linéaires e_j^* et $\sum_i a_{ji} f_i^*$ coïncident sur la base (f_k) de E.

D'une part,

$$e_j^*(f_k) = e_j^*(\sum_i a_{ik} e_i) = \sum_i a_{ik} \delta_{ij} = a_{jk}.$$

D'autre part,

$$\big(\sum_i a_{ji} f_i^*\big)(f_k) = \sum_i a_{ji} f_i^*(f_k) = \sum_i a_{ji} \delta_{ik} = a_{jk}.$$

Ceci établit notre égalité.

◇

Quand on aura vu les matrices, et en particulier les matrices de passage, on pourra mettre ce résultat sous la forme suivante :

Corollaire 3.3.2. *Avec les notations précédentes, soit P la matrice de passage de la base (e_i) vers la base (f_j), et soit Q la matrice de passage de la base (e_j^*) vers la base (f_i^*). Alors, $Q = {}^t P^{-1}$.*

3.4 Bidual

Le fait que E et E^* sont de même dimension nous dit que E et E^* sont isomorphes. Pour en fournir un isomorphisme, il suffit de prendre une base (e_i) de E et d'envoyer cette base sur sa duale (e_i^*). On construit alors un morphisme [2] α entre E et E^*. Ce morphisme dépend-il de la base (e_i) choisie ? C'est-à-dire, est-ce que si je prends une autre base (f_j) de E, α envoie-t-il (f_j) sur (f_j^*) ? La proposition 3.3.1 nous dit que non. En effet, par l'absurde, si c'était le cas, on aurait

$$f_j^* = \alpha\big(\sum_i a_{ij} e_i\big) = \sum_i a_{ij} e_i^*,$$

et, par la proposition 3.3.1, cela reviendrait à dire que la matrice de changement de base de (e_i) à (f_j) est égale à la matrice de changement de base de (e_i^*) à (f_j^*), ce qui est faux par le corollaire précédent.

2. On verra par la suite que $\mathcal{L}(E, E^*)$ s'identifie naturellement à l'espace des formes bilinéaires sur E, et ceci sera utilisé à bon escient pour l'étude des formes quadratiques sur E.

Si l'on regarde de plus près ce corollaire, on voit que l'idée de dualiser E^* permet de retomber sur nos pattes, dans le sens que si $Q =^t P^{-1}$, alors $P =^t Q^{-1}$. Donc, on peut espérer obtenir un isomorphisme entre E et le dual E^{**} de E^*. C'est l'idée du bidual, mais nous allons le faire avec une approche différente, en nous dispensant du choix d'une base.

Théorème 3.4.1. *On construit une application ι en assignant à x de E une forme linéaire ι_x sur E^* par $\iota_x(\varphi) = \varphi(x)$ pour tout φ de E^*. L'application $x \mapsto \iota_x$ définit un isomorphisme de E dans E^{**}.*

Démonstration. On a bien $\iota_x \in E^{**}$. En effet, l'application ι_x va bien de E^* dans $\mathbb{K}$. Elle est linéaire car

$$\iota_x(\lambda\varphi + \mu\psi) = (\lambda\varphi + \mu\psi)(x) = \lambda\varphi(x) + \mu\psi(x) = \lambda\iota_x(\varphi) + \mu\iota_x(\psi).$$

L'application $x \mapsto \iota_x$ est linéaire. En effet,

$$\iota_{\lambda x+\mu y}(\varphi) = \varphi(\lambda x + \mu y) = \lambda\varphi(x) + \mu\varphi(y) = \lambda\iota_x(\varphi) + \mu\iota_y(\varphi) = (\lambda\iota_x + \mu\iota_y)(\varphi),$$

et donc $\iota_{\lambda x+\mu y} = \lambda\iota_x + \mu\iota_y$.

L'application $x \mapsto \iota_x$ est injective. En effet, on peut regarder son noyau (elle est linéaire). Si x est dans le noyau, alors, pour tout φ de E^*,

$$\varphi(x) = \iota_x(\varphi) = 0(\varphi) = 0.$$

Par l'absurde, si x était non nul, on pourrait poser $e_1 = x$ et le compléter en une base (e_i). On aurait alors, $e_1^*(x) = e_1^*(e_1) = 1$, et il existerait donc une forme linéaire φ telle que $\varphi(x) \neq 0$. Conclusion, $x = 0$ et l'application est bien injective.

L'application est un isomorphisme car elle est injective et $\dim E^{**} = \dim(E)$. ◇

On notera dans la suite ι l'isomorphisme que l'on vient de construire entre E et E^{**}.

Remarque 3.4.2. On commence à mieux comprendre le côté involutif de notre dictionnaire, puisque déjà $(E^*)^*$ s'identifie naturellement à E.

Contre-exemple 3.4.3. Il n'est pas évident, dans un contexte du programme de l'agrégation où l'axiome du choix n'est qu'une vague philosophie, de donner rigoureusement un exemple précis où E n'est pas isomorphe à son bidual. Mais pour rester dans le vague, disons que si E est l'espace des suites réelles nulles presque partout, alors E^* est isomorphe à l'espace des suites réelles (tout court). Résultat des courses, E possède une base dénombrable, E^* possède une base non dénombrable, et E^* s'injecte dans E^{**}. Donc, un isomorphisme entre E et E^{**} est non-négociable car il impliquerait une bijection entre un ensemble dénombrable et un ensemble qui contient un ensemble non dénombrable.

3.5 Sous-espaces et orthogonaux

Définition 3.5.1. Soit F un sous-espace de E. On note

$$F^\perp := \{\varphi \in E^*,\ \varphi(x) = 0,\ \forall x \in F\}$$

Proposition 3.5.2 (Dimension de l'orthogonal)**.** *Pour tout sous-espace F de E, $F^\perp$ est un sous-espace vectoriel de E^* de dimension* $\dim E - \dim F$.

Démonstration. On considère l'application $r_F : E^* \to F^*$, qui envoie la forme linéaire φ de E^* sur sa restriction $r_F(\varphi) := \varphi\,|_F$ sur F. Il est clair que $\varphi\,|_F$ reste linéaire et donc, r_F est bien définie. De plus, r_F est clairement linéaire (écrivez-le pour vous convaincre).

Le noyau de r_F est l'ensemble des φ de E^* tels que $\varphi\,|_F$ est nul ; c'est donc $F^\perp$. Il en résulte que $F^\perp$ est bien un sous-espace vectoriel de E^*.

L'application r_F est surjective par la proposition 2.4.5. Il en résulte par la formule du rang que $\dim \operatorname{Ker} r_F + \dim F^* = \dim E^*$, d'où l'égalité $\dim F^\perp = \dim E - \dim F$. ◇

Définition 3.5.3. Si Φ est un sous-espace de E^*, on note

$$\Phi^o := \{x \in E,\, \varphi(x) = 0,\, \forall \varphi \in \Phi\}.$$

On dit que Φ^o est l'antéorthogonal de Φ.

Le bidual évite de faire une seconde étude pour l'antéorthogonal ; ce dernier est un orthogonal qui s'ignore.

Proposition 3.5.4. *Avec les notation ci-dessus, $\iota(\Phi^o) = \Phi^\perp$. En particulier, Φ^o est un sous-espace vectoriel de E de dimension* $\dim E - \dim \Phi$.

Démonstration. Cette proposition est une tautologie qui résulte juste des définitions de ι, Φ^o et $\Phi^\perp$.

En effet, $x \in \Phi^o$ si et seulement si $\varphi(x) = 0$ pour tout $\varphi \in \Phi$, si et seulement si $\iota_x(\varphi) = 0$ pour tout $\varphi \in \Phi$ et donc si et seulement si $\iota_x \in \Phi^\perp$.

La dernière assertion provient alors du fait que ι est un isomorphisme, donc respecte les dimensions. ◇

Notation 3.5.5. On va à partir de maintenant identifier E avec E^{**}, ce qui va avoir pour effet d'être égalitaire entre la fonction et la variable, dans le sens que l'on voit $\varphi(x)$ comme un $x(\varphi)$ (on a en quelques sortes effacé la fonction ι) ; on fait jouer des rôles symétriques entre la fonction et la variable, contribuant ainsi un peu à la paix dans le monde. C'est peut-être un choix radical, mais nous n'utiliserons pas plus loin la notation Φ^o, puisque l'identification (via ι) entre E et E^{**} identifie Φ^o et $\Phi^\perp$.

Une fois que l'on voit l'orthogonalité "modulo l'identification entre E et E^{**}" comme c'est dit ci-dessus, on se rend compte que celle-ci est involutive.

Proposition 3.5.6. *Soit F un sous-espace de E, alors $(F^\perp)^\perp = F$.*

Démonstration. On a évidemment $F \subset (F^\perp)^\perp$, puisque si $x \in F$, et $\varphi \in F^\perp$, alors $x(\varphi) = \varphi(x) = 0$, et donc $x \in (F^\perp)^\perp$. L'égalité est affaire de dimension puisque $\dim(F^\perp)^\perp = n - \dim F^\perp = n - (n - \dim F) = \dim F$. ◇

3.6 Ordre et opérations sur les sous-espaces

Tout d'abord, l'orthogonalité est "décroissante" dans le sens qu'elle inverse l'ordre sur les sous-espaces :

Proposition 3.6.1. *Soit F, G deux sous-espaces vectoriels de E. Alors,*

$$F \subset G \Leftrightarrow G^\perp \subset F^\perp.$$

Démonstration. Il est clair que si $F \subset G$, alors toute forme qui annule G, annule en particulier F. D'où l'implication.

La réciproque résulte clairement de l'implication et du fait que $\perp$ est involutive. ◇

Nous allons voir que les deux opérations sur les sous-espaces sont en dualité.

Proposition 3.6.2. *Soit F, G deux sous-espaces vectoriels de E, alors*

$$(F+G)^\perp = F^\perp \cap G^\perp,\ (F \cap G)^\perp = F^\perp + G^\perp.$$

Démonstration. Montrons $(F+G)^\perp = F^\perp \cap G^\perp$. Comme $F, G \subset F+G$, on a par la proposition précédente $(F+G)^\perp \subset F^\perp, G^\perp$, et donc $(F+G)^\perp \subset F^\perp \cap G^\perp$.

Pour l'inclusion inverse, on suppose $\varphi \in F^\perp \cap G^\perp$ et $x \in F+G$. Alors, x peut se décomposer en $x_F + x_G$, avec $x_F \in F$ et $x_G \in G$. On en déduit $\varphi(x) = \varphi(x_F) + \varphi(x_G) = 0 + 0 = 0$. On a donc bien $\varphi \in (F+G)^\perp$.

Montrons maintenant $(F \cap G)^\perp = F^\perp + G^\perp$. Tout d'abord, $F \cap G \subset F, G$ et donc $F^\perp, G^\perp \subset (F \cap G)^\perp$. Il vient alors que $F^\perp + G^\perp \subset (F \cap G)^\perp$. Montrons l'égalité des dimensions en utilisant la formule de Grassmann et l'égalité que nous venons de montrer :

$$\begin{aligned}
\dim(F^\perp + G^\perp) &= \dim F^\perp + \dim G^\perp - \dim(F^\perp \cap G^\perp) \\
&= (n - \dim F) + (n - \dim G) - \dim(F+G)^\perp \\
&= 2n - \dim F - \dim G - \big(n - \dim(F+G)\big) \\
&= n - \big(\dim F + \dim G - \dim(F+G)\big) \\
&= n - \dim(F \cap G) = \dim(F \cap G)^\perp.
\end{aligned}$$

◇

Remarque 3.6.3. On pouvait aussi montrer la dernière égalité par la première en utilisant le bidual. En effet, on veut montrer que $(F \cap G)^\perp = F^\perp + G^\perp$. Comme $\perp$ est une involution, il suffit de montrer $F \cap G = (F^\perp + G^\perp)^\perp$, et donc $(F^\perp)^\perp \cap (G^\perp)^\perp = (F^\perp + G^\perp)^\perp$. Mais ceci n'est rien d'autre que la première égalité, appliquée aux deux sous-espaces $F^\perp$ et $G^\perp$ de E^*.

3.7 Cas particulier de la droite et de l'hyperplan

Une droite est un sous-espace engendré par un vecteur non nul. C'est une particule élémentaire des sous-espaces vectoriels d'un espace E dans le sens que tout sous-espace est somme de droites. Par définition, un hyperplan un sous-espace de codimension 1. Il

s'agit donc de l'orthogonal d'une droite, et c'est même une définition plus intéressante, parce que plus exploitable.

Pour donner une illustration de l'utilité de la version duale de l'hyperplan, disons que, par ce qui a été dit plus haut sur les droites, tout sous-espace est intersection d'hyperplans.

On peut même quantifier un peu les choses :

Proposition 3.7.1. *Soit E un espace vectoriel de dimension n. Alors,*

i. tout sous-espace F de dimension m de E peut s'écrire comme intersection d'une famille de $n-m$ hyperplans.

ii. soit $H_1, \cdots, H_k$ une famille de k hyperplans d'un espace vectoriel E de dimension n. Alors, $\dim H_1 \cap H_2 \cap \cdots \cap H_k \geq n-k$.

Démonstration. Montrons (i). Le sous-espace $F^\perp$ de E^* est de dimension $n-m$, donc possède une base (φ_i), pour $1 \leq n-m \leq n$. Le sous-espace $F^\perp$ est donc somme (directe) des droites $D_i := \mathbb{K}\varphi_i$. Donc $F = (F^\perp)^\perp = \cap_{i=1}^{n-m} D_i^\perp$. Ceci prouve la première assertion.

Montrons (ii). On pose $F := H_1 \cap H_2 \cap \cdots \cap H_k$. Il vient $F^\perp = \sum_{i=1}^k H_i^\perp$. Or, $H_i^\perp$ est une droite engendrée par, disons φ_i pour tout i, et donc, les φ_i, $1 \leq i \leq k$, est une famille génératrice de $F^\perp$. Il en découle que $\dim \mathbb{F}^\perp \leq k$, et donc $\dim F \geq n-k$. ◇

Remarque 3.7.2 (et pourquoi pas une intersection directe ?). Cela fait penser que la somme et l'intersection jouent des rôles symétriques puisque ces deux opérations sont en dualité. On sait qu'il existe une notion de somme directe qui précise la somme. Peut-on imaginer une notion d'intersection directe ? Quelle serait sa propriété principale ?

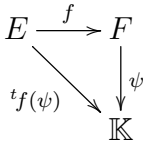
On peut y répondre comme ceci. On définit (en dualisant la notion de somme directe) une "intersection directe" de sous-espaces F_i de E une intersection $\cap_i F_i$ telle que pour tout i, $F_i + \cap_{j\neq i} F_j = E$. On a alors $\operatorname{codim} F = \sum_i \operatorname{codim} F_i$. Si, comme moi, un rien vous amuse, cette idée poétique d'intersection directe apportera un petit rayon de soleil à votre journée [3].

3.8 Transposée

Il reste encore le meilleur : dualiser les applications linéaires.

Définition 3.8.1. Soit E, F deux $\mathbb{K}$-espaces vectoriels (de dimension finie) et $f \in \mathcal{L}(E,F)$. On définit l'application ${}^t f$ de F^* dans E^* qui envoie $\psi \in F^*$ sur ${}^t f(\psi) = \psi \circ f$.

Il peut être plus agréable de voir cette définition avec un diagramme commutatif :



Exemple 3.8.2 (la dualité injection/restriction). Si E est un sous-espace de l'espace vectoriel F, on obtient une application linéaire $\iota \in \mathcal{L}(E,F)$ naturelle par l'injection : pour

3. "La dualité, à vous d'imaginer la vie qui va avec !"

$x \in E$, $\iota(x) = x \in F$. Soit ψ une forme linéaire sur F, alors ${}^t\iota(\psi) \in E^*$ et pour tout x de E, ${}^t\iota(\psi)(x) = \psi(\iota(x)) = \psi(x)$. Ce qui signifie que ${}^t\iota(\psi)$ est la restriction de ψ au sous-espace E de F. Moralité : le dual de l'injection est la restriction [4].

Proposition 3.8.3. *Avec les hypothèses ci-dessus, l'application ${}^tf$ est bien dans $\mathcal{L}(F^*, E^*)$.*

Démonstration. Tout d'abord, ${}^tf(\psi)$ est bien une application de E dans $\mathbb{K}$ par construction, et c'est une application linéaire comme composée d'applications linéaires. On a donc bien ${}^tf(\psi) \in E^*$ pour tout $\psi \in F^*$.

Ensuite, la linéarité de ${}^tf$ est claire puisque

$$ {}^tf(\mu_1\psi_1 + \mu_2\psi_2) = (\mu_1\psi_1 + \mu_2\psi_2) \circ f = \mu_1\psi_1 \circ f + \mu_2\psi_2 \circ f = \mu_1^t f(\psi_1) + \mu_2^t f(\psi_2). $$

◇

On dira dans la suite que ${}^tf$ est la transposée de f. De la même manière que l'orthogonalité envoie bijectivement les sous-espaces de E vers les sous-espaces de E^*, on va voir que la transposée envoie bijectivement les éléments les éléments de $\mathcal{L}(E, F)$ vers les éléments de $\mathcal{L}(F^*, E^*)$. Il est temps (déjà ?) d'étudier la transposée pour elle-même.

Proposition 3.8.4. *La transposée, qui envoie $f \in \mathcal{L}(E, F)$ sur ${}^tf \in \mathcal{L}(F^*, E^*)$, est une bijection linéaire. Mieux ! Si l'on identifie naturellement E avec E^{**}, ainsi que F avec F^{**}, alors la transposée est sa propre application réciproque.*

Démonstration. On a vu que la transposée définissait une application de $\mathcal{L}(E, F)$ vers $\mathcal{L}(F^*, E^*)$. Montrons la linéarité. Il suffit de montrer que ${}^t(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1^t f_1 + \lambda_2 4 f_2$, ce qui est clair car $\psi \circ (\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 \psi \circ f_1 + \lambda_2 \psi \circ f_2$, grâ à la linéarité de $\psi \in F^*$.

Il reste à montrer qu'en assimilant $x \in E$ à $x^{**} \in E$, et $y \in F$ à $y^{**} \in F$, alors ${}^t({}^tf) = f$. Tout d'abord, on a bien ${}^t({}^tf) \in \mathcal{L}(E^{**}, F^{**}) = \mathcal{L}(E, F)$. Il reste à montrer que pour tout $x \in E$, ${}^t({}^tf)(x^{**}) = f(x)$.

Soit $\psi \in F^*$, alors $f(x)$ (vu dans F^{**}) envoie $\psi \in F^*$ sur $\psi(f(x))$. Il reste donc à montrer que $({}^t({}^tf)(x^{**}))(\psi) = \psi(f(x))$. On respire un bon coup, et on y va. Tout d'abord ${}^t({}^tf)(x^{**} = x^{**} \circ^t f$. Donc,

$$ \begin{aligned} {}^t({}^tf)(x^{**}(\psi) &= (x^{**} \circ^t f)(\psi) = x^{**}({}^tf(\psi)) \\ &= x^{**}(\psi \circ f) = (\psi \circ f)(x) \\ &= \psi(f(x)). \end{aligned} $$

◇

Remarque 3.8.5. Le fait que la transposée soit involutive trouvera une raison peut-être plus convaincante pour la plupart : matriciellement, elle correspond (si on choisi bien les bases) à la transposée des matrices, qui est clairement involutive.

4. Ben v'la aut'chose !

3.9 Noyau et image d'une transposée

Les notions d'image et de noyau se retrouvent en dualité en algèbre linéaire. Avec elles, les notions d'équations paramétriques et d'équations cartésiennes, et plus philosophiquement, les notions de liberté et de contrainte, sont également en dualité dans ce contexte.

Proposition 3.9.1. *Soit E, F deux espaces vectoriels de dimension finie, et $f \in \mathcal{L}(E,F)$. Alors,* $\mathrm{rg}({}^t f) = \mathrm{rg}(f)$ *et*

$$\mathrm{Ker}({}^t f) = \mathrm{Im}(f)^\perp,\ \mathrm{Im}({}^t f) = \mathrm{Ker}(f)^\perp.$$

Démonstration. On va attaquer la proposition avec l'égalité $\mathrm{Ker}({}^t f) = \mathrm{Im}(f)^\perp$.

Soit φ une forme linéaire telle que ${}^t f(\varphi) = 0$. Alors, pour tout $y \in F$ de la forme $y = f(x)$, il vient $\varphi(f(x)) =^t f(\varphi)(x) = 0(x) = 0$. On vient de montrer l'inclusion $\mathrm{Ker}({}^t f) \subset \mathrm{Im}(f)^\perp$. L'inclusion réciproque est juste le chemin inverse puisque si $\varphi \in \mathrm{Im}(f)^\perp$, alors, pour tout x de E, ${}^t f(\varphi)(x) = \varphi(f(x)) = 0$, ce qui implique ${}^t f(\varphi) = 0$.

Il résulte de cette égalité, de la formule du rang et de la dimension de l'orthogonal, que

$$\mathrm{rg}({}^t f) = \dim F^* - \dim \mathrm{Ker}({}^t f) = \dim F^* - \dim \mathrm{Im}(f)^\perp = \dim \mathrm{Im}(f) = \mathrm{rg}(f).$$

Il reste à montrer que $\mathrm{Im}({}^t f) = \mathrm{Ker}(f)^\perp$. On montre tout d'abord l'inclusion, et l'égalité résultera directement de l'égalité des dimensions, vu que $\mathrm{rg}({}^t f) = \mathrm{rg}(f)$.

Soit $\varphi \in \mathrm{Im}({}^t f)$ et $x \in \mathrm{Ker}(f)$. Alors, il existe $\psi \in F^*$ tel que $\varphi =^t f(\psi)$, et donc

$$\varphi(x) =^t f(\psi)(x) = \psi(f(x)) = 0.$$

Ceci achève la preuve de la proposition. ◇

Remarque 3.9.2. La version matricielle des applications linéaires permet une jolie preuve du fait que $\mathrm{rg}({}^t f) = \mathrm{rg}(f)$. En effet, on peut faire provenir l'égalité $\mathrm{rg}({}^t A) = \mathrm{rg}(A)$ du fait que deux matrices sont équivalentes si et seulement si elles ont même rang.

Remarque 3.9.3. Soit E un sous-espace de F et prenons $f = \iota$, l'injection naturelle de E dans l'espace F. On a vu que ${}^t\iota$ est le morphisme de restriction de F^* vers E^*. On a donc $\mathrm{Im}({}^t\iota) = \mathrm{Ker}(\iota)^\perp = 0^\perp = E^*$. Et ceci veut dire que le morphisme de restriction est surjectif. On le savait déjà comme conséquence du théorème de la base incomplète. Cela nous rappelle qu'en algèbre linéaire, on n'est jamais bien loin du théorème de la base incomplète.

Corollaire 3.9.4. *Soit E et F deux espaces vectoriels de dimension finie et $f \in \mathcal{L}(E,F)$. Alors,*

i. f est injective si et seulement si ${}^t f$ est surjective,

ii. f est surjective si et seulement si ${}^t f$ est injective.

Démonstration. Le morphisme f est injectif ssi son noyau est nul dans E, ssi l'orthogonal de son noyau est E^*, ssi l'image de ${}^t f$ est E^*, ssi ${}^t f$ est surjective.

L'autre assertion est laissée en exercice. ◇

Ce dernier résultat pourra être utile à plus d'un titre. La preuve est laissée à la sagacité du lecteur.

Proposition 3.9.5. *Soit u un endomorphisme de E et F un sous-espace de E stable par u. Alors, $F^\perp$ est stable par ${}^t u$.*

3.10 Applications de la dualité

Avant d'énumérer quelques applications de la dualité, voici un lemme bien pratique dans nombre de situations.

Lemme 3.10.1. *Soit $(e_1, \cdots, e_m)$ une famille dans un espace vectoriel E et $(\varphi_1, \cdots, \varphi_m)$ une famille dans E^*. On suppose que $\varphi_i(e_j) = \delta_{ij}$ pour tout $1 \leq i, j \leq m$. Alors, les deux familles sont libres. Par conséquent, si $m = \dim E$, ce sont des bases, forcément duales l'une de l'autre.*

Démonstration. Si on applique φ_i à la relation $\sum_{j=1}^m \lambda_j e_j = 0$, on obtient $\lambda_i = 0$. Donc, la famille (e_j) est libre et la seconde assertion est analogue.

La dernière assertion provient du fait qu'une famille libre de cardinal $\dim E$ est une base. ◇

Voici quelques applications de la dualité.

• Polynômes interpolateurs de Lagrange. On choisit $x_1, \cdots, x_{n+1}$ deux à deux distincts dans $\mathbb{K}$ et on trouve le polynôme L_i qui vaut 1 sur x_i et 0 sur x_j, $j \neq i$. C'est la base duale dans $\mathbb{K}[X]_n$ de la base des évaluation en les x_j. Voir [**?**, 1.1.2].

• Formule de Taylor polynomiale en $a \in \mathbb{K}$ à l'ordre n. Elle provient de la base duale de la base des formes linéaires $P \mapsto P^{(i)}(a)$, pour $1 \leq i \leq n$. Voir [**?**, 1.1.2].

• Calculer la somme $P(0) + P(1) + \cdots P(n)$. On décompose P dans la base des polynômes de Hilbert. Pour construire la base des polynômes de Hilbert, on note Δ l'application (linéaire) qui envoie le polynôme P sur le polynôme $P(X+1) - P(X)$, et on calcule la base duale de la base des formes $\Delta^{(i)}(0)$. Voir Cours 5 agreg interne: Dualité 4

• Tout sous-espace est fermé pour la topologie (normique) de l'espace réel E. En effet, on a vu que tout sous-espace est intersection d'hyperplans, donc intersection de noyaux de formes linéaires, donc intersection de fermés (car une forme linéaire est continue en dimension finie), donc fermé.

• Tout hyperplan de $\mathcal{M}_n(\mathbb{K})$ contient une matrice inversible (et en fait une matrice de rang r pour tout rang, $0 \leq r \leq n$). Voir [**?**, 1.1.3].

• Tout hyperplan de $\mathcal{M}_n(\mathbb{R})$ contient une matrice orthogonale. Ces deux résultats proviennent des propriétés de la forme trace : toute forme linéaire sur $\mathcal{M}_n(\mathbb{K})$ peut s'écrire $X \mapsto \mathrm{tr}(AX)$ pour une matrice A (unique) de $\mathcal{M}_n(\mathbb{K})$. Voir [**?**, 1.1.3].

• On considère deux entiers m et n, avec $m \leq n$. Il y a autant de familles libres à m éléments dans $\mathbb{K}^n$ que de familles génératrices à n éléments dans $\mathbb{K}^m$. Voir [**?**, 1.1.6].

• Soit u un endomorphisme de l'espace E. Alors, pour tout polynôme P, Q $\mathrm{Im}(\mathrm{pgcd}(P,Q)(u)) = \mathrm{Im}\, P(u) + \mathrm{Im}\, Q(u)$ et $\mathrm{Im}(\mathrm{ppcm}(P,Q)(u)) = \mathrm{Im}\, P(u) \cap \mathrm{Im}\, Q(u)$. Voir [**?**, 1.1.9].

• On retrouve la dualité en analyse dans le calcul des intégrales de polynômes, par exemple dans la méthode des sécantes, cf. [**?**, Exercice 79], ou la méthode de quadrature de Gauss, cf. [**?**, Exercice 83].

• Mais la principale application de la dualité est dans la théorie des formes quadratiques en caractéristique différente de 2, où une forme quadratique sur un espace E est étudiée à partir d'une application linéaire de E vers son dual E^*.

3.11 Dualité : le mot de la fin

Il peut être éclairant de voir le dual de l'espace E comme l'ensemble des formes linéaires qui définissent les sous-espaces de E. En effet, on peut définir un sous-espace F par une famille génératrice (v_j) mais également par un système d'équations linéaires (φ_i) telles que ce sous-espace est défini comme l'annulateur de ces équations.

Cela signifie que l'on peut écrire, soit $F = \langle v_j \rangle$, soit $F = \langle \varphi_i \rangle^\perp$.

Mieux, l'ensemble de ces équations linéaires est lui-même doté d'une structure d'espace vectoriel. Le monde des espaces vectoriel définit des objets dans le monde dual : aux sous-espaces de E, muni des opérations $+$ et $\cap$, et de la relation inclusion, on associe les sous-espaces orthogonaux, et les opérations et relations correspondantes. Aux bases, on associe les bases duales et aux applications linéaires, les transposées.

Le bidual devient un objet théorique fondamental lorsqu'il est utilisé pour renvoyer la balle. En effet, grâce au bidual, tout espace W (en dimension finie) peut être vu comme le dual d'un espace V, et cet espace est canoniquement W^*. De la même manière, tout application linéaire peut être vue comme la transposée d'une application linéaire. C'est comme cela qu'on peut multiplier par deux les résultats et les théorèmes. Pour voir une utilisation de cela, on pourra se rapporter à [**?**, Exercice 1.1.9]

Chapitre 4

Matrices

4.1 Matrice d'une application linéaire

Soit E, F deux $\mathbb{K}$-espaces vectoriels de dimension respective n et m, soit $\underline{e} = (e_j)$ une base de E et $\underline{f} = (f_i)$ une base de F. Pour tout $f \in \mathcal{L}(E, F)$, on peut poser

$$f(e_j) := \sum_{i=1}^{m} a_{ij} f_i.$$

On note $\mathrm{mat}_{\underline{f},\underline{e}}(f)$ la matrice de f pour ces deux bases, il s'agit de la matrice (a_{ij}), ayant m lignes et n colonnes, où $a_{ij} \in \mathbb{K}$ est placé à la i-ème ligne, j-ème colonne. Cela revient à *écrire le vecteur $f(e_j)$ en colonne dans la base $\underline{f}$.*

Réciproquement, le théorème fondamental de l'algèbre linéaire assure que la donnée de l'image d'une base détermine l'application linéaire de façon unique. On a donc prouvé qu'il y a une bijection entre $\mathcal{L}(E, F)$ et l'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices ayant m lignes, n colonnes et à coefficients dans $\mathbb{K}$. Mieux, l'ensemble des matrices est muni d'une structure de $\mathbb{K}$-espace vectoriel naturel (on l'identifie en fait à $\mathbb{K}^{mn}$). On montre facilement que

Proposition 4.1.1. *On fixe une base $\underline{e}$ de E et une base $\underline{f}$ de F. L'application de $\mathcal{L}(E, F)$ vers $\mathcal{M}_{m,n}(\mathbb{K})$ qui envoie f sur $\mathrm{mat}_{\underline{f},\underline{e}}(f)$ est un isomorphisme d'espaces vectoriels.*

Dans toute la suite, on devra être concentré sur la correspondance entre matrices et applications linéaires, dit autrement, il faudra comprendre à tout moment ce qu'une matrice donnée raconte sur l'application linéaire qu'elle représente (pour des bases fixées).

Exemple 4.1.2. Par exemple, l'endomorphisme nul est envoyé sur la matrice nulle (normal, on envoie l'élément neutre sur l'élément neutre).

Exemple 4.1.3. On suppose $E = F$ et $\underline{e} = \underline{f}$. On voit que l'identité Id_E de E est envoyée sur la matrice $\mathrm{I}_n = \mathrm{diag}(1, \cdots, 1)$ dont la diagonale principale est constituée de 1 et avec de 0 partout ailleurs. On peut noter ce fait malheureusement rare que la matrice ici ne dépend que de f et non pas de la base $\underline{e}$ choisie. Toujours avec les mêmes hypothèses, on a la notion de matrice diagonale $\mathrm{diag}(a_1, \cdots, a_n)$ pour une matrice carrée de taille n. Cela signifie que $f(e_j) = a_j e_j$ pour tout j. Cette situation est tellement pratique et géométriquement simple qu'elle fera l'objet d'un pan entier de la théorie : la diagonalisation des endomorphismes.

On a aussi la notion de matrice diagonale par blocs. Cela signifie qu'un sous-espace de type $E_{k,k'} := \langle e_k, e_{k+1}, \cdots, e_{k'} \rangle$ est stable par f.

Notons que le $\mathbb{K}$-espace vectoriel E s'identifie naturellement avec $\mathcal{L}(\mathbb{K}, E)$. En effet, à un vecteur x de E, on associe l'application linéaire f_x de $\mathbb{K}$ dans E qui envoie 1 sur x. Le théorème fondamental de l'algèbre linéaire assure que l'on définit ainsi un isomorphisme entre E et $\mathcal{L}(\mathbb{K}, E)$.

Si $\underline{e}$ est une base de E, alors la matrice de f_x $\mathrm{mat}_{\underline{e},1}(f_x)$ n'est rien d'autre que le vecteur colonne x dans la base $\underline{e}$. On le note $\mathrm{mat}_{\underline{e}}(x)$.

4.2 Version duale de la matrice d'une application linéaire : la transposée

On vient de voir que la matrice d'une application linéaire se lisait en colonnes. On va maintenant donner un sens à ses lignes.

Tout d'abord, si $\varphi \in E^*$, alors, dans la base duale $\underline{e}^*$, on peut décomposer φ en $\varphi = \sum_{i=1}^n \varphi_i e_i^*$, avec $\varphi_i \in \mathbb{K}$. Comme $\varphi \in E^* = \mathcal{L}(E, \mathbb{K})$, et que $\varphi(e_i) = \varphi_i$ par construction, on a par définition,

$$\mathrm{mat}_{1,\underline{e}}(\varphi) = (\varphi_1 \ \cdots \ \varphi_n).$$

On comprend donc pourquoi les vecteurs s'écrivent en colonnes et les formes linéaires en lignes.

Si f est une application linéaire de E dans F, et si $\mathrm{mat}_{\underline{f},\underline{e}}(f) = A = (a_{ij})$, alors $f_i^* \circ f$ est une forme linéaire de E dans $\mathbb{K}$ et

$$(f_i^* \circ f)(e_j) = f_i^*(f(e_j)) = f_i^*(\sum_k a_{kj} f_k) = a_{ij}.$$

Résultat des courses, $\mathrm{mat}_{1,\underline{e}}(f_i^* \circ f)$ est la i-ème ligne de la matrice A.

On peut résumer ainsi : les colonnes de A sont les matrices colonnes des $f(e_j)$ dans la base $\underline{f}$ et les lignes de A sont les matrices lignes des $f_i^* \circ f$ dans la base $\underline{e}^*$.

Définition 4.2.1. Ceci appelle à une définition. La transposée d'une matrice $A = (a_{ij})$ de $\mathcal{M}_{m,n}(\mathbb{K})$ est une matrice notée ${}^tA$ de $\mathcal{M}_{n,m}(\mathbb{K})$ dont les colonnes sont les lignes de A, c'est-à-dire dont le coefficient en (i,j), $1 \leq i \leq n$, $1 \leq j \leq m$, est égal à a_{ji}.

Proposition 4.2.2. *Soit $f \in \mathcal{L}(E, F)$, avec E, F deux espaces vectoriels munis de bases respectives $\underline{e}$ et $\underline{f}$. Alors,*

$$\mathrm{mat}_{\underline{e}^*,\underline{f}^*}({}^tf) = {}^t\mathrm{mat}_{\underline{f},\underline{e}}(f).$$

Démonstration. Supposons que E soit de dimension n et F de dimension m. Alors, pour tout i de 1 à m, ${}^tf(_i^*$ est, par définition $f_i^* \circ f$. Or, on a vu que $f_i^* \circ f = \sum_{j=1}^n a_{ij} e_j^*$. Il en découle que la i-ème colonne de $\mathrm{mat}_{\underline{e}^*,\underline{f}^*}({}^tf)$ est la i-ème ligne de $\mathrm{mat}_{\underline{f},\underline{e}}(f)$. ◇

Remarque 4.2.3. Au passage, la transposée des matrices est linéaire, c'est-à-dire que ${}^t(\lambda A + \mu B) = \lambda {}^tA + \mu {}^tB$. Elle est bien sûr involutive car ${}^t({}^tA)) = A$, donc, bijective. Elle vérifie, à l'instar de l'inversion, la propriété ${}^t(AB) = {}^t B {}^tA$.

Proposition 4.2.4. *La transposée d'une matrice inversible de $\mathcal{M}_n(\mathbb{K})$ est inversible.*

Démonstration. Soit A une matrice inversible et B son inverse, alors ${}^tB{}^tA = {}^t(AB) = {}^t\mathrm{I}_n = \mathrm{I}_n$, ce qui prouve que ${}^tB$ est l'inverse (à gauche, donc à droite car ce sont des matrices carrées) de ${}^tA$. ◇

4.3 Matrices multipliables et multiplication matricielle

On a vu jusqu'à présent que l'espace vectoriel $\mathcal{L}(E,F)$ était isomorphe à $\mathcal{M}_{m,n}(\mathbb{K})$. Dans cette correspondance des structures, il ne reste plus qu'à ajouter une opération : la loi $\circ$.

Tout d'abord, notons que si $\varphi \in F^*$ a pour matrice ligne $(\varphi_1 \cdots \varphi_m)$ et $y \in F$ a pour coordonnées, dans la base $\underline{f}$ le vecteur colonne ${}^t(y_i \cdots y_m)$. Alors, $\varphi = \sum_j \varphi_j f_j^*$ et $y = \sum_j y_j f_j$. On a donc :

$$\varphi(y) = \sum_j \varphi_j y_j.$$

C'est notre première multiplication matricielle, et la plus élémentaire, en multipliant une matrice (ligne) de $\mathcal{M}_{1,m}$ par une matrice (colonne) de $\mathcal{M}_{m,1}$. On multiplie une ligne par une colonne en multipliant la j-ème colonne de la ligne, par la j-ème ligne de la colonne et en sommant le tout pour obtenir le scalaire $\varphi(y)$. La multiplication matricielle en général provient de cette opération élémentaire.

Soit E, F, G trois espaces vectoriels de dimension respective n, m, p, munis des bases respectives $\underline{e}$, $\underline{f}$, $\underline{g}$. Soit $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$, de sorte que $g \circ f \in \mathcal{L}(E,G)$. Alors,

Proposition 4.3.1. *Si $A = \mathrm{mat}_{\underline{f},\underline{e}}(f) = (a_{jk})$ et $B = \mathrm{mat}_{\underline{g},\underline{f}}(g) = (b_{ij})$, alors $\mathrm{mat}_{\underline{g},\underline{e}}(g \circ f) = (c_{ik})$, avec*

$$c_{ik} = \sum_{j=1}^{m} b_{ij} a_{jk}.$$

Démonstration. Le scalaire c_{ik} est par définition la i-ème coordonnée dans la base $\underline{g}$ de $g(f(e_k))$. C'est donc $g_i^*(g(f(e_k)))$, c'est-à-dire $(g_i^* \circ g)(f(e_k))$. D'après ce qui précède, c'est bien la i-ème ligne de B multipliée par la j-ème colonne de A. ◇

On peut donc définir la multiplication entre deux matrices A et B : on peut multiplier BA si $B \in \mathcal{M}_{p,m}$ et $A \in \mathcal{M}_{m,n}$ (penser à la relation de Chasles) ; on dit que les matrices sont multipliables. Le résultat est une matrice de $\mathcal{M}_{p,n}$ donnée par la formule de la proposition.

Deux matrices carrées de même taille sont donc multipliables. En fait, $(\mathcal{M}_n(\mathbb{K}), +, ., \times)$ est ce que l'on appelle une $\mathbb{K}$-algèbre et, d'après ce que l'on a vu, elle est isomorphe à la $\mathbb{K}$-algèbre $(\mathrm{End}(E), +, ., \circ)$. On associe tout simplement à un endomorphisme f de E, muni de la base $\underline{e}$, la matrice $\mathrm{mat}_{\underline{e},\underline{e}}(f)$.

4.4 Changement de bases

Ce "protocole Descartes" qui consiste à écrire sous forme d'objets de calcul (les matrices) les vecteurs et transformations géométriques, demande des bases. Et comme les

bases sont des objets de consommation courante, il faut s'attendre à changer de bases comme on change de chemise. Les formules de changement de bases deviennent primordiales dès que l'on rentre dans ce petit jeu, c'est-à-dire, depuis à peu près 1637 [1].

Soit E un espace vectoriel muni d'une base $\underline{e}$. On veut changer de base et remplacer l'ancienne base $\underline{e}$ par une nouvelle base $\underline{e}'$. On code ce changement de base par la matrice de passage de la base $\underline{e}$ à la base $\underline{e}'$ définie par $P := \mathrm{mat}_{\underline{e},\underline{e}'}(\mathrm{Id})$. Si l'on revient à la définition de la matrice d'une application linéaire, cela signifie que l'on écrit la nouvelle base en colonnes en fonction de l'ancienne.

Notons que $\mathrm{mat}_{\underline{e},\underline{e}'}(\mathrm{Id})\,\mathrm{mat}_{\underline{e}',\underline{e}}(\mathrm{Id}) = \mathrm{mat}_{\underline{e},\underline{e}}(\mathrm{Id}) = \mathrm{I}_n$, ce qui veut dire que la matrice de passage de la nouvelle base vers l'ancienne est la matrice $Q = P^{-1}$, c'est-à-dire l'inverse de la matrice P.

Il était clair qu'une matrice de passage est une matrice carrée, mais maintenant il devient clair qu'il s'agit d'une matrice inversible.

Soit E muni d'une ancienne base $\underline{e}$ et d'une nouvelle base $\underline{e}'$, avec P comme matrice de passage. Soit F muni d'une ancienne base $\underline{f}$ et d'une nouvelle base $\underline{f}'$, avec Q comme matrice de passage. On a la formule suivante dite de changement de base :

$$\mathrm{mat}_{\underline{f}',\underline{e}'}(f) = \mathrm{mat}_{\underline{f}',\underline{f}}(\mathrm{Id})\,\mathrm{mat}_{\underline{f},\underline{e}}(f)\,\mathrm{mat}_{\underline{e},\underline{e}'}(\mathrm{Id}) = Q^{-1}\,\mathrm{mat}_{\underline{f},\underline{e}}(f)P.$$

Définition 4.4.1. On dira que deux matrices A et B de $\mathcal{M}_{m,n}$ sont équivalentes si elles codent la même application linéaire f (dans des bases *a priori* différentes).

Remarque 4.4.2. On peut dire aussi que les deux matrices A et B sont équivalentes s'il existe une matrice inversible P de taille n et une matrice inversible Q de taille m telles que $B = Q^{-1}AP$.

On laisse le soin au lecteur de montrer que si $x \in E$:

$$\mathrm{mat}_{\underline{e}}(x) = P\,\mathrm{mat}_{\underline{e}'}(x).$$

On a vu aussi, cf. corollaire 3.3.2, que si E est muni d'une ancienne base $\underline{e}$ et d'une nouvelle base $\underline{e}'$, avec P comme matrice de passage. Alors, la matrice de passage de $\underline{e}^*$ vers $\underline{e}'^*$ est ${}^tP^{-1}$.

4.5 Le théorème du rang

Dans $\mathcal{M}_{m,n}(\mathbb{K})$, on a désormais la relation d'équivalence "est équivalente à". On va donc regrouper les matrices en classes d'équivalences. D'une part, il s'agit là d'un outil de classification, d'autre part, si f est dans $\mathcal{L}(E,F)$, la classe d'équivalence de l'ensemble des matrices qui "codent l'application linéaire f" est le seul bon objet matriciel que l'on peut associer à f dans faire de choix de base.

Le problème est maintenant de caractériser les classes d'équivalence. Pour cela, on va définir le rang d'une matrice.

Définition 4.5.1. Le rang d'une matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ est égal à la dimension sous-espace de $\mathbb{K}^m$ engendré par ses colonnes. Si A est une telle matrice, on notera son rang $\mathrm{rg}(A)$.

1. Date du *Discours de la méthode*.

Soit f dans $\mathcal{L}(E,F)$ et $A = \mathrm{mat}_{\underline{f},\underline{e}}(f)$. Comme $f(e_j)$, $1 \leq j \leq n$, engendre $\mathrm{Im}(f)$ et l'application "coordonnées dans la base $\underline{f}$" définit un isomorphisme de F vers $\mathbb{K}^m$, il vient que $\mathrm{Im}(f)$ est isomorphe au sous-espace engendré par les colonnes de A. En particulier

Proposition 4.5.2. *Soit f dans $\mathcal{L}(E,F)$ et $A = \mathrm{mat}_{\underline{f},\underline{e}}(f)$, alors* $\mathrm{rg}(f) = \mathrm{rg}(A)$.

On peut maintenant prouver le théorème du rang, qui peut être vu comme une version matricielle de la formule du rang. Si I_r est la matrice identité $r \times r$, on définit $\mathrm{I}_{m,n,r}$ comme étant la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ décrite par blocs de la façon suivante :

$$\mathrm{I}_{m,n,r} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

La matrice $\mathrm{I}_{m,n,r}$ est clairement de rang r.

Théorème 4.5.3. *Deux matrices de $\mathcal{M}_{m,n}(\mathbb{K})$ sont équivalentes si et seulement si elles ont même rang. En particulier, une matrice est de rang r si et seulement si elle est équivalente à $\mathrm{I}_{m,n,r}$.*

Démonstration. Si A et B sont équivalentes, elles codent la même application linéaire f, et donc $\mathrm{rg}(A) = \mathrm{rg}(f) = \mathrm{rg}(B)$.

Pour montrer la réciproque, il suffit de montrer que si A est de rang r, alors A est équivalente à $\mathrm{I}_{m,n,r}$. En effet, si B est de même rang que A, les deux matrices seront équivalentes à une matrice commune, et donc équivalentes entre elles par transitivité.

On va travailler dans les espaces vectoriels pour plus de souplesse (on peut changer de base sans changer d'application linéaire !). En fixant des bases respectives $\underline{e}$, $\underline{f}$ de E et F, soit $f \in \mathcal{L}(E,F)$ tel que $A = \mathrm{mat}_{\underline{f},\underline{e}}(f)$, dont l'existence est assurée par la proposition 4.1.1.

La formule du rang prouve que $\dim \mathrm{Ker}(f) = \dim E - \mathrm{rg}(f) = n - r$. Soit donc $(e'_{r+1}, \cdots, e'_n)$ une base de $\mathrm{Ker}(f)$. Il s'agit donc d'une partie libre de E que l'on complète en une base $\underline{e}' = (e'_1, \cdots, e'_r, e'_{r+1}, \cdots, e'_n)$.

Maintenant, soit $f'_i = f(e'_i)$ pour $1 \leq i \leq r$. Comme $S' := \langle e'_1, \cdots, e'_r \rangle$ est un sous-espace supplémentaire au noyau $\mathrm{Ker}(f)$, on sait que $f\,|_{S'}$ est injective. Donc, elle transforme une famille libre en une famille libre. Il en résulte que $(f'_1, \cdots, f'_r)$ est une famille libre que l'on peut compléter en une base $(f'_1, \cdots, f'_m)$ de F.

Tout est construit alors pour avoir $\mathrm{mat}_{\underline{f}',\underline{e}'}(f) = \mathrm{I}_{m,n,r}$. Ainsi, A et $\mathrm{I}_{m,n,r}$ sont équivalentes, ce qui achève la preuve.

◇

La matrice $\mathrm{I}_{m,n,r}$ sera par la suite la matrice "de référence" de rang r, on dit aussi qu'il s'agit d'une matrice *de forme normale* pour la relation d'équivalence. C'est si l'on veut un "bon" représentant de la classe. Bon, dans le sens que la matrice est jolie, légère, avec ses 1 élégamment disposés et ses zéros, et surtout, qu'il n'y a qu'une seule matrice de ce type par classe d'équivalence.

Corollaire 4.5.4. *Dans $\mathcal{M}_{m,n}(\mathbb{K})$, il y a exactement* $\min\{m,n\}+1$ *classes d'équivalence pour la relation "est équivalente à".*

Corollaire 4.5.5. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et ${}^tA$ sa matrice transposée dans $\mathcal{M}_{n,m}(\mathbb{K})$. Alors,* $\mathrm{rg}({}^tA) = \mathrm{rg}(A)$.

Démonstration. Par le théorème du rang, on a $A = P\mathrm{I}_{m,n,r}Q$ avec $P \in \mathcal{M}_m(\mathbb{K})$ et $Q \in \mathcal{M}_n(\mathbb{K})$ inversibles. Donc, ${}^tA = {}^tQ\,{}^t\mathrm{I}^t_{m,n,r}P = {}^tQ\mathrm{I}^t_{n,m,r}P$. Comme on l'a vu, ${}^tQ$ et ${}^tP$ sont inversibles, ce qui prouve que ${}^tA$ et $\mathrm{I}_{n,m,r}$ sont équivalentes, et donc $\mathrm{rg}({}^tA) = r = \mathrm{rg}(A)$. $\diamond$

Chapitre 5

Déterminant

Le déterminant est un objet profond qui possède de multiples avatars ainsi qu'un grand nombre d'interprétations. On le retrouve dans bon nombre de situations, tant calculatoires que théoriques, en algèbre linéaire, systèmes linéaires, polynômes, formes quadratiques, analyse sur les fonctions à plusieurs variables, calcul intégral... Nous allons ici tenter de lister ses définitions et ses principales propriétés. On se réfèrera à un bon livre de licence (Gourdon ou Grifone) pour les preuves lacunaires.

5.1 Définitions

Il existe plusieurs définitions du déterminant qu'il ne faut pas confondre.

5.1.1 Les formes n-linéaires alternées

Afin de partir d'un bon pied (comme dans Youpi Matin !), il est bon de démarrer avec le

Théorème 5.1.1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Il existe, à scalaire près, une unique application f non nulle de E^n dans $\mathbb{K}$, linéaire en chaque variable, et vérifiant $f(x_1, \cdots, x_n) = 0$ si $x_i = x_j$ pour $i \neq j$.*

Définition 5.1.2. Si f est linéaire en chaque variable, on dira que f est multilinéaire ou n-linéaire. Il est important de ne pas la confondre avec la linéarité [1]. Dans les calculs, la n-linéarité s'apparente à la distributivité.

Si f vérifie $f(x_1, \cdots, x_n) = 0$ si $x_i = x_j$ pour $i \neq j$, on dira que f est alternée.

Remarque 5.1.3. Ce théorème peut être interprété ainsi : l'ensemble $\mathrm{Alt}^n(E)$ des formes n-linéaires alternées sur un espace de dimension n est une droite vectorielle.

Dans ce théorème, il faut voir que l'unicité (à scalaire près) est la partie "facile" et l'existence difficile.

Pour ce qui est de l'unicité, on va dire qu'une forme n-linéaire alternée est entièrement déterminée par l'image d'une base. Dit autrement, soit $\underline{e}$ une base de E, alors si $f, f' \in \mathrm{Alt}^n(E)$, $f(\underline{e}) = f'(\underline{e})$ si et seulement si $f = f'$.

1. L'application f n'est pas une une forme linéaire de l'espace vectoriel E^n

Regardons ce qu'il se passe en dimension 2. En utilisant la multilinéarité et le caractère alterné, on trouve pour $f \in \mathrm{Alt}^2(\mathbb{K}^2)$

$$\begin{aligned} f(x,y) = f(ae_1 + be_2, ce_1 + de_2) &= acf(e_1,e_1) + adf(e_1,e_2) + bcf(e_2,e_1) + bdf(e_2,e_2) \\ &= adf(e_1,e_2) + bcf(e_2,e_1) \end{aligned}$$

Or, on peut calculer de deux manières $\alpha := f(e_1+e_2, e_1+e_2)$. D'une part, en utilisant le caractère alterné : $\alpha = 0$. D'autre part, en utilisant d'abord la n-linéarité :

$$\alpha = f(e_1,e_1) + f(e_1,e_2) + f(e_2,e_1) + f(e_2,e_2) = f(e_1,e_2) + f(e_2,e_1)$$

En comparant les deux résultats, il vient le résultat de "redressement"

$$f(e_2,e_1) = -f(e_1,e_2)$$

On trouve alors au final :

$$f(x,y) = (ad-bc)f(e_1,e_2)$$

Et cela prouve que f ne dépend que de l'image de la base $f(e_1,e_2)$. Donc, deux formes 2-linéaires alternées en dimension 2 sont proportionnelles, car elles sont toutes deux proportionnelles à la forme qui envoie (ae_1+be_2, ce_1+de_2) sur $ad-bc$.

On voit alors se profiler, telle la queue du loup, une généralisation en dimension n.

D'abord, $f(x_1,\cdots,x_n)$ se change en son opposé dès que l'on échange x_i et x_j pour $i \neq j$. Ensuite, si on fixe une base $\underline{e}$ de E, alors, après n-linéarité et redressement, on se convainc que l'on va obtenir, comme pour le cas $n=2$, que $f(x_1,\cdots,x_n)$ sera sous la forme $\delta((x_{ij}))f(e_1,\cdots,e_n)$, où x_{ij} est la j-ième coordonnée de x_i dans la base $\underline{e}$ et où δ est une fonction de la matrice (x_{ij}). Donc, deux formes n-linéaires alternées en dimension n sont proportionnelles.

Tout dans le théorème réside en fait dans le côté "non nul". En effet, nous n'avons pas montré qu'il existait une n-forme linéaire alternée telle que $f(e_1,\cdots,e_n) \neq 0$.

Dans le cas $n=2$, on a un bon candidat : $f(ae_1+be_2, ce_1+de_2) = ad-bc$. Encore faut-il prouver que ceci est bien 2-linéaire alterné. L'application f est linéaire en la première variable puisque, si on fixe c et d, on trouve une forme linéaire (homogène de degré 1), et idem pour la seconde variable. Pour le caractère alterné, il suffit de voir que si $(a,b) = (c,d)$, alors $ad-bc=0$, ce qui est clair.

Mais l'existence, dans le cas où n est quelconque, est plus retord. Pour s'en convaincre, redressons de trois manières différentes, dans le cas $n=3$, $f(x_3,x_2,x_1)$.

$$f(x_3,x_2,x_1) = -f(x_3,x_1,x_2) = f(x_1,x_3,x_2) = -f(x_1,x_2,x_3)$$

$$f(x_3,x_2,x_1) = -f(x_2,x_3,x_1) = f(x_2,x_1,x_3) = -f(x_1,x_2,x_3)$$

$$f(x_3,x_2,x_1) = -f(x_1,x_2,x_3)$$

Ouf ! Dans les trois cas, on tombe sur la même chose. Mais on voit qu'il n'y a pas unicité du mode de redressement ! Pourtant, on est tombés à la fin sur le même résultat.

Si on était tombés sur $-f(x_1, x_2, x_3)$ dans un cas, et $f(x_1, x_2, x_3)$ dans l'autre, on aurait conclu que f est nulle (en caractéristique différente de 2).

En fait, l'existence d'une fonction non nulle, dont on est convaincu à ce stade qu'elle est loin d'être évidente, repose principalement sur l'existence de la signature d'une permutation. Oui, ce fameux morphisme du groupe $\mathcal{S}_n$ sur $\{1, -1\}$ va se porter garant de l'existence du déterminant. Plus précisément, la parité du nombre de transpositions à effectuer pour construire une permutation ne dépend que de la permutation et non des transpositions utilisées [2].

Si on veut calculer $\alpha := f(x_{i_1}, \cdots, x_{i_n})$, alors, soit il y a répétition dans les i_k (l'application $k \mapsto i_k$ n'est pas injective), et dans ce cas $\alpha = 0$ car f est alternée ; soit, il n'y a pas répétition et dans ce cas, l'application $\sigma : k \mapsto i_k$ est bijective (pourquoi ?) et

$$f(x_{i_1}, \cdots, x_{i_n}) = \epsilon(\sigma) f(x_1, \cdots, x_n)$$

Définition 5.1.4. On définit dans ce contexte le déterminant $\det_{\underline{e}}$ comme la forme n-linéaire alternée qui vaut 1 sur $\underline{e}$.

Attention : le déterminant, défini ainsi, dépend d'une base, ce qui le distingue du déterminant directement défini sur l'espace des matrices carrées de taille n, (qui elle, a la chance de posséder une base canonique).

Remarque 5.1.5. Si $\mathbb{K} = \mathbb{R}$, cette dépendance de f à l'image de $\underline{e}$ correspond à ce que l'on attend d'une *forme-volume*. Le déterminant $\det_{\underline{e}}$ correspond à l'application qui, à une famille de n vecteurs, associe le volume du n parallélépipède engendré par cette famille, dans l'unité de volume où la base $\underline{e}$ engendre un parallélépipède de volume 1. On comprend bien que définir une forme-volume demande le caractère alterné (le volume d'un parallélépipède plat est nul), mais il faut un peu plus de temps (à peine en fait) pour se convaincre de la n-linéarité.

5.1.2 Définition matricielle

La définition matricielle découle de la définition précédente en posant $E = \mathbb{K}^n$. L'avantage est, d'une part, que E est doté de la base canonique $\underline{e}$, et d'autre part, que l'espace E^n s'identifie naturellement à l'espace des matrices carrées de taille n, colonne par colonne. On définit alors le déterminant d'une matrice A, donnée par ses colonnes $A_j \in \mathbb{K}^n$, j de 1 à n, par $\det(A) = \det_{\underline{e}}(A_1, \cdots, A_n)$, qu'il sera agréable d'écrire plus directement (par canonicité, *i.e.* indépendance de choix de base) $\det(A) = (A_1, \cdots, A_n)$.

On trouve donc

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

Plus généralement, si $A = (a_{ij})_{1 \leq i,j, \leq n}$, alors

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j)j}$$

2. Finalement, cela tient à peu de choses : si pas de signature, pas de déterminant, et si pas de déterminant... pas de déterminant.

Il est bon de voir σ comme la permutation qui envoie la colonne j vers la ligne $i := \sigma(j)$

Par exemple, dans le cas du déterminant de la matrice de taille 3, on reconnaîtra les signatures respectives des permutations Id, (123), (132), (23), (13), (12).

Attention, il s'agit là d'une formule théorique, rarement simple dans la pratique, à utiliser avec modération.

On peut montrer par exemple que le déterminant d'une matrice diagonale est le produit des a_{jj}, j de 1 à n. En effet, le seul terme (éventuellement) non nul dans le déterminant correspond à $\sigma : j \mapsto j$ pour tout j, donc $\sigma = \mathrm{Id}$, de signature 1.

De même, le déterminant d'une matrice triangulaire supérieure est encore une fois produit des a_{jj} diagonaux. On trouve encore une fois que $\sigma = \mathrm{Id}$ fournit le seul terme non nul $\prod_{j=1}^{n} a_{\sigma(j)j}$. Voyez plutôt : il est clair que σ doit envoyer 1 sur 1 (sinon, on tombe sur un terme nul), puis, 2 sur 1 ou 2, mais 1 est déjà pris, et donc, il ne reste plus que 2, ainsi de suite, on montre que σ envoie k sur k.

Voilà pour la contribution au calcul de cette formule. Mais cette formule possède une importance théorique majeure :

Proposition 5.1.6. *La fonction* det *sur* $\mathcal{M}_n(\mathbb{K})$ *est une fonction polynomiale à* n^2*- variables, de degré homogène* n*. En particulier, si* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$*, elle sera continue pour la topologie normique de l'espace des matrices.*

Cette propriété va être essentielle dans le cadre de la réduction (nombre fini de valeurs propres, densité des matrices inversibles, densité des matrices diagonalisables sur $\mathbb{C}$...)

Une autre conséquence immédiate de ce que l'on peut appeler "la formule explicite du déterminant" est

Proposition 5.1.7. *Soit* $A \in \mathcal{M}_n(\mathbb{K})$*, alors*

$$\det({}^tA) = \det(A)$$

Démonstration. Si on pose $A = (a_{ij})$ et ${}^tA = (b_{ij})$ avec $b_{ij} = a_{ji}$, alors

$$\det({}^tA) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} b_{\sigma(j)j} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma(j)}.$$

Par commutativité dans le produit, en posant $k = \sigma(j)$, on obtient

$$\det({}^tA) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma^{-1}(k)k}.$$

En posant le changement de variables $\tau = \sigma^{-1}$ et en remarquant que $\varepsilon(\tau^{-1}) = \varepsilon(\tau)$, il vient

$$\det({}^tA) = \sum_{\tau \in \mathcal{S}_n} \varepsilon(\tau) \prod_{j=1}^{n} a_{\tau(k)k} = \det(A)$$

et le tour est joué.

◇

Remarque 5.1.8. Ne soyons pas dupe : cette formule est un grossier avatar de la dualité des espaces vectoriels de dimension finie.

On termine avec une remarque évidente, qui sera utile par la suite :

Remarque 5.1.9. A l'aide de la formule explicite du déterminant, on peut définir le déterminant d'une matrice de $\mathcal{M}_n(\mathcal{A})$, où $\mathcal{A}$ est un anneau, et $\det(A) \in \mathcal{A}$. Ceci semble anodin mais peut devenir important, par exemple, au moment de définir le polynôme caractéristique d'un endomorphisme.

Remarque 5.1.10 (Calcul du déterminant par opérations sur les lignes).

Ce qui est vrai sur les colonnes va devenir vrai sur les lignes : on ne change pas le déterminant en remplaçant une ligne L_i par L_i plus une combinaison des autres lignes.

5.1.3 Déterminant d'un endomorphisme

C'est ici que le déterminant va prendre le leadership de l'algèbre linéaire, au tout début du 19-ème siècle avec Gauss, Cauchy et Binet, ce qui révolutionnera toutes les mathématiques. Un peu comme le coup d'état du 18 brumaire, sauf qu'en maths, les choses finissent mieux car la discipline protège de la mégalomanie.

Définition 5.1.11. Soit E un espace de dimension n, $\underline{e}$ une base de u, et u un endomorphisme de E. On va (provisoirement) noter $\det_{\underline{e}}(u)$ le déterminant dans la base $\underline{e}$ de la famille de vecteurs $(u(e_i)_{1\leq i\leq n})$.

Proposition 5.1.12. *Soit u un endomorphisme de E et A sa matrice dans une base $\underline{e}$. Alors le déterminant de u coïncide avec le déterminant de A.*

Démonstration. On a par définition $\det_{\underline{e}}(u)$ est l'image de $(u(e_i))$ par l'unique forme n-linéaire alternée qui envoie $\underline{e}$ sur 1. On identifie E et $\mathbb{K}^n$ en assimilant tout vecteur de E avec ses coordonnées dans la base $\underline{e}$.

On a

$$\det_{\underline{e}}(u) = \det_{\underline{e}}(u(e_1), \cdots, u(e_n)) = \det_{\underline{e}}(A_1, \cdots, A_n),$$

où les A_j sont les vecteurs $u(e_j)$ vus en colonnes dans $\mathbb{K}^n$. Donc, $\det_{\underline{e}}(u)$ est par définition le déterminant de la matrice A de l'endomorphisme u dans la base $\underline{e}$.

On obtient $\det_{\underline{e}}(u) = \det(A)$.

◇

Scoop La grande force du déterminant provient du fait que cette définition *ne dépend pas de la base choisie.*

Théorème 5.1.13. *Avec les hypothèses précédentes, pour deux bases $\underline{e}$ et $\underline{e}'$, on a* $\det_{\underline{e}}(u) = \det_{\underline{e}'}(u)$.

Par exemple, le déterminant de l'application nulle est 0, et le déterminant de l'application identité est 1.

5.2 Propriétés du déterminant

Une première propriété du déterminant d'une matrice et provenant directement de la n-linéarité est que

$$\begin{aligned}(C_1, \cdots, C_{i-1}, C_i' + C_i'', C_{i+1}, \cdots, C_n) = (C_1, \cdots, C_{i-1}, C_i', C_{i+1}, \cdots, C_n) \\ + (C_1, \cdots, C_{i-1}, C_i'', C_{i+1}, \cdots, C_n)\end{aligned}$$

$$(C_1, \cdots, C_{i-1}, \lambda C_i, C_{i+1}, \cdots, C_n) = \lambda(C_1, \cdots, C_{i-1}, C_i, C_{i+1}, \cdots, C_n)$$

Il faudra faire bien attention à ne pas confondre cette propriété avec de la linéarité. Par exemple, on a $\det(\lambda A) = \lambda^n \det(A)$ car det est, comme on l'a vu, une fonction polynomiale à n^2- variables, de degré homogène n.

Remarque 5.2.1 (Calcul du déterminant par opérations sur les colonnes).

En ajoutant à cela le caractère alterné du déterminant, on voit que l'on ne change pas le déterminant en remplaçant une colonne C_j par C_j plus une combinaison des autres colonnes. Ceci va nous permettre de calculer le déterminant par une méthode plus soft que le calcul brutal, appelée généralement "calcul du déterminant par opérations sur les colonnes".

Voici une propriété théorique du déterminant sous diverses formes :

Théorème 5.2.2. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. Alors,* $\det(AB) = \det(A)\det(B)$.

Démonstration. On pose $E := \mathbb{K}^n$ et f_A la forme sur E^n donnée par

$$f_A(u_1, \cdots, u_n) = \det(Au_1, \cdots, Au_n).$$

On montre facilement que f_A est une forme n-linéaire et alternée. Donc, il existe une constante λ_A telle que

$$f_A = \lambda_A \det$$

Si on applique cette égalité à la base canonique de $\mathbb{K}^n$, il vient $\det(A) = \lambda_A$.
Donc, si les u_j sont les colonnes de B, on voit que

$$\det(AB) = \det(Au_1, \cdots, Au_n) = \det(A)\det(u_1, \cdots, u_n) = \det(A)\det(B).$$

◇

Dites-le avec des morphismes de groupes

Corollaire 5.2.3. *Le déterminant définit un morphisme du groupe linéaire* $\mathrm{GL}_n(\mathbb{K})$ *dans le groupe multiplicatif* $\mathbb{K}^*$.

Corollaire 5.2.4. *Soit $A, P \in \mathcal{M}_n(\mathbb{K})$, avec P inversible. Alors,* $\det(PAP^{-1}) = \det(A)$.

On atteint alors le résultat annoncé sur le déterminant d'un endomorphisme.

Corollaire 5.2.5. *Le déterminant d'un endomorphisme ne dépend pas de la base $\underline{e}$ choisie dans sa définition.*

A ce propos, dites-le avec des endomorphismes :

Corollaire 5.2.6. *Soit $u, v \in \mathrm{End}(E)$, alors* $\det(u \circ v) = \det(u)\det(v)$.

Et au fait, le déterminant, il détermine quoi ?

Corollaire 5.2.7. *Une matrice est inversible si et seulement si son déterminant est non nul.*

Démonstration. Si A est inversible, il existe B telle que $AB = \mathrm{I}_n$, en appliquant le déterminant, il vient $\det(A)\det(B) = 1$ et donc $\det(A) \neq 0$.

Si A est non inversible, alors le théorème du rang nous dit que A est équivalente à une matrice diagonale D avec au moins un 0 sur la diagonale, donc, $A = PDQ$, avec $\det(D) = 0$. Ceci implique $\det(A) = 0$. ◇

5.3 Calcul du déterminant

Le calcul "à la main" d'un déterminant se fait généralement à l'aide d'opérations sur les colonnes (ou les lignes) afin de faire apparaître un maximum de zéros, et on finit à l'aide du "développement" par rapport à une colonne (ou une ligne). C'est ce développement que nous allons voir.

Tout d'abord, on remarque que si $A \in \mathcal{M}_n(\mathbb{K})$, alors

$$M = \begin{pmatrix} A & 0 \\ * & 1 \end{pmatrix}, \ \det(M) = \det(A)$$

Pour le montrer, on peut évoquer la formule explicite du déterminant qui dit que $\det(M) = \sum_{\sigma \in \mathcal{S}_{n+1}} \varepsilon(\sigma) \prod_{j=1}^{n+1} m_{\sigma(j)j}$. Or, vu la structure triangulaire par bloc de M, tous les termes correspondant à des permutations σ tels que $\sigma(n+1) \neq n+1$ sont nuls. On se ramène donc aux seuls termes où $\sigma(n+1) = n+1$ et donc $\sigma([1,n]) = [1,n]$. Comme $m_{ij} = a_{ij}$ pour $1 \leq i, j \leq n$, il vient immédiatement $\det(M) = \det(A)$.

Par permutation de lignes et de colonnes, on voit que

$$M = \begin{pmatrix} A_1 & 0 & A_2 \\ * & 1 & * \\ A_3 & 0 & A_4 \end{pmatrix}, \ \det(M) = (-1)^{i+j}\det(A), \text{ avec } A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix},$$

où le coefficient 1 se trouve sur la ligne i colonne j.

Définition 5.3.1. Soit A une matrice carré, on va noter $\Delta_{ij} = \Delta_{ij}(A)$ le déterminant de la matrice A dont on a retiré la ligne i et la colonne j. On note $m_{ij} := m_{ij}(A) = (-1)^{i+j}\Delta_{ij}(A)$ le cofacteur (i,j) de A. La matrice $(m_{ij})_{1 \leq i,j \leq n}$ est appelée comatrice de A et notée $\mathrm{Com}(A)$.

Le résultat précédent, prouve par multilinéarité du déterminant la formule du développement du déterminant par rapport à la colonne j

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}.$$

On a bien entendu une formule analogue de développement du déterminant par rapport à une ligne i.

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}.$$

On déduit de tout cela le théorème

Théorème 5.3.2. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors,*

$$A {}^t\mathrm{Com}(A) = \det(A) \mathrm{I}_n.$$

En particulier, si A est inversible, alors son inverse est donné par

$$A^{-1} = \frac{1}{\det(A)} {}^t\mathrm{Com}(A).$$

Démonstration. On va noter $N = {}^t\mathrm{Com}(A)$, c'est à dire $n_{ij} = m_{ji}(A) = m_{ji} = (-1)^{i+j}\Delta_{ij}$, et $B = AN$.

On a $b_{ik} = \sum_{j=1}^{n} a_{ij} n_{jk} = \sum_{j=1}^{n} a_{ij} m_{kj}$.

Le but est donc de montrer que $b_{ik} = \delta_{ik} \det(A)$.

Tout d'abord, si $i = k$, alors $b_{ii} = \sum_{j=1}^{n} a_{ij} m_{ij} = \det(A)$, d'après la formule de développement par rapport à une ligne.

Maintenant, supposons $i \neq k$. En notant $L_s(C)$ la ligne s d'une matrice C, on considère la matrice A' définie par $L_s(A') = L_s(A)$, si $s \neq k$, et $L_k(A') = L_i(A)$.

On calcule le déterminant de A' de deux manières différentes. D'une part, $\det(A') = 0$ puisque A' possède deux lignes égales. D'autre part, le développement de $\det(A')$ par rapport à sa k-ième ligne donne

$$\det(A') = \sum_{j=1}^{n} a_{ij} m_{ij}(A') = \sum_{j=1}^{n} a_{ij} m_{kj}(A) = \sum_{j=1}^{n} a_{ij} m_{kj} = b_{ik}.$$

Conclusion $b_{ik} = 0$ et on a rempli notre mission.

◇

Rappelons au passage que l'inverse à gauche d'une matrice carrée est égale à son inverse à droite.

Voici un corollaire important sur un anneau.

Corollaire 5.3.3. *Soit $\mathcal{A}$ un anneau et $A \in \mathcal{M}_n(\mathcal{A})$. Alors, A est inversible dans l'anneau $\mathcal{M}_n(\mathcal{A})$ (c'est-à-dire qu'il existe $B \in \mathcal{M}_n(\mathcal{A})$ telle que $AB = BA = \mathrm{I}_n$) si et seulement si $\det(A)$ est un inversible de l'anneau.*

Démonstration. Si $AB = \mathrm{I}_n$, alors $\det(A)\det(B) = \det(AB) = \det(\mathrm{I}_n) = 1$, ce qui fait de $\det(A)$ un inversible de l'anneau.

Réciproquement, si $\det(A)$ est inversible dans l'anneau $\mathcal{A}$, alors $\frac{1}{\det(A)} \in \mathcal{A}$. Comme tous les cofacteurs de A appartiennent tous à A, on a ${}^t\mathrm{Com}(A) \in \mathcal{M}_n(\mathcal{A})$. Conclusion, $\frac{1}{\det(A)} {}^t\mathrm{Com}(A) \in \mathcal{M}_n(\mathcal{A})$, et donc A est inversible dans $\mathcal{M}_n(\mathcal{A})$. ◇

Remarque 5.3.4. Le calcul effectif de A^{-1} par la transposée de la comatrice n'est pas raisonnable. C'est en résolvant un système par la méthode du pivot que l'on pourra trouver une méthode dont la complexité de calcul (en $\frac{2}{3}n^3$) est bien plus légère.

5.4 Systèmes de Cramer

Définition 5.4.1. Un système de la forme d'inconnue $AX = Y$, où $X \in \mathcal{M}_{n,1}(\mathbb{K})$, $Y \in \mathcal{M}_{n,1}(\mathbb{K})$, $A \in \mathrm{GL}_n(\mathbb{K})$ est appelé système de Cramer.

Ce système possède une solution unique $X = A^{-1}Y$. Mais, on peut trouver une façon de calculer X sans passer par la matrice inverse A^{-1}.

En effet, si on pose x_i la i-ème coordonnée du vecteur colonne X et si on note C_j la j-ième colonne de A, alors le système peut s'écrire

$\sum_{j=1}^{n} x_j C_j = Y$

Notons $D := \det(A)$ (qui est non nul par hypothèse) et $D_j = \det(A_j)$ où l'on a substitué la j-ième colonne de A par Y. Alors, comme le déterminant est n-linéaire et alterné, on sait que

$$D_j = x_j \det(A) = D.$$

Ceci implique $x_j = \frac{D_j}{D}$.

On constate une fois de plus la puissance de l'association "multi-linéaire" et "alterné".

5.5 Critère du rang par les mineurs

Voici un résultat (assez peu pratique, mais intéressant sur le plan théorique) de caractérisation du rang d'une matrice de $\mathcal{M}_{m,n}$ par ses mineurs (c'est-à-dire par les déterminants de ses sous-matrices carrées).

Théorème 5.5.1. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Alors, le rang de A est la taille maximum d'un mineur de A non nul.*

Démonstration. Soit r le rang de A et t la taille maximum d'un mineur non nul. On va montrer que $t \leq r$, puis, que $r \leq t$.

Montrons que $t \leq r$. Considérons les colonnes $j_1, \cdots, j_t$ de la sous-matrice de A dont on peut extraire un mineur non nul. Ces colonnes forment une sous-matrice B de A (de taille (m, t)). On voit que l'ensemble des solutions du système $BX = 0$ est réduit au sous-espace nul, puisque ce système contient, par hypothèse sur le mineur non nul, un sous-système de Cramer. On en déduit que la matrice B est de rang t, et donc, que le sous-espace engendré par les colonnes de A est de dimension au moins t. Ainsi, $t \leq r$.

Montrons que $r \leq t$. On considère le système $(C_1, \cdots, , C_n)$ formé des n colonnes de A. Par définition du rang, ce système engendre un espace de dimension r et par le théorème de la base extraite, on peut en extraire une base $(C_{j_1}, \cdots, C_{j_r})$. On porte alors notre regard sur la matrice B formée de ces seules colonnes, qui est une sous-matrice (m, r) de A : cette matrice B a des colonnes linéairement indépendantes, donc est de rang r. Comme le rang est invariant par transposition, on en déduit que ses lignes engendrent un sous-espace de dimension r. Ainsi, des lignes de B, on peut extraire, une base de r lignes de cet espace, comme on l'a déjà fait avec les colonnes.

On a extrait donc r lignes $(L_{i_1}, \cdots, L_{i_r})$ de B. On a ainsi un sous-matrice carrée de taille r de A, qui, par construction, est restée de rang r. On en déduit donc bien l'inégalité $r \leq t$.

◇

Voici un résultat inattendu qui en découle :

Corollaire 5.5.2. *Invariance du rang par extension Soit $\mathbb{K} \subset \mathbb{L}$ deux corps et $(v_1, \cdots, v_k)$ une famille de vecteurs de $\mathbb{K}^n$. Alors, le rang de la famille est la même, que l'on voie cette famille dans $\mathbb{K}^n$ ou dans $\mathbb{L}^n$.*

Démonstration. En effet, le rang de cette famille est égal au rang de la matrice de v_i écrits en colonne. Il suffit de montrer que cette matrice a le même rang, qu'elle soit vue sur $\mathbb{K}$ ou sur $\mathbb{L}$.

La caractérisation du rang par les mineurs rend trivial cette assertion (être nul dans $\mathbb{K}$ et dans $\mathbb{L}$ revient au même). ◇

5.6 Calcul du déterminant par blocs

On considère deux matrices carrées $A \in \mathcal{M}_m(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$. On développe le déterminant diagonal par blocs :

Comme

$$M := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \mathrm{I}_m & 0 \\ 0 & B \end{pmatrix},$$

on a après développement par rapport aux colonnes :

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

Considérons maintenant une matrice $C \in \mathcal{M}_{n,m}$. On veut calculer le déterminant de $N := \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$.

Voici deux méthodes instructives :

Méthode 1 : Le pivot de Gauss par blocs.

Si A est inversible, en utilisant le déterminant d'une matrice triangulaire ainsi que le développement d'une matrice diagonale par blocs.

$$N = \begin{pmatrix} \mathrm{I}_m & 0 \\ CA^{-1} & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det(A)\det(B).$$

Si A n'est pas inversible, alors, encore une fois, $\det(N) = \det(A)\det(B)$.

En effet, le membre de droite est nul, puisque $\det(A) = 0$. Le membre de gauche est également nul, puisque les m premières lignes de N sont liées (puisque les m premieres lignes de A le sont).

Dans les deux cas, on conclut que $\det(N) = \det(A)\det(B)$.

Méthode 2 : La fonction de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ qui, à B associe $\det \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$ est multilinéaire et alternée sur les colonnes de B. Donc, par le théorème fondamental du déterminant, il existe un scalaire λ tel que

$$\det \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \lambda \det(B).$$

En évaluant la fonction en $B = \mathrm{I}_n$, il vient

$$\lambda \det(\mathrm{I}_n) = \lambda = \det \begin{pmatrix} A & 0 \\ C & \mathrm{I}_n \end{pmatrix} = \det(A).$$

On conclut encore une fois que $\det(N) = \det(A)\det(B)$.

Proposition 5.6.1 (Calcul du déterminant par blocs)**.**

$$\det \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \det(A)\det(B).$$

5.7 Les douze travaux du déterminant

i. Inversibilité, calcul de l'inverse (notion de comatrice)

ii. Détermination du rang d'une matrice

iii. Résolution de systèmes

iv. Multiplicativité, Morphisme de groupe, Invariant de conjugaison

v. Polynomialité (en réduction, valeurs propres finies, polynômes caractéristique, produit des valeurs propres, $\mathrm{GL}_n(\mathbb{R})$ est dense, connexité de $\mathrm{GL}_n(\mathbb{C})$)

vi. Continuité sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ ($\mathrm{GL}_n(\mathbb{K})$ est un ouvert)

vii. Invariant de similitude et de congruence : condition nécessaire pour que deux matrices soient semblables, congruentes (notion générale de discriminant). Condition sur les mineurs principaux pour avoir un produit scalaire, pour calculer la signature. Critère de Sylvester.

viii. Equation de courbes et de surface...

ix. Définition de la norme sur $\mathbb{C}$ (et sur d'autres extensions)

x. Formes volumes

xi. Orientation de l'espace, similitudes directes, indirectes

xii. Jacobien

Exercice 5.7.1. Soit f une application de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ qui vérifie :

a) $f(AB) = f(A)f(B)$

b) Si A est diagonale, $f(A)$ est le produit des éléments diagonaux a_{ii}.

On veut montrer que f est le déterminant.

i. Montrer que si A est inversible, $f(A) \neq 0$ et $f(A^{-1}) = f(A)^{-1}$.

ii. Montrer réciproquement, que si $f(A)$ est non nul, A est inversible. *On pourra utiliser la contraposée et le théorème du rang.*

iii. Montrer que f est alternée.

iv. On rappelle qu'une matrice de transposition est une matrice de la forme $T_{ij}(\lambda) = \mathrm{I}_n + \lambda E_{ij}$, où les E_{ij} sont les matrices élémentaires et $i \neq j$. Montrer que $f(T_{ij}(\lambda)) = 1$. *En caractéristique différente de* 2*, on montrera que* $T_{ij}(\lambda)^2 = T_{ij}(2\lambda)$ *et que* $T_{ij}(2\lambda)$ *et* $T_{ij}(\lambda)$ *sont semblables. En caractéristique* 2*, on pourra adapter la preuve.*

v. En déduire que pour $i \neq j$,

$$f(C_1, \cdots, C_{i-1}, C_i + \lambda C_j, \cdots, C_{i+1}, \cdots, C_n) = f(C_1, \cdots, C_{i-1}, C_i, \cdots, C_{i+1}, \cdots, C_n).$$

vi. Montrer que f est linéaire en chacune des colonnes.

vii. Conclure que f est bien le déterminant.

Exercice 5.7.2 (Variations sur le rang)**.**
Comparer le rang d'une matrice à coefficients entiers sur $\mathbb{Q}$, sur $\mathbb{C}$, puis sur le corps $\mathbb{F}_p$.

Le rang d'une matrice, sur un corps quelconque, est égal à la taille du plus grand mineur non nul.

Comme la matrice est à coefficients entiers, chaque mineur est un entier.

Qu'il soit vu dans $\mathbb{Z}$, dans $\mathbb{Q}$ ou dans $\mathbb{C}$, un entier non nul reste non nul (car nous sommes en caractéristique nulle). Donc, les mêmes mineurs sont non nuls, qu'on les considère dans $\mathbb{Q}$ ou $\mathbb{C}$.

En revanche, si on considére un entier dans $\mathbb{F}_p$, c'est-à-dire, si l'on prend sa classe modulo p, il est possible qu'il s'annule.

Soit donc, $A = (a_{ij})$ une matrice à coefficients dans $\mathbb{Z}$. On peut réduire A modulo p pour obtenir $\overline{A} = (\bar{a}_{ij})$. Comme le déterminant est polynomial à coefficients entiers, on a $\det(\overline{A}) = \overline{\det(A)}$.

Ceci implique que si un mineur est nul, il reste nul après réduction de la matrice. Mais on voit facilement que la réciproque est fausse. Le rang devient plus petit après réduction modulo p.

Par exemple, la matrice $\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$ a pour déterminant 15. Elle est de rang 2 sur $\mathbb{Q}$ et sur $\mathbb{C}$, mais de rang 1 sur $\mathbb{Z}/3\mathbb{Z}$ et sur $\mathbb{Z}/5\mathbb{Z}$.

Exercice 5.7.3 (Déterminant et équation de cercle)**.**
On considère trois points non alignés, $A_i := (x_i, y_i)$, $1 \leq i \leq 3$, du plan affine euclidien $\mathbb{A}^2$. Montrer que l'équation du cercle passant par les points A_i (x_i, y_i), $1 \leq i \leq 3$ est donnée par

$$\begin{vmatrix} 1 & x & y & x^2+y^2 \\ 1 & x_1 & y_1 & x_1^2+y_1^2 \\ 1 & x_2 & y_2 & x_2^2+y_2^2 \\ 1 & x_3 & y_3 & x_3^2+y_3^2 \end{vmatrix} = 0$$

Pour tout i, les coordonnées des points A_i vérifient l'équation proposée. En effet, si on remplace (x, y) par (x_i, y_i), le déterminant est celui d'une matrice dont deux lignes sont égales ; il est donc nul.

De plus, en développant le déterminant par rapport à la première ligne, on obtient bien une équation de cercle, puisque l'on reconnaît la forme $ax^2 + ay^2 + bx + cy + d = 0$. Attention au piège, tout de même : il faut prouver que $a \neq 0$.

Or, le développement par rapport à la première ligne donne

$$-a = \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}.$$

En substituant la ligne L_2, resp. L_3, par $L_2 - L_1$, resp. $L_3 - L_1$, et en développant par rapport à la première colonne, il vient

$$-a = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}.$$

Or, comme les A_i ne sont pas alignés, $\overrightarrow{A_1A_2}$ et $\overrightarrow{A_1A_3}$ ne sont pas proportionnels, et donc $a \neq 0$.

Conclusion : comme il n'existe qu'un unique cercle passant par trois points non alignés (« le » cercle circonscrit du triangle), on obtient bien l'équation voulue.

Remarque 5.7.4. Si les trois points avaient été alignés (mais deux à deux distincts), les termes en degré 2 en x et y auraient disparu et on aurait obtenu l'équation de la droite passant par ces trois points.

En manipulant légèrement le déterminant 3×3 ci-dessus, on obtient au passage que les trois points A_i sont alignés si et seulement si

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

Ceci est facilement visible en plaçant les points A_i de $\mathbb{A}^2$ sur le plan affine d'équation $z = 1$, plongé dans l'espace vectoriel $\mathbb{R}^3$, en identifiant $A = (x, y)$ à $\overrightarrow{OA} = (x, y, 1)$. On voit alors que A_1, A_2, A_3 sont alignés si et seulement si $\overrightarrow{OA_1}$, $\overrightarrow{OA_2}$, $\overrightarrow{OA_3}$ sont coplanaires.

Remarque 5.7.5. Avec sa multilinéarité et son caractère alterné, le déterminant a plus d'un tour dans son sac. Il permet donc de trouver dans $\mathbb{R}^n$ des équations de cercles, mais aussi de droites, de plans, de coniques... passant par des points donnés.

Par exemple, si deux points (x_1, y_1) et (x_2, y_2) du plan sont distincts, alors l'équation de la droite passant par ces deux points est

$$\begin{vmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix} = 0$$

La conique passant par 5 points (x_i, y_i), $1 \leq i \leq 5$ a pour équation

$$\begin{vmatrix} 1 & x & y & x^2 & xy & y^2 \\ 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3y_3 & y_3^2 \\ 1 & x_4 & y_4 & x_4^2 & x_4y_4 & y_4^2 \\ 1 & x_5 & y_5 & x_5^2 & x_5y_5 & y_5^2 \end{vmatrix} = 0$$

"Le déterminant. A vous d'inventer la vie qui va avec !"

Chapitre 6

Réduction des endomorphismes

Nous avons vu que deux matrices sont équivalentes si et seulement si elles ont même rang. Il s'agit là d'un théorème qui culmine dans la théorie des applications linéaires. Si on veut maintenant comprendre les endomorphismes, c'est une autre paire de manche, puisqu'on ne peut jouer que sur une seule base, et non deux : la base de départ doit être la même que la base d'arrivée. Deux matrices A et B de $\mathcal{M}_n(\mathbb{K})$ sont semblables si et seulement si il existe P de $\mathrm{GL}_n(\mathbb{K})$ telle que $B = PAP^{-1}$. Deux matrices semblables sont équivalentes, mais la réciproque est clairement fausse. Même en dimension 1 (sauf sur $\mathbb{F}_2$), puisque puisque deux matrices équivalentes sont juste deux scalaires, soit tous deux nuls, soit tous deux non nuls ; et deux matrices semblables sont juste deux scalaires égaux.

6.1 Polynôme d'endomorphismes

On définit ici les polynômes d'endomorphismes par la proposition suivante.

Proposition 6.1.1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathrm{End}(E)$. Il existe un unique morphisme de $\mathbb{K}$-algèbres (unitaires)* ev_u

$$\mathrm{ev}_u \ : \ \mathbb{K}[X] \longrightarrow \mathrm{End}(E)$$

qui envoie X sur u.

Cette proposition repose sur la propriété universelle des algèbres de polynômes à une indéterminée.

On rappelle que $\mathbb{K}[X]$ est une $\mathbb{K}$-algèbre pour les lois $+$, $\cdot$ (loi externe), et $\times$ (multiplication interne) et que $\mathrm{End}(E)$ est une $\mathbb{K}$-algèbre pour les lois $+$, $\cdot$ (loi externe), et $\circ$ (multiplication interne).

Donc, si $P = \sum_{i=0}^{n} a_i X^i$, on a

$$P(u) := \mathrm{ev}_u(P) = a_n u^n + a_{n-1} u^{n-1} + \cdots + a_1 u + a_0 \,\mathrm{Id}\,.$$

Attention surtout à ne pas laisser a_0 tout seul sans le Id, ça fait désordre.

On a donc, en particulier,

$$(P+Q)(u) = P(u) + Q(u),\ (PQ)(u) = P(u) \circ Q(u),\ P^m(u) = P(u)^m.$$

Et surtout, il ne faut pas oublier le degré 0 : $1(u) = \mathrm{Id}$. En particulier, 1 n'est jamais annulateur (et encore moins minimal) d'un endomorphisme, sauf en dimension nulle.

On tire d'un morphisme une connexion entre deux mondes. Ici, ce n'est pas rien : on connecte le monde des polynômes (et son arithmétique, lemme de Gauss, décomposition en facteurs irréductibles, identité de Bezout...) avec l'algèbre linéaire (sous-espaces en somme directe, existence et unicité d'une décomposition, images et noyaux).

On a d'une part l'image de ev_u qui est la sous-algèbre $\mathbb{K}[u]$ des polynômes en u. Et d'autre part, le noyau de ev_u qui est l'idéal des polynômes annulateurs de u.

En dimension $n > 1$, le morphisme ev_u n'est pas surjectif puisque que son image $\mathbb{K}[u]$ est une sous-algèbre commutative de $\mathrm{End}(E)$.

En dimension finie, le morphisme ev_u n'est pas injectif puisque la dimension de l'espace de départ est infinie alors que la dimension de l'espace d'arrivée est finie (n^2).

Comme le noyau, qui est un idéal, de l'évaluation en u est non nul, on peut définir, de façon unique, un générateur unitaire.

Définition 6.1.2. Il existe un unique polynôme unitaire, appelé polynôme minimal de u, que l'on notera μ_u, tel que $\mathrm{Kerev}_u = (\mu_u)$, ou, dit autrement, tel que pour tout polynôme P de $\mathbb{K}[u]$

$$P(u) = 0 \Longleftrightarrow \mu_u \mid P.$$

Remarque 6.1.3. On peut faire de même avec les matrices. Attention, pour l'instant, le polynôme minimal d'une matrice A dépend du corps $\mathbb{K}$, et on devrait la noter $\mu_{A,\mathbb{K}}$. Si $\mathbb{K}$ est un sous-corps du corps $\mathbb{L}$, on devrait avoir $\mu_{A,\mathbb{L}} \mid \mu_{A,\mathbb{K}}$. On a en fait égalité, voir [**?**, Remarque III-5.2]. Cela provient de l'invariance du rang par extension du sous-espace engendré par les puissances A^k, $k \in \mathbb{N}$, comme on peut le voir dans le corollaire 5.5.2 et la proposition qui suit.

Proposition 6.1.4. *Le degré de μ_u est égal à la dimension de $\mathbb{K}[u]$.*

Démonstration. Soit d le degré de μ_u. On considère le morphisme de l'espace $\mathbb{K}[X]_{d-1}$ des polynômes de degré inférieur ou égal à $d-1$ vers $\mathbb{K}[u]$, qui envoie le polynôme P sur $P(u)$. Si on montre que c'est un isomorphisme, on aura montré que $\dim \mathbb{K}[u] = \dim \mathbb{K}[X]_{d-1} = d$, comme demandé.

Il suffit donc de montrer que pour tout élément v de $\mathbb{K}[u]$, il existe un unique P de $\mathbb{K}[X]_{d-1}$ tel que $v = P(u)$.

Unicité. On suppose $P_1(u) = P_2(u)$, avec $P_1, P_2 \in \mathbb{K}[X]_{d-1}$. Alors, $P_1 - P_2$ annule u et donc μ_u divise $P_1 - P_2$. Or, le premier polynôme est de degré d et le second, de degré $< d$. La seule possibilité est que $P_1 - P_2 = 0$. D'où l'unicité.

Existence. Comme $v \in \mathbb{K}[u]$, il existe un polynôme $P \in \mathbb{K}[X]$ tel que $P(u) = v$. On effectue la division euclidienne de P par μ_u :

$$P = \mu_u \cdot Q + R, \ \deg(R) < d.$$

On évalue en u pour obtenir

$$v = P(u) = \mu_u(u) \circ Q(u) + R(u) = R(u),$$

d'où l'existence. ◇

De la même manière on définit l'évaluation en une matrice, et toutes les résultats qui précèdent possèdent une version matricielle qu'il n'est pas difficile de deviner.

6.2 Polynôme minimal local et matrice compagnon

Nous avons défini l'évaluation d'un polynôme en un endomorphisme. On peut facilement définir l'évaluation d'un endomorphisme en un vecteur d'un espace E. Il s'agit de l'application ev_x qui envoie un endomorphisme u de $\mathrm{End}(E)$ sur $u(x)$. Il s'agit d'un morphisme d'espaces vectoriels (et non plus de $\mathbb{K}$-algèbres).

En concaténant les deux morphismes d'évaluation (en u, puis en x) on obtient un morphisme de $\mathbb{K}[X]$ dans E qui envoie le polynôme P sur $P(u)(x)$. On a l'impression de perdre beaucoup en structures (ce n'est qu'un morphisme d'espaces, alors que $\mathbb{K}[X]$ a tant d'arithmétique à donner), mais on peut se rassurer en montrant que le noyau de ce morphisme est un idéal de $\mathbb{K}[X]$. Il suffit pour cela de montrer que si P et Q sont dans le noyau, et $A \in \mathbb{K}[X]$, alors $(P+Q)(u)(x) = 0$, $(AP)(u)(x) = 0$. La première égalité est évidente (elle découle du morphisme d'espace), et la seconde se voit facilement puisque :

$$(AP)(u)(x) = A(u)(P(u)(x)) = A(u)(0) = 0.$$

On en déduit que le morphisme $\mathrm{ev}_x \circ \mathrm{ev}_u$ possède un noyau non trivial (déjà μ_u y appartient !) qui est un idéal principal, engendré par un unique polynôme unitaire que nous noterons $\mu_{u,x}$.

Proposition 6.2.1. *Soit $u \in \mathrm{End}(E)$ et $x \in E$. Il existe un unique polynôme $\mu_{u,x}$ tel que pour tout polynôme P de $\mathbb{K}[X]$, $P(u)(x) = 0$ si et seulement si $\mu_{u,x}$ divise P.*

Définition 6.2.2. Le polynôme $\mu_{u,x}$ est appelé polynôme minimal local de u en x. Le sous-espace engendré par les $u^k(x)$, $k \in \mathbb{N}$, est le sous-espace u-stable (ou u-cyclique) engendré [1] par x.

Remarque 6.2.3. Comme dans la preuve de la proposition 6.1.4, on prouve que le rang de $\mathrm{ev}_x \circ \mathrm{ev}_u$ est égal à $\deg \mu_{u,x}$. Or, l'image de $\mathrm{ev}_x \circ \mathrm{ev}_u$ est le sous-espace engendré par les $u^k(x)$, $k \in \mathbb{N}$. On en déduit

$$\deg \mu_{u,x} = \dim\langle u^k(x),\, k \in \mathbb{N}\rangle$$

Définition 6.2.4. Soit $P \in \mathbb{K}[X]$ unitaire, avec $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0$. On note C_P la matrice compagnon de P donnée par

$$C_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & 0 & -a_2 \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Selon le contexte, on pourra être amené à appeler matrice compagnon la transposée de cette matrice.

1. On voit clairement que c'est le plus petit sous-espace u-stable contenant x.

Proposition 6.2.5. *Soit $u \in \mathrm{End}(E)$ et $x \in E$ non nul. Soit s l'entier maximal tel que la famille $(x, u(x), \cdots, u^{s-1}(x))$ soit libre. Alors*

i. la famille $(x, u(x), \cdots, u^{s-1}(x))$ est une base du sous-espace u-stable $\langle u^k(x),\, k \in \mathbb{N}\rangle$ engendré par x.

ii. la matrice de l'induit de u sur ce sous-espace (qui est u-stable) dans cette base est la matrice compagnon de $\mu_{u,x}$.

Démonstration. Tout d'abord, il n'est pas inutile de noter que s existe puisque l'on est en dimension finie, et donc, l'ensemble des k tels que la famille $(x, u(x), \cdots, u^{k-1}(x))$ est libre est finie. De plus, elle est non vide, car x est non nul. Elle possède donc bien un élément maximal.

Notons $F = \langle u^k(x),\, k \in \mathbb{N}\rangle$. La famille $(x, u(x), \cdots, u^{s-1}(x))$ est bien dans F, et libre par construction. Montrons qu'elle est génératrice. Par maximalité, $(x, u(x), \cdots, u^{k-1}(x), u^s(x))$ est liée, et si $b_0x + b_1u(x) + \cdots + b_su^s(x) = 0$ est une relation de liaison, alors $b_s \neq 0$ (sinon par l'absurde, on aurait une relation de liaison dans la famille libre $(x, u(x), \cdots, u^{s-1}(x))$). On peut donc écrire

$$u^s(x) = -a_0x - a_1u(x) - \cdots - a_{s-1}u^{s-1}(x), \quad \text{avec } a_i = \frac{b_i}{b_r}.$$

On a donc pour tout $k \geq s$

$$u^k(x) = a_0u^{k-s}(x) + a_1u^{k-s+1}(x) + \cdots + a_{s-1}u^{k-1}(x).$$

Par récurrence, on voit alors que pour tout k, $u^k(x)$ est combinaison linéaire de la famille $(x, u(x), \cdots, u^{s-1}(x))$. Donc, F est bien engendrée par cette famille.

Pour la dernière assertion, on remarque que $u(u^m(x)) = u^{m+1}(x)$, pour tout m, $0 \leq m \leq s-2$, et

$$u(u^{s-1}(x)) = u^s(x) = -a_0x - a_1u(x) - \cdots + a_{s-1}u^{s-1}(x).$$

Il suffit donc de montrer que le polynôme minimal local de u en x est

$$\mu_{u,x} = X^d + a_{d-1}X^{d-1} + \cdots + a_0$$

.

Soit P le polynôme du membre de droite. Il est clair que $\mu_{u,x}$ divise P, puisque par construction $P(u)(x) = 0$. Comme $\mu_{u,x}$ et P sont unitaires, il suffit de voir que $d \leq \deg(\mu_{u,x})$.

Par l'absurde, comme $\mu_{u,x}(x) = 0$, si $\deg(\mu_{u,x}) < d$, on aurait une relation de liaison dans la famille libre $(x, u(x), \cdots, u^{s-1)(x))}$, ce qui est impossible.

◇

6.3 Lemme des noyaux

Le lemme des noyaux va nous permettre d'étudier un endomorphisme sur un espace E, à l'aide de plusieurs "études locales", c'est-à-dire l'étude de la restriction de E à des sous-espaces stables, plus simples à comprendre. Ce lemme est un exemple type de "passage du local au global" tel qu'on les aime en mathématiques.

La preuve du lemme des noyaux est basée sur le morphisme d'évaluation en u suivi de l'endomorphisme d'évaluation en x, que nous venons de voir, et qui permet de basculer de l'arithmétique des polynômes vers les propriétés géométriques dans l'espace E.

On commence par un petit lemme qui sert tout le temps :

Lemme 6.3.1. *Si $u, v \in \mathrm{End}(E)$ commutent, le noyau de v est stable par u.*

Démonstration. Si $x \in \mathrm{Ker}(v)$, $v(u(x)) = u(v(x)) = u(0) = 0$, donc $u(x) \in \mathrm{Ker}(v)$. ⋄

Exercice 6.3.2. Montrer de même que si $u, v \in \mathrm{End}(E)$ commutent, l'image de v est stable par u.

Proposition 6.3.3. *Soit P et Q deux polynômes premiers entre eux dans $\mathbb{K}[X]$, et $u \in \mathrm{End}(E)$. On a la décomposition en somme directe :*

$$\mathrm{Ker}(PQ)(u) = \mathrm{Ker}P(u) \oplus \mathrm{Ker}Q(u).$$

Démonstration. On part de l'identité de Bezout, qui assure l'existence de polynômes U, V de $\mathbb{K}[X]$ tels que $UP + VQ = 1$.

Voici un enchaînement fondamental dans l'étude des endomorphismes :

$UP + VQ = 1$ par le morphisme (d'algèbres) ev_u devient $U(u)P(u) + V(u)Q(u) = \mathrm{Id}$, qui, par le morphisme (d'espaces) devient à son tour, pour $x \in$Ker(PQ)(u), $U(u)(P(u)(x)) + V(u)(Q(u))(x) = x$. $(*)$

Or, par une petite manipulation, on obtient

$$Q(u)\big(U(u)(P(u)(x))\big) = U(u)\big((PQ)(u)(x)\big) = 0,$$

donc $U(u)(P(u)(x)) \in \mathrm{Ker}Q(u)$. Et, de même, $V(u)(Q(u)(x)) \in \mathrm{Ker}P(u)$. On a donc montré que $\mathrm{Ker}(PQ)(u) = \mathrm{Ker}P(u) + \mathrm{Ker}Q(u)$.

Il reste à montrer $\mathrm{Ker}P(u) \cap \mathrm{Ker}Q(u) = \{0\}$. On suppose donc $x \in \mathrm{Ker}P(u) \cap \mathrm{Ker}Q(u)$, et $(*)$ fournit $0 = x$, ce qui achève la preuve. ⋄

Remarque 6.3.4. Dans [**?**, Exercice 1.1.9], on trouvera une généralisation de ce résultat lorsque P et Q ne sont pas nécessairement premiers entre eux. Dans ce cas, si D et M désignent respectivement leur PGCD et leur PPCM :

$$\mathrm{Ker}(D(u)) = \mathrm{Ker}(P(u)) \cap \mathrm{Ker}(Q(u)), \;\; \mathrm{Ker}(M(u)) = \mathrm{Ker}(P(u)) + \mathrm{Ker}(Q(u)).$$

En effet, si $D = 1$, alors $\mathrm{Ker}(D(u)) = \mathrm{Ker}(\mathrm{Id}) = \{0\}$, et $M = PQ$, et ainsi, $\mathrm{Ker}(PQ(u)) = \mathrm{Ker}(P(u)) + \mathrm{Ker}(Q(u))$. On retrouve bien la somme directe.

Notez au passage cette propriété de croissance du noyau qui peut être bien utile : si P divise Q, alors $\mathrm{Ker}(P(u))$ est inclus dans $\mathrm{Ker}(Q(u))$. C'est la propriété dite des noyaux emboîtés. En revanche, l'application qui envoie P sur $\mathrm{Im}(P(u))$ est décroissante.

Corollaire 6.3.5. *Soit E un espace vectoriel muni d'un endomorphisme u. On suppose que le polynôme PQ est un polynôme annulateur de u, avec P et Q premiers entre eux. Alors,*

$$E = \mathrm{Ker}P(u) \oplus \mathrm{Ker}Q(u).$$

De plus, la projection de E sur $\mathrm{Ker}P(u)$, parallèlement à $\mathrm{Ker}Q(u)$, est un polynôme en u. De même, la projection de E sur $\mathrm{Ker}Q(u)$, parallèlement à $\mathrm{Ker}P(u)$, est un polynôme en u.

Démonstration. C'est juste la proposition avec $\mathrm{Ker}(PQ)(u) = \mathrm{Ker}(0) = E$.

Pour la dernière assertion, appelons $K_P = \mathrm{Ker}P(u)$ et $K_Q = \mathrm{Ker}Q(u)$. Les sous-espaces K_P et K_Q sont en somme directe, donc tout x de E se décompose en $x_P + x_Q$, avec $x_P \in K_P$, $x_Q \in K_Q$, et ce, de façon unique. Par unicité, on a, par la preuve de la proposition précédente, que $(VQ)(u)(x) = x_P$, et $(UP)(u)(x) = x_Q$. Ce qui prouve notre assertion. ◇

Par récurrence sur le nombre de facteurs, on en déduit le lemme des noyaux

Théorème 6.3.6 (Lemme des noyaux, version 1)**.** *On suppose que P annule l'endomorphisme u de E avec $P = \prod_i P_i$ et les P_i deux à deux premiers entre eux. Alors*

$$E = \bigoplus_i \mathrm{Ker}(P_i(u)).$$

De plus, les projecteurs sur chaque composante sont des polynômes en u.

Pour la preuve, on applique juste une récurrence qui part du fait que si les P_i, $1 \leq i \leq k$, deux à deux premiers entre eux, alors P_k est premier à $P_1 \cdots P_{k-1}$. Pour la derniere assertion, on se sert du fait que la composée de deux polynômes en u est un polynôme en u : $R(u) \circ S(u) = (RS)(u)$.

Remarque 6.3.7. On peut donner des précisions sur ces projecteurs. Soit Π_i le projecteur sur la composante $\mathrm{Ker}(P_i(u))$ parallèlement aux autres composantes. Alors, Π_i peut se retrouver ainsi :

On note $Q_i = \prod_{j \neq i} P_j$ de sorte que les Q_i sont *globalement* premiers entre eux. On a donc une relation de Bezout

$$\sum_i U_i Q_i = 1,$$

et $\Pi_i = (U_i Q_i)(u)$.

Théorème 6.3.8 (Lemme des noyaux, version 2)**.** *On suppose que P annule l'endomorphisme u de E avec $P = \prod_\lambda (X - \lambda)^{k_\lambda}$, où les λ sont deux à deux distincts. Alors*

$$E = \bigoplus_i \mathrm{Ker}(u - \lambda \,\mathrm{Id})^{k_\lambda}.$$

De plus, les projecteurs sur chaque composante sont des polynômes en u.

Définition 6.3.9. Dans le cadre du théorème, μ_u divise P, et donc, en particulier, toute valeur propre (voir définition 6.4.4) est racine de P. Si λ est valeur propre, alors le projecteur sur $\mathrm{Ker}(u - \lambda \,\mathrm{Id})^{k_\lambda}$ est appelé projecteur spectral.

Les sous-espaces $K_\lambda := \mathrm{Ker}(u - \lambda \,\mathrm{Id})^{k_\lambda}$ sont appelés sous-espaces spectraux (ils ne dépendent pas du polynôme annulateur utilisé).

Notons au passage que Si λ n'est pas valeur propre, $\mathrm{Ker}(u - \lambda \,\mathrm{Id})^{k_\lambda}$ est le sous-espace nul.

Pourquoi s'intéresse-t-on tant au fait que les projecteurs spectraux sont des polynômes en u ? Tout simplement parce que les polynômes en u commutent avec u. Par le

lemme 6.3.1, les sous-espaces spectraux sont stables par u. Le lemme des noyau est donc un lemme de passage du local au global, dans le sens que l'on arrive à comprendre l'endomorphisme u à partir de restrictions u_λ "spectralement simple". En effet, la restriction u_λ de u à K_λ est un endomorphisme annulé par le polynôme $(X-\lambda)^{k_\lambda}$, qui ne possède qu'une seule racine.

6.4 Polynôme caractéristique

Le polynôme caractéristique vient à point pour détecter les valeurs de λ dans $\mathbb{K}$, telles que le sous-espace caractéristique K_λ est non nul.

Définition 6.4.1. On définit le polynôme caractéristique $\chi_u \in \mathbb{K}[X]$ de l'endomorphisme u de E par

$$\chi_u = \det(X\,\mathrm{Id} - u).$$

Tout d'abord, on a bien fait de définir le déterminant sur un anneau, plutôt que de rester cantonné à un corps. Ici, on travaille sur l'anneau de polynômes $\mathbb{K}[X]$.

Comme le déterminant est un invariant de similitude, on voit que si l'on définit le polynôme caractéristique d'une matrice carrée A par

$$\chi_A = \det(X\,\mathrm{Id} - A),$$

alors, on a $\chi_u = \chi_A$ pour toute matrice A qui code l'endomorphisme u dans une base quelconque. C'est une très bonne nouvelle parce qu'on va pouvoir calculer χ_u dans la base qui nous arrange le mieux.

Remarque 6.4.2. Notons que cette version du polynôme caractéristique est assez récente ; elle a remplacé l'ancienne $\chi_A = \det(A - X\,\mathrm{Id})$. En gros, on a préféré avoir un polynôme unitaire plutôt qu'un déterminant calculable à la main sans erreurs de signe. C'est une bonne nouvelle pour la théorie, mais une mauvaise pour le calcul à la main. Mais bon, qui calcule encore à la main de nos jours. Ah oui, les candidats dans les écrits de concours !

Proposition 6.4.3. *Soit u un endomorphisme de E. Les conditions suivantes sur le scalaire $\lambda \in \mathbb{K}$ sont équivalentes.*

i. $\chi_u(\lambda) = 0$,

ii. $\mathrm{Ker}(u - \lambda\,\mathrm{Id}) \neq 0$,

iii. $u - \lambda\,\mathrm{Id}$ *n'est pas injective,*

iv. $u - \lambda\,\mathrm{Id}$ *n'est pas surjective,*

v. $u - \lambda\,\mathrm{Id}$ *n'est pas bijective,*

vi. le sous-espace caractéristique K_λ est non nul.

Démonstration. La preuve est assez immédiate, basée sur la caractérisation de la bijectivité d'un endomorphisme par le déterminant. Pour la dernière assertion, on a juste à remarquer que $\det(A) = 0$ si et seulement si $\det(A)^k = 0$, pour tout $k > 0$. ◇

Définition 6.4.4. Les $\lambda \in \mathbb{K}$ vérifiant ces propriétés équivalentes sont appelées valeurs propres de u. Par extension (c'est le cas de dire), on dira que λ est valeur propre de u si $\lambda \in \mathbb{L}$, vérifie $\chi_u(\lambda) = 0$, où $\mathbb{L}$ est un corps qui contient $\mathbb{K}$. Il sera bon de préciser que λ est une "valeur propre dans $\mathbb{L}$" de u.

Alors que le polynôme minimal de u est un polynôme facile à obtenir "en théorie", le polynôme caractéristique χ_u est le polynôme annulateur de u le plus facilement calculable dans la pratique.

Théorème 6.4.5 (Théorème de Cayley-Hamilton)**.** *Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors, $\chi_A(A)$ est la matrice nulle. En particulier, μ_A divise χ_A.*

Démonstration. Il suffit de montrer que $\chi_u(u) = 0$ pour tout endomorphisme de E, et donc que pour tout x de E, $\chi_u(u)(x) = 0$. Pour $x = 0$, c'est clair.

On suppose donc x non nul. Soit F le sous-espace u-stable engendré par x et $P := \mu_{u,x}$ le polynôme minimal local de u en x.

Par la proposition 6.2.5, on peut construire une base de F, que l'on complète en une base de E, telle que la matrice de u dans cette base est de la forme $\left(\begin{smallmatrix} C_P & B \\ 0 & C \end{smallmatrix}\right)$.

En calculant χ_u dans cette base, on obtient $\chi_u = \chi_{C_P}\chi_C$. Si on montre que $\chi_{C_P} = P$, alors on a achevé la preuve. En effet, on a alors :

$$\chi_u(u)(x) = \chi_C(u)(\chi_{C_P}(u)(x)) = \chi_C(u)(P(u)(x)) = 0,$$

puisque $P = \mu_{u,x}$.

Lemme 6.4.6. *On a $\chi_{C_P} = P$ pour tout polynôme unitaire P de $\mathbb{K}[X]$.*

Démonstration.

En changeant la première ligne L_1 par $L_1 + XL_2 + \cdots + X^{d-1}L_d$, et en développant le déterminant par rapport à la première ligne obtenue, on obtient

$$\chi_{C_P} = \begin{vmatrix} X & 0 & \cdots & 0 & 0 & a_0 \\ -1 & X & \cdots & 0 & 0 & a_1 \\ 0 & -1 & \ddots & 0 & 0 & a_2 \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & -1 & X & a_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & X + a_{n-1} \end{vmatrix} = (-1)^{n-1}P(-1)^{n-1} = P$$

◇

◇

Vu que le polynôme minimal d'un endomorphisme, par construction, divise tout polynôme annulateur cet endomorphisme, le corollaire suivant est immédiat.

Corollaire 6.4.7. *Le polynôme minimal d'un endomorphisme divise son polynôme caractéristique.*

On peut dire mieux sur la matrice compagnon.

Remarque 6.4.8. [Pour tout savoir sur la matrice compagnon]

On peut dire mieux : une matrice compagnon C_P vérifie $\mu_{C_P} = \chi C_P = P$.

La preuve est assez simple. Soit e_1 le premier vecteur de la base canonique, on voit rapidement sur les colonnes de C_P que pour tout entier naturel $k < n$,

$$X^k(C_P)(e_1) = C_P^k(e_1) = e_{k+1}.$$

Comme les e_i forment une partie libre, cela signifie que le polynôme minimal local de C_P en e_i est de degré au moins n, et donc, n puisqu'il divise le polynôme caractéristique. On a donc $\mu_{C_P,e_1} = \chi_{C_P} = P$. Or, μ_{C_P} divise le polynôme caractéristique et il est divisible par μ_{C_P,e_1}. Donc, $\mu_{C_P} = P$.

Le théorème de Cayley-Hamilton peut encore être amélioré :

Corollaire 6.4.9. *[Cayley-Hamilton avec protections renforcées]*
Soit u un endomorphisme d'un espace E de dimension n, alors

$$\mu_u \mid \chi_u \mid \mu_u^n.$$

Démonstration.

Il suffit de prouver la seconde inégalité.

On considère l'identité remarquable suivante dans l'anneau $\mathbb{K}[X, Y]$ à deux indéterminées X et Y.

$$Y^m - X^m = (Y - X)(Y^{n-1} + Y^{n-2}X + \cdots + YX^{n-2} + X^{n-1}).$$

Soit P un polynôme à une indéterminée à coefficients dans $\mathbb{K}$. Par combinaisons linéaires de l'identité suivante, et en mettant en facteur $(Y - X)$, il vient

$$P(Y) - P(X) = (Y - X)Q(Y, X),$$

pour un polynôme Q de $\mathbb{K}[X, Y]$ (qui dépend bien sûr de P).

On va considérer $\mathbb{K}[X, Y]$ comme l'anneau des polynômes en l'indéterminée Y, à coefficients dans l'anneau $\mathbb{K}[X]$, c'est-à-dire par l'isomorphisme canonique $\mathbb{K}[X, Y] \simeq \mathbb{K}[X][Y]$.

Ceci va nous permettre d'évaluer Y en l'endomorphisme u.

$$P(u) - P(X)\,\mathrm{Id} = (u - X\,\mathrm{Id})Q_X(u),$$

où $Q_X(Y) = Q(X, Y)$ vu dans $\mathbb{K}[X][Y]$.

Posons maintenant $P = \mu_u$. Il vient

$$-\mu_u(X)\,\mathrm{Id} = (u - X\,\mathrm{Id})Q_X(u),$$

pour le polynôme Q_X correspondant à μ_u.

Il suffit alors d'utiliser la multiplicativité du déterminant, et l'affaire est dans sac ! ◇

On retrouve ici que toute racine du polynôme minimal est aussi racine du polynôme caractéristique et inversement.

6.5 Multiplicité algébrique et multiplicité géométrique

Rien de mieux que des nombres pour mesurer les problèmes qui nous attendent. Nous voulons, en gros, comprendre des phénomènes géométriques (endomorphismes d'un espace vectoriel), par des moyens algébriques (racines du polynôme minimal, polynôme caractéristique). Mais l'algèbre est-elle soluble dans la géométrie ? Grande question. Nous allons voir dans cette section qu'il y a loin de la coupe aux lèvres.

Définition 6.5.1. Soit λ une valeur propre d'un endomorphisme u, c'est-à-dire une racine du polynôme caractéristique. On appelle multiplicité algébrique de λ la multiplicité $m_a(\lambda)$ de la racine λ dans le polynôme caractéristique χ_u. On appelle sous-espace propre E_λ le noyau de $u - \lambda \mathrm{Id}$, c'est-à-dire, le sous-espace des x de E tels que $u(x) = \lambda x$. Enfin, on appelle multiplicité géométrique de λ la dimension E_λ.

Attention, un vecteur propre est un élément *non nul* du sous-espace propre ! Même si cela peut paraître contre-intuitif, on distinguera vecteur propre (qui se doit d'être non nul), et "élément d'un sous-espace propre" (qui peut éventuellement être nul).

Une première propriété importante des sous-espaces propres est qu'ils sont en somme directe, c'est-à-dire que toute décomposition d'un vecteur en somme d'éléments des sous-espaces propres deux à deux distincts est unique (mais elle n'existe pas forcément pour tout élément de E).

Proposition 6.5.2. *Soit u un endomorphisme de E. Les sous-espaces propres de u sont en somme directe.*

Démonstration. Soient λ_j, $1 \leq j \leq m$, les valeurs propres de u. On veut montrer que si $\sum_{j=1}^m x_j = 0$, avec $x_j \in E_{\lambda_j}$, alors tous les x_j sont nuls. Pour cela, on applique u^i à l'égalité, avec $0 \leq i \leq m-1$. Ce qui donne

$$\sum_{j=1}^{m} \lambda_j^i x_j = 0$$

On obtient un système que l'on peut écrire audacieusement $VX = 0$, où V est la matrice de Vandermonde (λ_j^i), X la matrice colonne $X = (x_i)$, et où le membre de droite est la matrice colonne dont toutes les composantes sont le vecteur nul.

Comme les λ_j sont deux à deux distincts, la matrice V est inversible, ce qui donne $X = 0$, ce qui achève la preuve.

◇

Proposition 6.5.3. *Soit λ une valeur propre de l'endomorphisme u. On a l'inégalité*

$$1 \leq m_g(\lambda) \leq m_a(\lambda).$$

Démonstration. Comme λ est une valeur propre, le sous-espace propre est non nul, ce qui nous donne la première inégalité.

Pour la seconde, on va calculer encore une fois la matrice de u dans une base adaptée. On considère une base du sous-espace E_λ, de dimension $m := m_g(\lambda)$, que l'on complète en une base de l'espace E.

La matrice de l'endomorphisme u s'écrit alors matriciellement sous la forme triangulaire par blocs $\left(\begin{smallmatrix} \mathrm{I}_m & B \\ 0 & A \end{smallmatrix}\right)$.

On obtient donc, $\chi_u = (X-\lambda)^m \chi_A$. Ainsi, λ est racine de χ_u avec multiplicité au moins $m = m_g(\lambda)$. ◇

Il existe une autre multiplicité "parallèle" à la multiplicité géométrique :

Notation 6.5.4. Soit u un endomrophisme de l'espace E. Pour toute valeur propre λ de u, on note $n_a(\lambda)$ la multiplicité de λ comme racine du polynôme caractéristique.

Proposition 6.5.5. *Soit λ une valeur propre de l'endomorphisme u. On a l'inégalité*

$$1 \leq n_a(\lambda) \leq m_a(\lambda).$$

Démonstration. Cela découle directement du corollaire 6.4.9. ◇

Remarque 6.5.6 (Situations extrêmes). On a, dans le monde des matrices carrées, deux situations extrêmes liées à ces inégalités. Et comme toutes les situations extrêmes, elles permettent de donner facilement des contre-exemples.

Tout d'abord la matrice scalaire $\lambda \mathrm{I}_n$. On vérifie que son polynôme minimal est $X-\lambda$, elle a donc une seule valeur propre λ, et le sous-espace propre associé est $E_\lambda = E$. Donc, $n_a(\lambda) = 1$ et $m_g(\lambda) = n$.

A l'opposé, on a la matrice compagnon C_P du polynôme $P = \prod_\lambda (X-\lambda)^{m_a(\lambda)}$. On montre que $\mu_{C_P} = \chi_{C_P} = P$, et donc $n_a(\lambda) = m_a(\lambda)$.

6.6 Diagonalisabilité

C'est probablement le mot sur lequel les candidats ont le plus souvent dérapé à l'oral. En revanche, le concept est souvent bien compris. A l'opposé du dual finalement.

Commençons par une définition naturelle (what else ?)

Définition 6.6.1. Un endomorphisme est dit diagonalisable s'il peut s'écrire matriciellement sous la forme d'une matrice diagonale, dans une certaine base. On dira aussi qu'une matrice carrée est diagonalisable si elle code un endomorphisme diagonalisable, ou, de manière équivalente, si elle est semblable à une matrice diagonale.

On va donner tous les critères de diagonalisabilité pour un endomorphisme, et on laisse le soin au lecteur de transcrire ces critères en termes de matrices.

Théorème 6.6.2. *Soit E un espace de dimension finie n sur un corps $\mathbb{K}$, et u un endomorphisme de E. Les conditions suivantes sont équivalentes :*

i. l'endomorphisme u est diagonalisable,

ii. l'espace E est somme directe de ses sous-espaces propres,

iii. le polynôme caractéristique est scindé sur $\mathbb{K}$ et pour toute valeur propre λ de $\mathbb{K}$, $m_g(\lambda) = m_a(\lambda)$,

iv. le polynôme minimal de u est scindé simple,

v. l'endomorphisme u possède un polynôme annulateur scindé simple.

Démonstration. (i)⇔ (ii). Si E est somme directe de ses sous-espaces propres, on choisit une base de E compatible avec cette somme directe, et u d'écrit diagonalement dans cette base. Réciproquement, si u est diagonalisable, disons que sa matrice est $\mathrm{diag}(\lambda_i)$ dans une base (e_i), alors pour chaque λ que l'on trouve sur la diagonale, on trouve un sous-espace propre engendré par les e_i qui correspondant à cette valeur de λ. On voit que la somme directe des sous-espaces propres est égale à E.

On sait que, pour toute valeur propre λ, $m_g(\lambda) \leq m_a(\lambda)$, et que les sous-espaces propres sont en somme directe. Donc, si u possède k valeurs propres λ_i, $\bigoplus_{i=1}^k E_{\lambda_i} \subset E$.

Dans une base adaptée à la décomposition de E en sous-espaces propres, un calcul simple prouve que

$$\prod_i (X-\lambda_i)^{m_g(\lambda_i)}\chi_u = \prod_i (X-\lambda_i)^{m_a(\lambda_i)}.$$

On en déduit (iii) par unicité de la décomposition en irréductibles.

Réciproquement, si on suppose (iii), $\chi_u = \prod_i (X-\lambda_i)^{m_a(\lambda_i)}$ et

$$\sum_i m_g(\lambda_i) = \sum_i m_a(\lambda_i) = n.$$

Comme $\sum_i m_g(\lambda_i) = \dim \bigoplus_i E_{\lambda_i}$, il vient l'égalité (par inclusion+dimension) $E = \bigoplus_i E_{\lambda_i}$. On a montré (ii)⇔ (iii).

Montrons que (ii) implique (iv). On suppose donc (ii). On sait donc que χ_u est scindé par ce qui précède, disons $\chi_u = \prod_i (X-\lambda_i)^{m_a(\lambda_i)}$. Par le corollaire 6.4.9, $\mu_u = \chi_u = \prod_i (X-\lambda_i)^{k_i}$, avec $1 \leq k_i \leq m_a(\lambda)$.

Or, si l'on fixe i, $X-\lambda_i$ annule l'endomorphisme induit u_i de u sur E_{λ_i}. Donc, tout multiple de $X-\lambda_i$ annule cet induit. Il en résulte que $\prod_i (X-\lambda_i)$ annule u_j pour tout j. Comme E est somme directe de ses sous-espaces propres, on déduit que $\prod_i (X-\lambda_i)$ annule u. Donc, forcément, $k_i = 1$ pour tout i, ce qui prouve (iv).

La réciproque (iv)⇒ (ii) provient du lemme des noyaux appliqué au polynôme annulateur μ_u.

Maintenant, (iv)⇒ (v) est évident et (v)⇒ (iv) vient du fait que μ_u divise tout polynôme annulateur. Or, on sait qu'un diviseur d'un polynôme scindé simple est encore scindé simple (encore une application de la factorialité de $\mathbb{K}[X]$).

◇

Il faut bien voir que l'outil le plus puissant dans la réduction est ce action d'algèbre des polynômes sur les endomorphismes par $P \cdot u = P(u)$, et mieux, si l'endomorphisme u est fixé, sur la géométrie, par $P.x = P(u)(x)$. On le penser ainsi : l'action de l'algèbre des polynôme introduit de l'arithmétique (Bezout, factorialité...) dans la géométrie. Un corollaire assez spectaculaire qui montre la puissance de l'action de l'algèbre des polynômes est le suivant :

Corollaire 6.6.3. *Soit E un espace muni d'un endomorphisme diagonalisable et F un sous-espace stable par u. Alors, l'endomorphisme induit u_F est diagonalisable.*

Démonstration. Comme u est diagonalisable, il possède un polynôme annulateur scindé simple P. Comme P annule u, il annule son induit u_F et donc u_F est diagonalisable. ◇

Vous pouvez essayer chez vous de trouver une preuve sans l'utilisation des polynômes, mais ne tentez jamais cela devant un jury car les pièges sont nombreux. Le piège principal est que l'on pourrait croire, vu que E est somme directe des E_λ, qu'un sous-espace F quelconque est somme directe de ses intersections avec les E_λ. Or, on ne le répètera jamais assez, l'intersection n'est pas distributive sur la somme. Pourtant, cela fonctionne très bien si F est stable par u, car si on note F_λ les sous-espaces propres pour u_F

$$\sum_\lambda (F \cap E_\lambda) = \sum_\lambda F_\lambda = F = F \cap E = F \cap \sum E_\lambda.$$

Corollaire 6.6.4. *Si deux endomorphismes u et v commutent, alors, il existe une base qui les diagonalise simultanément. En particulier, leur somme $u+v$ est diagonalisable.*

Démonstration. Soit $E := \bigoplus_\lambda E_\lambda$ la décomposition en sous-espaces propres pour u. Comme u et v commutent, E_λ est stable par v pour tout λ. En effet, si $x \in E_\lambda$,

$$u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x),$$

ce qui prouve que l'on a bien $v(x) \in E_\lambda$.

Par le corollaire précédent, l'induit de v sur E_λ est encore diagonalisable pour tout λ. Dans une base de diagonalisation de l'induit v_{E_λ}, v est diagonale (par construction), mais u également puisque u est une homothétie (de rapport λ) sur E_λ. En concaténant toutes ces bases quand λ parcourt le spectre de u, on obtient une base qui les diagonalise simultanément.

Cette base est également une base de diagonalisation pour $u+v$.

Contre-exemple 6.6.5. Le premier contre-exemple de matrice non diagonalisable (et valable sur tout corps, est la matrice compagnon de X^2, $J := \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, puisque son polynôme minimal est X^2, qui est scindé, certes, mais non simple.

Il est intéressant de voir que l'on a une condition suffisante de diagonalisabilité.

Proposition 6.6.6. *Supposons que le polynôme caractéristique de u soit scindé simple, alors u est diagonalisable.*

Démonstration. En effet, dans ce cas, on a un polynôme scindé simple annulateur, par Cayley-Hamilton.

Remarque 6.6.7. Sur le corps des complexes, c'est une condition générique, c'est-à-dire, une condition valable sur une partie dense de la topologie normique, voir [**?**, 1.3.35], et de plus, une partie ouverte (non annulation du discriminant du polynôme caractéristique). On termina donc cette section sur la diagonalisabilité sur une note positive : sur $\mathbb{C}$, on est toujours très proche d'une matrice diagonalisable.

6.7 Trigonalisabilité

On voit donc que toute matrice n'est pas forcément diagonalisable et ce, pour plusieurs raisons. 1) tout polynôme unitaire P possède une matrice compagnon C_P associée, et son polynôme minimal est P, 2) tout polynôme n'est pas forcément scindé (si on est sur un

corps non algébriquement clos) comme X^2+1 sur $\mathbb{R}$, 3) un polynôme scindé n'est pas forcément scindé simple (voir le contre-exemple ci-dessus).

On va voir que sur un corps algébriquement clos, tout endomorphisme (ou toute matrice carrée) est trigonalisable, *i.e.* s'écrit sous forme triangulaire dans une base bien choisie. La trigonalisation est tout de même une bonne consolation à la non diagonalisabilité. Par exemple :

i. On peut facilement calculer le déterminant, et le polynôme caractéristique d'une matrice triangulaire, ils valent respectivement $\prod_i d_i$ et $\prod_i (X-d_i)$, où les d_i sont les éléments diagonaux de la matrice.

ii. l'inverse d'une matrice triangulaire (dont les éléments diagonaux sont notés d_i) est encore triangulaire. On voit donc ses valeurs propres avec multiplicité, et ce sont éléments diagonaux.

iii. Si Q est un polynôme et T une matrice triangulaire, alors $Q(T)$ est encore triangulaire avec comme éléments diagonaux les $Q(d_i)$.

iv. Maintenant, si $A = PTP^{-1}$ (A est alors trigonalisable), alors $Q(A) = PQ(T)P^{-1}$. On voit donc toutes les valeurs propres de $Q(A)$: ce sont les images par Q du spectre de A. On peut faire pareil avec, à la place du polynôme Q, une fonction analytique (comme exp et log, par exemple).

On a deux critères de trigonalisabilité : un, de nature géométrique, et un critère algébrique (beaucoup plus puissant !)

Théorème 6.7.1 (Critère géométrique de trigonalisabilité)**.** *Un endomorphisme u d'un espace E, de dimension n, est trigonalisable si et seulement si il existe un drapeau complet stable par u, c'est-à-dire une suite emboîtée de sous-espaces $F_0 \subset F_1 \cdots \subset F_n$, avec $\dim F_k = k$ pour tout k, telle que $u(F_k) \subset F_k$.*

Démonstration. Si $F_0 \subset F_1 \cdots \subset F_n$ est un drapeau complet, on peut construire, par récurrence et en utilisant le théorème de la base incomplète, une base $(e_i)_{1\leq i\leq n}$ de E telle que $(e_i)_{1\leq i\leq k}$ est une base de F_k. Si u stabilise le drapeau, alors $\underline{e}$ est une base qui trigonalise u.

Réciproquement, si $\underline{e}$ est une base qui trigonalise u, alors $F_k := (e_i)_{1\leq i\leq k}$ définit un drapeau complet stable par u. ◇

Théorème 6.7.2 (Critère algébrique de trigonalisabilité)**.** *Un endomorphisme u est trigonalisable si et seulement si une de ces condition est vérifiée :*

i. le polynôme caractéristique χ_u est scindé,

ii. le polynôme minimal μ_u est scindé,

iii. il existe un polynôme scindé qui annule u.

Démonstration. Tout d'abord, il est clair que les trois points sont équivalents. Cela provient du corollaire 6.4.9, du fait que μ_u divise tout polynôme annulateur, et du fait que tout polynôme qui divise un polynôme scindé est lui-même scindé.

Si u est trigonalisable avec les scalaires d_i sur la diagonale, alors $\chi_u = \prod_i (X-d_i)$ et donc, χ_u est scindé.

Montrons donc par récurrence sur la dimension n de l'espace ambiant que si u possède un polynôme annulateur scindé, alors u est trigonalisable.

Pour l'initialisation, c'est clair puisque, en dimension 1, tout endomorphisme est trigonalisable (et même diagonalisable).

Pour l'hérédité, supposons l'assertion vraie dimension $n-1$. Soit Q un polynôme annulateur scindé de u. Comme μ_u divise Q et qu'il est non constant (en dimension > 0), il possède une racine λ, qui est ainsi une valeur propre.

Soit x un vecteur propre associé, il est non nul, et donc peut se prolonger en une base $\underline{e}$, avec $e_1 = x$. Dans cette base u s'écrit matriciellement sous la forme $A := \left(\begin{smallmatrix} \lambda & X \\ 0 & B \end{smallmatrix}\right)$, où X est une matrice ligne et B une matrice carrée de taille $n-1$.

Par un calcul par blocs,

$$0 = Q(A) = \begin{pmatrix} Q(\lambda) & Y \\ 0 & Q(B) \end{pmatrix}.$$

Il vient $Q(B) = 0$ (et $Q(\lambda) = 0$, mais ça on s'en doutait déjà !). Donc, B est trigonalisable et il existe une matrice inversible P_{n-1} de taille $n-1$ telle que $T_{n-1} := P_{n-1}BP_{n-1}^{-1}$ est triangulaire supérieure.

On construit alors la matrice de taille n : $P_n = \left(\begin{smallmatrix} 1 & 0 \\ 0 & P_{n-1} \end{smallmatrix}\right)$. Par un calcul par blocs, on voit que $T_n := P_nAP_n^{-1}$ est triangulaire. La matrice A, et donc, u, est trigonalisable.

◇

Remarque 6.7.3 (Plaidoyer pour les espaces-quotient). On peut voir une preuve alternative plus canonique en utilisant les structures quotient. On utilise le fait que si un sous-espace F de E est stable par u, alors, on a un endomorphisme coinduit $u_{E/F}$ sur l'espace quotient E/F, défini par $u_{E/F}(x+F) = u(x)+F$, et le fait que F est stable par u implique que cet endomorphisme est bien défini. La matrice B est tout simplement une matrice de $u_{E/F}$, avec $F = \langle x \rangle$.

Remarque 6.7.4. Quand on travaille sur $\mathbb{C}$, et plus généralement sur un corps algébriquement clos, tout endomorphisme est trigonalisable. Mais quand on travaille sur $\mathbb{R}$, on peut voir toutes les matrices dans $\mathbb{C}$, sans changer le polynôme caractéristique, puisque le déterminant est le même sur $\mathbb{R}$ et sur $\mathbb{C}$. DE la même manière, tout corps peut être plongé dans un corps algébriquement clos, ce qui est bien pratique pour la trigonalisation.

6.8 La décomposition de Dunford

Nous allons travailler dans le cadre où l'endomorphisme u de E est trigonalisable, c'est-à-dire, quand le polynôme caractéristique est scindé. Notons tout de même que certains théorèmes (hors programme mais bien connus) nous permettent de toujours travailler dans ce cadre.

Théorème 6.8.1. *Deux matrices de $\mathcal{M}_n(\mathbb{R})$ sont $\mathrm{GL}_n(\mathbb{C})$-semblables si et seulement si elles sont $\mathrm{GL}_n(\mathbb{R})$-semblables.*

Et dans le même ordre d'idée :

Théorème 6.8.2. *Soit $\mathbb{K} \subset \mathbb{L}$ deux corps. Deux matrices de $\mathcal{M}_n(\mathbb{K})$ sont $\mathrm{GL}_n()$-semblables si et seulement si elles sont $\mathrm{GL}_n(\mathbb{K})$-semblables.*

Supposons donc le polynôme caractéristique de u scindé. Le lemme des noyaux ramène alors l'étude d'un endomorphisme u à son étude locale sur $K_\lambda = \mathrm{Ker}(u - \lambda \mathrm{I}_n)^{k_\lambda}$.

Dans cette étude locale, les choses sont plus simples (et c'est le cas de dire) puisque l'induit u_λ de l'endomorphisme u sur K_λ est annulé par le polynôme $(X - \lambda)^{k_\lambda}$. Ainsi, le polynôme minimal de u_λ est une puissance de $(X - \lambda)$, et donc λ est l'unique valeur propre de u_λ.

Si l'on note $d_\lambda = \lambda \,\mathrm{Id}$ et $n_\lambda = u_\lambda - d_\lambda$ les endomorphismes de K_λ. Alors, on voit que $u_\lambda = d_\lambda + \lambda$. L'endomorphisme u_λ se décompose en un endomorphisme scalaire (en l'occurence d_λ) et un endomorphisme nilpotent, n_λ. En effet, comme la seule valeur propre de u_λ est λ, la seule valeur propre de n_λ est 0, ce qui implique, par Cayley-Hamilton, que n_λ est bien nilpotent.

On recompose maintenant l'endomorphisme u à l'aide de la décomposition du lemme des noyaux. On va écrire $u = \oplus_\lambda u_\lambda$, ce qui signifie que $u(x) = \sum_\lambda u(x_\lambda) = \sum_\lambda u_\lambda(x_\lambda)$, où $x = \sum_\lambda x_\lambda$ est la décomposition sur les sous espaces caractéristiques K_λ de x. On construit alors $d := \oplus_\lambda d_\lambda$ et $n = \oplus_\lambda n_\lambda$, pour obtenir $u = d + n$.

L'endomorphisme n est nilpotent car il est nilpotent sur chaque K_λ. L'endomorphisme d est diagonalisable car il est scalaire sur chaque K_λ. On a donc décomposé u en un endomorphisme diagonalisable et un endomorphisme nilpotent.

On va montrer que l'on a mieux que ça car, malheureusement, la décomposition d'un endomorphisme en un endomorphisme diagonalisable et un nilpotent n'est pas unique.

Contre-exemple 6.8.3.

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

sont deux décompositions en une matrice diagonalisable et une matrice nilpotente.

En fait, le petit plus (et même le gros plus) qu'apporte la décomposition $u = d + n$ que l'on vient de trouver vérifie que n et d commutent. En effet, par construction $d(x) = \sum_\lambda \lambda x_\lambda$, et donc $d = \sum_\lambda \lambda \Pi_\lambda$, où les Π_λ sont les projecteurs spectraux. Or, ceux-ci sont des polynômes en u, donc, d est un polynôme en u. Il vient que n est également un polynôme en u, ce qui fait que d et n commutent.

Théorème 6.8.4 (Décomposition de Dunford)**.** *Soit u un endomorphisme d'un espace E de dimension finie dont le polynôme caractéristique est scindé. Alors, il existe une unique paire d'endomorphismes (d, n) avec*

i. d diagonalisable,

ii. n nilpotent,

iii. d et n commutent.

De plus, d et n sont des polynômes en u.

Démonstration. On vient de montrer l'existence et la dernière assertion. Montrons l'unicité.

On suppose $d+n=d'+n'$, avec d' et n' qui commutent et d, n les deux polynômes en u que nous venons de trouver. Comme d' commute avec d' et avec n', d' commute avec $d'+n'=u$. Donc d' commute avec d, puisque d est polynomial en u. De même, n' commute avec n. Donc $d-d'=n'-n$, avec $d-d'$ diagonalisable, par le corollaire 6.6.4 et $n'-n$ nilpotent. En effet, si n^k et n'^k s'annulent simultanément à partir d'un certain k_0, alors, pour $m \geq 2k_0$ $n^k n'^{m-k}$ est toujours nul (car un des deux facteurs est nul, et donc $(n-n')^m$ est nul par le binôme de Newton.

Maintenant, un endomorphisme v diagonalisable et nilpotent est nul. En effet, ses valeurs propres sont toutes égales à 0, donc sa matrice diagonalisée est nulle ; l'endomorphisme v est donc nul.

On a donc $d=d'$ et $n=n'$, d'où l'unicité.