Agrégation : algèbre linéaire

Cyril Demarche

3 novembre 2020

1 Introduction

L'algèbre linéaire est l'étude des espaces vectoriels, et des applications linéaires entre espaces vectoriels. Elle permet par exemple de modéliser et analyser les phénomènes linéaires en mathématiques et dans les autres sciences. On peut voir les notes qui suivent comme une généralisation en dimension supérieure de la notion de proportionnalité en dimension 1.

2 Espaces vectoriels et applications linéaires : généralités

2.1 Définitions et premiers exemples

On commence par définir les objets d'étude (les espaces vectoriels), puis les morphismes entre ces objets (les applications linéaires) :

Définition 2.1. Soit K un corps. Un K-espace vectoriel est la donnée d'un triplet $(E, +, \cdot)$, où E est un ensemble, muni d'une loi de composition interne $+ : E \times E \to E$ et d'une loi de composition externe $\cdot : K \times E \to E$, de sorte que

1. $(E, +)$ est un groupe commutatif.
2. pour tous $(\lambda, \mu) \in K^2$, $(x, y) \in E^2$, $(\lambda+\mu)\cdot(x+y) = \lambda\cdot x+\mu\cdot y$ et $(\lambda\mu)\cdot x = \lambda\cdot(\mu\cdot x)$.

On définit la notion de sous-espace vectoriel :

Définition 2.2. Soit E un K-espace vectoriel. Une partie $F \subset E$ est un sous-espace vectoriel si elle est non vide et stable par $+$ et $\cdot$, i.e. F est un sous-groupe de $(E, +)$ stable par multiplication par K.

Exemples 2.3.
- Un ensemble a un élément est naturellement un K-espace vectoriel, appelé espace vectoriel trivial ou nul.
- K lui-même est naturellement un K-espace vectoriel.
- Plus généralement, pour tout ensemble X, K^X (l'ensemble des fonctions de X dans K) est un K-espace vectoriel. Par exemple, pour tout entier $n \geq 0$, K^n est un K-espace vectoriel.
- Si L est un corps contenant K comme sous-corps (L est une extension de K), alors L est naturellement un K-espace vectoriel.
- Si E est un K-espace vectoriel et X un ensemble, E^X est un K-espace vectoriel.
- L'ensemble des fonctions continue de $\mathbf{R}$ dans $\mathbf{R}$ est un sous-espace vectoriel du $\mathbf{R}$-espace vectoriel $\mathbf{R}^{\mathbf{R}}$.

— ? ?

Comme en théorie des groupes, il est fondamental de considérer les morphismes d'espaces vectoriels : ce sont les applications entre espaces vectoriels qui respectent les structures précédentes, à savoir l'addition des vecteurs et la multiplication par un scalaire.

Définition 2.4. Soient E et F deux K-espaces vectoriels. Une application $f; E \to F$ est dite linéaire si pour tous $\lambda \in K$, $(x, y) \in E^2$, $f(\lambda \cdot x + y) = \lambda \cdot f(x) + f(y)$.

En particulier, on voit que $f(0) = 0$.

Remarque 2.5. Le symbole + (resp. $\cdot$) désigne deux lois différentes sur deux ensembles différents. Sauf situation exceptionnelle, on notera abusivement de la même façon (+ et $\cdot$) les lois dans tous les espaces vectoriels considérés.

Définition 2.6. Soient E et F deux K-espaces vectoriels. On note $\mathcal{L}(E, F)$ ou $\mathrm{Hom}_K(E, F)$ l'ensemble des applications linéares de E dans F.

Lemme 2.7. *Soient E et F deux K-espaces vectoriels. Alors $\mathcal{L}(E, F)$ est un sous-espace vectoriel de F^E.*

De façon équivalente, la somme de deux applications linéaires et le produit d'une application linéaire par un scalaire, sont des applications linéaires.

Démonstration: Soient $f, G : E \to F$ deux applications linéaires, et $\lambda \in K$. Montrons que $\lambda \cdot f + g : E \to F$ est une application linéaire. On rappelle que par définition, pour tout $x \in E$, $(\lambda \cdot f + g)(x) = \lambda \cdot f(x) + g(x)$. On a donc, pour tout $(x, y) \in E^2$, tout $\mu \in K$,

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

On utilise alors la linéarité de f et g :

$$\begin{aligned}(\lambda \cdot f + g)(\mu \cdot x + y) &= \lambda \cdot (\mu \cdot f(x) + f(y)) + \mu \cdot g(x) + g(y)\\ &= \mu \cdot (\lambda \cdot f(x) + g(x)) + (\lambda \cdot f(y) + g(y))\\ &= \mu \cdot (\lambda \cdot f + g)(x) + (\lambda \cdot f + g)(y)\,,\end{aligned}$$

ce qui assure le résultat. □

Un peu de vocabulaire maintenant :

Définition 2.8. Soit $f : E \to F$ une application linéaire. On dit que f est un
— endomorphisme si $E = F$.
— isomorphisme si f est bijective.
— automorphisme si f est un endomorphisme et un isomorphisme.

Introduisons maintenant deux sous-espaces vectoriels importants associés à une application linéaire :

Définition 2.9. Soit $f : E \to F$ une applcation linéaire.
— Le noyau de f, noté $\ker(f)$, est l'ensemble $\ker(f) := \{x \in E : f(x) = 0\} \subset E$.
— L'image de f, notée $\mathrm{Im}\,(f)$, est l'ensemble $\mathrm{Im}\,(f) := \{f(x) : x \in E\} \subset F$.
Alors $\ker(f)$ est un sous-espace vectoriel de E et $\mathrm{Im}\,(f)$ est un sous-espace vectoriel de F.

Le lemme suivant est facile, mais très utile :

Lemme 2.10. *Soit $f : E \to F$ une application linéaire.*

1. *f est injective si et seulement si* $\ker(f) = \{0\}$.
2. *f est surjective si et seulement si* $\mathrm{Im}\,(f) = F$.

Démonstration: Seul le sens réciproque du premier point nécessite une démonstration. Supposons donc $\ker(f) = \{0\}$. Soient $(x, y) \in E^2$ tels que $f(x) = f(y)$. Alors par linéarité $f(x - y) = 0$, donc $x - y \in \ker(f)$, donc $x = y$. Donc f est injective. □

2.2 Produit, somme directe et quotient

On propose maintenant quelques méthode fondamentales de construction d'espaces vectoriels. Commençons par le produit d'espaces vectoriels :

Définition 2.11. Soit I un ensemble et $(E_i)_{i\in I}$ une famille de K-espaces vectoriels. On munit l'ensemble produit $E := \prod_{i\in I} E_i$ d'une structure naturelle de K-espace vectoriel en posant, pour tous $(x, y) \in E^2$ et $\lambda \in K$, pour tout $i \in I$,

$$\begin{cases} x + y := (x_i + y_i)_{i\in I} \\ \lambda \cdot x := (\lambda \cdot x_i)_{i\in I}\,. \end{cases}$$

L'espace vectoriel ainsi obtenu est appelé espace vectoriel produit des E_i.

On peut vérifier que l'espace vectoriel $E = \prod_{i\in I} E_i$, muni des projections (linéaires) $p_i : E \to E_i$ vérifie la propriété universelle suivante : pour tout K-espace vectoriel F, et toute famille $f_i : F \to E_i$ d'applications linéaires, il existe une unique application linéaire $f : F \to E$ telle que pour tout $i \in I$, $f_i = p_i \circ f$. Autrement dit, on a un isomorphisme canonique de K-espaces vectoriels

$$\mathcal{L}\left(F, \prod_i E_i\right) \xrightarrow{\sim} \prod_i \mathcal{L}(F, E_i)\,.$$

En outre, cette propriété caractérise l'espace vectoriel $\prod_i E_i$ (muni de ses projections) à unique isomorphisme près.

On introduit maintenant la somme directe de K-espaces vectoriels :

Définition 2.12. Soit I un ensemble et $(E_i)_{i\in I}$ une famille de K-espaces vectoriels. On définit $E := \bigoplus_{i\in I} E_i$ comme le sous-ensemble de $E := \prod_{i\in I} E_i$ formé des éléments $(x_i)_{i\in I}$ tels qu'il existe $J \subset I$ fini vérifiant $x_i = 0$ pour tout $i \in I \setminus J$. Alors E est un sous espace vectoriel de $\prod_{i\in I} E_i$, appelée somme directe de E_i.

Remarque 2.13. On voit tout de suite que si l'ensemble I est fini, alors $\bigoplus_{i\in I} E_i = \prod_{i\in I} E_i$.

On peut vérifier que l'espace vectoriel $E = \bigoplus_{i\in I} E_i$, muni des morphismes injectifs (linéaires) $j_i : E_i \to E$ vérifie la propriété universelle suivante : pour tout K-espace vectoriel F, et toute famille $f_i : E_i \to F$ d'applications linéaires, il existe une unique application linéaire $f : E \to f$ telle que pour tout $i \in I$, $f_i = f \circ j_i$. Autrement dit, on a un isomorphisme canonique de K-espaces vectoriels

$$\mathcal{L}\left(\bigoplus_i E_i, F\right) \xrightarrow{\sim} \prod_i \mathcal{L}(E_i, F)\,.$$

En outre, cette propriété caractérise l'espace vectoriel $\bigoplus_i E_i$ (muni des morphismes j_i) à unique isomorphisme près.

Mentionnons également la construction des espaces vectoriels quotients :

Définition 2.14. Soit E un K-espace vectoriel et $F \subset E$ un sous-espace vectoriel. Alors le groupe quotient E/F est naturellement muni d'une structure de K-espace vectoriel, et la projection $\pi : E \to E/F$ est une application linéaire.

Il suffit en effet de définir le produit $K \times E/F \to E/F$, et pour cela on vérifie que la formule

$$\lambda \cdot \bar{x} := \overline{\lambda \cdot x}$$

est bien définie.

Comme en théorie des groupes, l'espace vectoriel quotient E/F et l'application π sont caractérisés par la propriété universelle suivante :

Proposition 2.15. *Pour tout K-espace vectoriel G et toute application linéaire $f : E \to G$, si $\ker(f) \subset F$, alors il existe une unique application linéaire $\overline{f} : E/F \to G$ telle que le diagramme suivant commute :*

$$\begin{array}{ccc} E & \xrightarrow{f} & G \\ \downarrow{\scriptstyle \pi} & \nearrow & \\ E/F & & \end{array} ,$$

i.e. $f = \overline{f} \circ \pi$.

En corollaire de la construction du quotient, on dispose du fameux "théorème du rang", variante du premier théorème d'isomorphisme en théorie des groupes :

Corollaire 2.16. *Soit $f : E \to F$ une application linéaire.*
Alors f induit un isomorphisme $\overline{f} : E/\ker(f) \xrightarrow{\sim} \mathrm{Im}\,(f)$.

2.3 somme de sous-espaces, familles libres, familles génératrices

Introduisons d'abord la notion de sous-espace vectoriel engendré par une partie d'un espace vectoriel :

Définition 2.17. Soit E un K-espace vectoriel et $P \subset E$ une partie.
Le sous-espace vectoriel de E engendré par P est le sous-esapce vectoriel minimal (pour l'inclusion) de E contenant P. On le note $\mathrm{Vect}(P)$ ou $\langle P \rangle$.

Ce sous-espace est bien défini. On peut en donner deux descriptions : la première construction est "externe", "par l'extérieur". On voit facilement que

$$\mathrm{Vect}(P) = \bigcap_{\substack{F \subset E \text{ sev} \\ F \supset P}} F ,$$

où F décrit les sous-espaces vectoriels de E contenant P. La seconde description est "interne" : on peut décrire $\mathrm{Vect}(P)$ comme l'ensemble des combinaisons linéaires (à coefficients dans K) de vecteurs de P, i.e. $\mathrm{Vect}(P) = \{\sum_{i=1}^{n} \lambda_i \cdot x_i : n \in \mathbf{N}, x_i \in P, \lambda_i \in K\}$.
On introduit alors la notation suivante :

Définition 2.18. Soit E un K-espace vectoriel et $(F_i)_{i\in I}$ une famille de sous-espace vectoriels de E. On note $\sum_{i\in I} E_i$, et on appelle somme des sous-espaces E_i, le sous-espace vectoriel de E engendré par la réunion des E_i, i.e. $\sum_{i\in I} E_i := \langle E_i, i \in I\rangle$.

Ainsi, par définition, les éléments de $\sum_{i\in I} E_i$ sont exactement les vecteurs de E qui s'écrivent $\sum_{j\in J} x_j$, avec $J \subset I$ finie et $x_j \in E_j$ pour tout $j \in J$. En revanche, une telle écriture n'est pas unique en général. Pour remédier à ce défaut d'unicité, on s'intéressera particulièrement dans la suite au cas où les E_i sont en somme directe :

Définition 2.19. Avec les notations précédentes, on dit que les E_i, $i \in I$, sont en somme directe lorsque les conditions équivalentes suivantes sont vérifiées :

1. Tout vecteur $x \in \sum_{i\in I} E_i$ s'écrit de façon unique comme $x = \sum_{i\in I} x_i$, avec $x_i \in E_i$ pour tout i et $x_i = 0$ pour presque tout i.
2. Pour tous $(x_i)_{i\in I} \in \prod_i E_i$ tels que $x_i = 0$ pour presque tout i, on a $\sum_{i\in I} x_i = 0$ si et seulement si $x_i = 0$ pour tout $i \in I$.

Dans ce cas, le sous-espace $\sum_{i\in I} E_i$ sera souvent noté $\bigoplus_{i\in I} E_i$. Lorsque $\bigoplus_{i\in I} E_i = E$, on dira que les sous-espaces $(E_i)_{i\in I}$ sont supplémentaires.

Remarque 2.20. On dispose de deux définitions apparemment différentes pour la notation $\bigoplus_{i\in I} E_i$: celle de la définition 2.12 (somme directe externe d'espaces vectoriels) et celle de la définition 2.19 (somme directe interne de sous-espaces vectoriels). Il se trouve que ces deux notions de somme directe sont compatibles. En effet, si $(E_i)_{i\in I}$ est une famille de sous-espaces vectoriels de E, on dispose toujours d'une application linéaire surjective naturelle

$$\bigoplus_{i\in I} E_i \to \sum_{i\in I} E_i\,,$$

où le membre de gauche est défini en 2.12. Il se trouve que les sous-espaces $(E_i)_{i\in I}$ sont en somme directe au sens de 2.19 si et seulement si le morphisme précédent est un isomorphisme. Cela justifie l'utilisation de la même notation $\bigoplus_{i\in I} E_i$ dans deux contextes légérement différents : les deux espaces vectoriels considérés sont canoniquement isomorphes.

Dans le cas particulier de deux sous-espaces vectoriels, les notions précédentes se simplifient légérement :

Lemme 2.21. *Soient F et G deux sous-espaces vectoriels de E.*

Alors F et G sont en somme directe si et seulement si $F \cap G = \{0\}$. De même, F et G sont supplémentaires dans E si et seulement si $F + G = E$ et $F \cap G = \{0\}$.

On s'intéresse maintenant à la notion cruciale de base d'un espace vectoriel. On introduit d'abord les notions de familles libres et de familles génératrices.

Définition 2.22. Soient $(e_i)_{i\in I}$ une famille de vecteurs de E.

On dit que la famille est libre (resp. génératrice) si la somme $\sum_{i\in I} K \cdot e_i$ est directe (resp. si $\sum_{i\in I} K \cdot e_i = E$).

Proposition 2.23. *Soient $(e_i)_{i\in I}$ une famille de vecteurs de E.*

— *La famille est libre si et seulement si pour tous $(\lambda_i) \in K^{(I)}$, si $\sum_{i\in I} \lambda_i \cdot e_i = 0$ alors $\lambda_i = 0$ pour tout $i \in I$.*

— *La famille est génératrice si et seulement si tout vecteur de E s'écrit comme combinaison linéaire des e_i, $i \in I$.*

Définition 2.24. Une base de E est une famille libre et génératrice.

Proposition 2.25. *Soient $(e_i)_{i\in I}$ une famille de vecteurs de E.*

Cette famille est une base si et seulement si pour tout vecteur $x \in E$, il existe un unique $(x_i)_{i\in I} \in K^{(I)}$ tel que $x = \sum_{i\in I} x_i \cdot e_i$.

Les $(x_i)_{i\in I}$ sont appelées les coordonnées de x dans la base $(e_i)_{i\in I}$.

Corollaire 2.26. *Soit $\mathcal{B} = (e_i)_{i\in I}$ une base de E. Le morphisme naturel défini par $\mathcal{B}$:*

$$K^{(I)} \to E$$

est un isomorphisme de K-espaces vectoriels.

Ce morphisme est défini par $(x_i) \mapsto \sum_i x_i \cdot e_i$. Sa réciproque est l'application qui associe à un vecteur $x \in E$ ses coordonnées (x_i) dans la base $\mathcal{B}$.

Remarquons aussi comment les applications linéaires transforment familles libres, familles génératrices et bases :

Proposition 2.27. *Soient E et F deux K espaces vectoriels, $u \in \mathcal{L}(E, F)$. On suppose que E admet des bases.*

1. *L'application u est injective si et seulement si pour toute famille libre $\mathcal{B}$ de E, la famille $u(\mathcal{B})$ est libre dans F si et seulement si toute (resp. une) base de E est envoyée sur une famille libre de F.*
2. *L'application u est surjective si et seulement si pour toute (resp. pour une) famille génératrice $\mathcal{B}$ de E, la famille $u(\mathcal{B})$ est génératrice si et seulement si toute (resp. une) base de E est envoyée sur une partie génératrice de F.*
3. *L'application u est un isomorphisme si et seulement si pour toute (resp. il existe une) base $\mathcal{B}$ de E, $u(\mathcal{B}$ est une base de F.*

2.4 Sous-espaces vectoriels associés à un endomorphisme

Nous avons déjà introduits les notions de noyau et d'image d'un endomorphisme.

Définition 2.28. Soit E un K-espace vectoriel et $u \in \mathcal{L}(E)$ un endomorphisme.

Un sous-espace vectoriel F de E est dit stable par u si $u(F) \subset F$.

Par exemple, $\ker(u)$ et $\mathrm{Im}\,(u)$ sont des sous-espaces stables par u. Introduisons d'autres sous-espaces stables, très utiles pour la réduction des endomorphismes.

Définition 2.29. Soit $u \in \mathcal{L}(E)$, $\lambda \in K$.

— On dit que λ est valeur propre de u s'il existe $x \in E \setminus \{0\}$ tel que $u(x) = \lambda \cdot x$.
— Si λ est valeur propre de u, on note $E_\lambda := \ker(u - \lambda \cdot \mathrm{id})$. C'est un sous-espace vectoriel non nul de E, appelé espace propre de u associé à la valeur propre λ.
— Un vecteur propre associé à λ est un vecteur non nul de E_λ.

Remarquons que les espaces propres de u sont stables par u.

3 Espaces vectoriels de dimension finie ; matrices, dualité, rang, déterminant

Dans cette partie, on s'intéresse spécifiquement aux espaces de dimension finie, à leurs propriétés et à leurs endomorphismes.

3.1 Théorie de la dimension

On cherche ici à définir la dimension d'un espace vectoriel. On se limitera à la dimension finie.

Définition 3.1. Soit E un K-espace vectoriel.
On dit que E est de dimension finie s'il admet une partie génératrice finie.

Exemples 3.2.
1. L'espace K^n est de dimension finie.
2. Le K-espace vectoriel $K[X]$ des polynômes n'est pas de dimension finie.
3. ? ?

On montre maintenant que les espaces de dimension finie admettent des bases, et que toutes les bases d'un espace ont le même cardinal. On commence par un résultat immédiat.

Proposition 3.3. *Soit E un espace vectoriel de dimension finie.*
De toute famille génératrice finie de E, on peut extraire une base de E. En particulier, E admet une base.

Démonstration: Soit $(x_1, \dots, x_n)$ une famille génératrice.
Si la famille est libre, c'est une base et la preuve est terminée. Sinon, on peut exprimer l'un des vecteurs de la famille, par exemple x_n, comme combinaison linéaires des autres. On en déduit que la famille $(x_1, \dots, x_{n-1})$ est génératrice. On répète ce processus jusqu'à arriver à une famille libre, qui reste génératrice. Donc c'est une base. □

Lemme 3.4. *Soit E un espace de dimension finie, et $(x_1, \dots, x_n)$ une famille génératrice.*
Soient $y_1, \dots, y_{n+1}$ des vecteurs de E. Alors la famille $(y_1, \dots, y_{n+1})$ est liée.

Démonstration: On raisonne par récurrence sur n.
- Si $n = 1$, soit y_1 ou y_2 est nul, soit il existe λ_1 et λ_2 non nuls tels que $y_1 = \lambda_i \cdot x_1$. Alors $\lambda_2 \cdot y_1 - \lambda_1 y_2 = 0$. Dans les deux cas, la famille (y_1, y_2) est donc liée.
- On suppose maintenant $n > 1$ et l'hypothèse de récurrence connue pour un espace engendré par $n-1$ vecteurs. Puisque $(x_1, \dots, x_n)$ est génératrice, il existe des scalaires $a_{i,j} \in K$ tels que pour tout i,

$$y_i = \sum_{j=1}^{n} a_{i,j} \cdot x_j \,.$$

Si $y_1 = 0$, la conclusion est évidente. On peut donc supposer $y_1 \neq 0$. Quitte à permuter les x_i, on peut supposer le coefficient $a_{1,1} \neq 0$. Alors la famille $\left(y_i - \frac{a_{i,1}}{a_{1,1}} \cdot y_1\right)_{2 \leq i \leq n+1}$ est une famille de n vecteurs dans le sous-espace vectoriel

$F := \mathrm{Vect}(x_2, \ldots, x_{n-1})$, lequel est engendré par $n-1$ vecteurs. Par hypothèse de récurrence, cette famille est liée, il existe donc $(\lambda_2, \ldots, \lambda_{n+1}) \in K^n \setminus \{0\}$ tel que $\sum_{i=2}^{n+1} \lambda_i \cdot \left(y_i - \frac{a_{i,1}}{a_{1,1}} \cdot y_1\right) = 0$. Donc

$$-\left(\sum_{i=2}^{n+1} \lambda_i \frac{a_{i,1}}{a_{1,1}}\right) y_1 + \sum_{i=2}^{n+1} \lambda_i \cdot y_i = 0\,.$$

Donc la famille $(y_1, \ldots, y_{n+1})$ est liée.

Le principe de récurrence permet de conclure la preuve. □

Lemme 3.5. *Soit E un espace vectoriel de dimension finie.*
Toute famille libre de E se complète en une base de E.

Démonstration: On sait qu'il existe une base $(e_1, \ldots, e_n)$ de E.

Soit $(x_1, \ldots, x_r)$ une famille libre. Si elle est génératrice, c'est terminé. Sinon, il existe $x_{r+1} \in E \setminus \mathrm{Vect}(x_1, \ldots, x_r)$. Vérifions que $(x_1, \ldots, x_{r+1})$ est libre. Soient (λ_i) tels que $\sum_{i=1}^{r+1} \lambda_i \cdot x_i = 0$. Si $\lambda_{r+1} = 0$, alors comme $(x_1, \ldots, x_r)$ est libre, on en déduit que tous les λ_i sont nuls. Si $\lambda_{r+1} \neq 0$, on peut écrire x_{r+1} comme combinaison linéaire de $(x_1, \ldots, x_r)$, i.e. $x_{r+1} \in \mathrm{Vect}(x_1, \ldots, x_r)$, ce qui est contradictoire. Donc les λ_i sont tous nuls et $(x_1, \ldots, x_{r+1})$ est libre.

On peut répéter le processus tant que la famille obtenue n'est pas génératrice. Puisque E admet une famille génératrice de cardinal n, le lemme 3.4 assure que $r \leq n$ et qu'après au plus $n - r$ étapes, on arrive à une famille $(x_1, \ldots, x_r, \ldots, x_p)$ (avec $r \leq p \leq n$) qui est libre et génératrice, donc une base. □

Théorème 3.6. *Soit E un espace vectoriel de dimension finie.*

1. *Toutes les bases de E ont même cardinal.*
2. *Les bases de E sont exactement les familles génératrices de cardinal minimal (resp. les familles libres de cardinal maximal).*

Démonstration:

1. C'est une conséquence du lemme 3.4.
2. C'est une conséquence de la propostion 3.3 et du lemme 3.4 (resp. de la proposition 3.3 et du lemme 3.5).

□

Définition 3.7. Soit E un espace vectoriel de dimension finie.
La dimension de E est le cardinal d'une base de E (donc de toute base de E).

C'est donc aussi le cardinal maximal d'une famille libre, ou encore le cardinal minimal d'une famille génératrice. La proposition suivante est évidente (cf corollaire 2.26) :

Proposition 3.8. *Soit E un K-espace vectoriel de dimension n.*
Alors toute base de E définit un isomorphisme $E \xrightarrow{\sim} K^n$.

Notez bien que cet isomorphisme dépend du choix d'une base, il n'est pas canonique. C'est tout l'intérêt du changement de base que nous verrons plus loin.

Une conséquence utile des rappels précédents :

Proposition 3.9. *Soit E un K-espace vectoriel de dimension finie et F un sous-espace vectoriel de K.*
Alors F est de dimension finie, $\dim(F) \leq \dim(E)$*, et F admet un supplémentaire.*

Démonstration: Notons n la dimension de E. Si $F = \{0\}$, c'est évident. Sinon, il existe $x_1 \in F \setminus \{0\}$. Si x_1 engendre F, alors F est de dimension finie. Sinon, il existe $x_2 \in F \setminus \mathrm{Vect}(x_1)$, et (x_1, x_2) est une famille libre de F. On poursuit ainsi tant que la famille libre $(x_1, \ldots, x_i)$ est libre et pas génératrice. Si on peut faire $n+1$ fois cette construction, on dispose d'une famille libre $(x_1, \ldots, x_{n+1})$ dans F, donc dans E, ce qui contredit le lemme 3.4. Donc le processus s'arrête en au plus $k \leq n$ étapes et fournit une base $(x_1, \ldots, x_k)$ de F. Donc F est de dimension finie $k \leq n$. En outre, le lemme 3.5 assure que la famille $(x_1, \ldots, x_k)$ se complète en une base de E, disons $(x_1, \ldots, x_k, \ldots, x_n)$. Alors on vérifie facilement que $G := \mathrm{Vect}(x_{r+1}, \ldots, x_n)$ est un supplémentaire de F dans E. □

Calculons maintenant les dimensions du produit et du quotients d'espaces vectoriels :

Proposition 3.10. *Soient $E_1, \ldots, E_n$ des K-espaces vectoriels de dimension finie.*

Alors $E := E_1 \bigoplus \ldots \bigoplus E_n$ est de dimension finie, et $\dim(E) = \dim(E_1) + \cdots + \dim(E_n)$. *Plus précisément, si pour tout i, $(x_1^i, \ldots, x_{k_i}^i)$ est une base de E_i, alors la famille $(\iota_i(x_j^i); 1 \leq i \leq n, 1 \leq j \leq k_i)$ est une base de E, où $\iota_i : E_i \to E$ est l'inclusion canonique.*

Proposition 3.11. *Soit E un K-espace vectoriel de dimension finie et F un sous-espace vectoriel.*

Alors E/F est de dimension finie, et $\dim(E/F) = \dim(E) - \dim(F)$. *Plus précisément, la surjection canonique $\pi : E \to E/F$ induit un ismorphisme entre tout supplémentaire de F dans E et E/F.*

L'énoncé suivant est une variante du lemme du rang :

Corollaire 3.12. *Soit E un espace vectoriel de dimension finie, et $u \in \mathcal{L}(E)$.*

Alors $\dim(\ker(u)) + \dim(\mathrm{Im}(u)) = \dim(E)$.

Démonstration: C'est la conjonction du corollaire 2.16 et de la proposition 3.11. □

3.2 Matrices d'une application linéaire

Dans cette partie, E désigne un K-espace vectoriel de dimension n.

Comme expliqué plus haut, le choix d'une base $\mathcal{B} := (e_1, \ldots, e_n)$ de E fournit un isomorphisme

$$\varphi_{\mathcal{B}} : E \xrightarrow{\sim} K^n .$$

En particulier, la base étant fixée, les vecteurs de E peuvent être identifiés à des vecteurs de K^n, appelés vecteurs colonnes. Ainsi, un vecteur $x = \sum_i x_i \cdot e_i$ de E est identifié au vecteur colonne de ses coordonnées, à savoir $X = \varphi_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. L'un des intérêts principaux de cette traduction est de faciliter le calcul algébrique avec les vecteurs et les applications linéaires. Remarquons que dans toute la suite, et sauf mention explicite du contraire, on représentera les vecteurs de K^n comme des vecteurs colonnes.

Rappelons d'abord le fait bien connu suivant :

Proposition 3.13. *Soit $A \in \mathrm{M}_{n,p}(K)$.*

Alors A induit une application linéaire $f_A : K^p \to K^n$ définie par le produit matriciel $f_A(X) := A \cdot X$, pour tout $X \in K^n$.

Ainsi une matrice définit-elle une application linéaire. Nous allons maintenant nous intéresser à la correspondance dans l'autre sens, qui est plus riche : il s'agit d'associer à une application linéaire donnée, non pas une matrice, mais une famille de matrices...

Définition 3.14. Soient E et F deux K-espaces vectoriels de dimensions respectives p et n, et $u \in \mathcal{L}(E,F)$. Soit $\mathcal{B} = (e_1, \ldots, e_p)$ (resp. $\mathcal{C} := (f_1, \ldots, f_n)$) une base de E (resp. F).

La matrice de u dans les bases $\mathcal{B}$ et $\mathcal{C}$, notée $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u)$, est la matrice $(a_{i,j}) \in \mathrm{M}_{n,p}(K)$ dont les coefficients sont définis par

$$u(e_j) = \sum_{i=1}^{n} a_{i,j} \cdot f_j \,.$$

Autrement dit, le j-ième vecteur colonne de la matrice de u est formé des coordonnées de $u(e_j)$ dansl la base $\mathcal{C}$.

Remarque 3.15. Si $E = F$ et $\mathcal{B} = \mathcal{C}$, $\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(u)$ sera notée $\mathrm{Mat}_{\mathcal{B}}(u)$.

Un premier résultat évident est fondamental pour la traduction calculatoire des applications linéaires :

Proposition 3.16. *Avec les notations précédentes, en posant $A = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}$, le diagramme suivant est commutatif :*

$$\begin{array}{ccc} E & \xrightarrow{u} & F \\ {\scriptstyle \varphi_{\mathcal{B}}}\downarrow{\scriptstyle \sim} & & {\scriptstyle \sim}\downarrow{\scriptstyle \varphi_{\mathcal{C}}} \\ K^p & \xrightarrow{f_A} & K^n \,. \end{array}$$

Autrement dit, pour tout $x \in E$, si l'on note X (resp. Y) le vecteur colonne des coordonnées de x (resp. $u(x)$) dans la base $\mathcal{B}$ (resp. $\mathcal{C}$), alors $Y = A \cdot X$.

Da façon concrète, une fois E et F identifiés à K^p et K^n par des choix de bases, l'application linéaire u devient une application matricielle $K^p \to K^n$, donnée par la formule $X \mapsto A \cdot X$ pour une certaine matrice A.

En outre, cette correspondance entre applications linéaires et matrices respecte les différentes opérations sur les espaces d'applications linéaires et de matrices :

Proposition 3.17. *Avec les notations précédentes, l'application*

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}} : \mathcal{L}(E,F) \to \mathrm{M}_{n,p}(K)$$

est un isomorphisme de K-espaces vectoriels.

Traduction : si $u, v \in \mathcal{L}(E,F)$ et $\lambda \in K$, alors $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\lambda \cdot u + v) = \lambda \cdot \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u) + \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(v)$.

De la même façon, la composition des applications linéaires va se traduire par une opération algébrique simple sur les matrices :

Proposition 3.18. *Soient E, F, G trois espaces vectoriels, de dimensions respectives m, n, p, et munis de bases respectives $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$. Soient $u \in \mathcal{L}(F,G)$ et $v \in \mathcal{L}(E,F)$.*

Alors

$$\mathrm{Mat}_{\mathcal{D},\mathcal{B}}(u \circ v) = \mathrm{Mat}_{\mathcal{D},\mathcal{C}}(u) \cdot \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(v)\,.$$

Autrement dit, la correspondance entre applications linéaires et matrices transforme la composition des applications linéaires en le produit matriciel, ce qui fournit un outil calculatoire très agréable pour calculer des composées.

Le cas particulier suivant est fondamental, et il résume les deux énoncés précédents dans le cas des endomorphismes :

Proposition 3.19. *Soit E un K-espace vectoriel de dimension n, muni d'une base $\mathcal{B}$. Alors l'application*

$$\mathrm{Mat}_{\mathcal{B}} : \mathcal{L}(E) \to \mathrm{M}_n(K)$$

est un isomorphisme de K-algèbres.

Il est important de retenir que toutes ces identifications entre applications linéaires et matrices dépendent fortement du choix d'une (ou deux) base(s). Il convient donc maintenant d'étudier comment ce choix des bases influe sur les identifications précédentes. C'est la notion de changement de bases.

Définition 3.20. Soit E un K-espace vectoriel de dimension n, muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$.

La matrice de changement de bases (ou matrice de passage) de $\mathcal{B}$ vers $\mathcal{B}'$ est la matrice $P_{\mathcal{B}\to\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}} := \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)$.

Pour obtenir la matrice de passage, il faut écrire en colonnes les coordonnées des vecteurs de $\mathcal{B}$ dans la base $\mathcal{B}'$.

Lemme 3.21. *On a $P_{\mathcal{B},\mathcal{B}'} \cdot P_{\mathcal{B}',\mathcal{B}} = \mathrm{I}_n$ et donc $P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}}^{-1}$.*

Proposition 3.22. *Avec les notations précédentes, si $P := P_{\mathcal{B}\to\mathcal{B}'}$, alors le diagramme suivant est commutatif :*

$$\begin{array}{ccc} E & \xrightarrow{\varphi_{\mathcal{B}}} & K^n \\ & \searrow^{\varphi'_{\mathcal{B}}} & \downarrow{\scriptstyle f_P} \\ & & K^n\,. \end{array}$$

Autrement dit, pour tout vecteur $x \in E$, si X (resp. X') désigne le vecteur colonnes des coordonnées de x dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$), alors $X' = P \cdot X$.

On est désormais en mesure d'énoncer la formule de changement de bases pour les applications linéaires :

Théorème 3.23. *Soient E, F deux K-espaces vectoriels. Soient $\mathcal{B}$, $\mathcal{B}'$ deux bases de E, et $\mathcal{C}$, $\mathcal{C}'$ deux bases de F. Soit $u \in \mathcal{L}(E,F)$. Notons $Q := P_{\mathcal{B}\to\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}}$ et $P := P_{\mathcal{C}\to\mathcal{C}'} = P_{\mathcal{C}',\mathcal{C}}$.*

Alors

$$\mathrm{Mat}_{\mathcal{C}',\mathcal{B}'}(u) = P \cdot \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u) \cdot Q^{-1}\,.$$

En toutes lettres, on a donc

$$\mathrm{Mat}_{\mathcal{C}',\mathcal{B}'}(u) = P_{\mathcal{C}',\mathcal{C}} \cdot \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u) \cdot P_{\mathcal{B},\mathcal{B}'} \,,$$

ce qui peut justifier la notation adoptée pour les matrices de changement de bases.

Définition 3.24. Deux matrices A et A' dans $\mathrm{M}_{n,p}(K)$ sont dites équivalentes s'il existe $P \in \mathrm{GL}_n(K)$ et $Q \in \mathrm{GL}_p(K)$ tels que

$$A' = P \cdot A \cdot Q^{-1} \,.$$

Autrement dit, deux matrices sont équivalentes si et seulement si elles représentent la même application linéaires dans des bases différentes (bases différentes aussi bien dans l'espace de départ que dans l'espace d'arrivée).

On peut reformuler la définition précédente en parlant d'une action du groupe $\mathrm{GL}_n(K) \times \mathrm{GL}_p(K)$ sur $\mathrm{M}_{n,p}(K)$, et en disant que deux matrices sont équivalentes si et seulement si elles sont dans la même orbite.

Le cas particulier le plus utile pour la réduction des endomorphismes est le suivant :

Corollaire 3.25. *Soit E un K-espace vectoriel de dimension n, muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$, et $u \in \mathcal{L}(E)$.*

En notant $P := P_{\mathcal{B}\to\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}}$, on a

$$\mathrm{Mat}_{\mathcal{B}'}(u) = P \cdot \mathrm{Mat}_{\mathcal{B}}(u) \cdot P^{-1} \,.$$

Autrement dit, les matrices associées à u dans deux bases différentes ne sont pas égales, elles sont conjuguées par la matrice $P \in \mathrm{GL}_n(K)$.

Définition 3.26. On dit que deux matrices A et A' dans $\mathrm{M}_n(K)$ sont semblables s'il existe $P \in \mathrm{GL}_n(K)$ telle que

$$A' = P \cdot A \cdot P^{-1} \,.$$

Autrement dit, deux matrices sont semblables si et seulement si elles représentent le même endomorphismes dans des bases différentes (mais avec la même base au départ et à l'arrivée).

3.3 Rang

Dans cette partie, on s'intéresse au rang d'une application linéaire, d'une matrice ou d'un système d'équations.

Définition 3.27. Soit $u \in \mathcal{L}(E, F)$.

Le rang de u, noté $\mathrm{rg}(u)$, est la dimension de $\mathrm{Im}\,(u)$.

On peut donc reformuler le lemme du rang (cf corollaire 3.12) :

Lemme 3.28. *Soit $u \in \mathcal{L}(E, F)$. On suppose E de dimension finie.*

Alors $\dim(\ker(u)) + \mathrm{rg}(u) = \dim(E)$.

On voit immédiatement que $\mathrm{rg}(u) \leq \min(\dim(E), \dim(F))$.

On dispose d'autres notions de rang en algèbre linéaire :

Définition 3.29. Soit E un espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs.

Le rang de la famille $(x_i)_{i\in I}$ est la dimension de $\mathrm{Vect}(x_i, i \in I)$.

Concrètement, le rang est le nombre maximal de vecteurs x_i indépendants, i.e. formant une famille libre. On voit tout de suite que $\mathrm{rg}(A) \leq \min(n, p)$.

Définition 3.30. Soit $A \in \mathrm{M}_{n,p}(K)$. On note C_1, ..., C_p les colonnes de A.

Le rang de la matrice A, noté $\mathrm{rg}(A)$, est la dimension de $\mathrm{Vect}(C_1, \ldots, C_p) \subset \mathbf{R}^n$, i.e. le rang de la famille $(C_1, \ldots, C_p)$.

La propriété suivante est facile, mais rassurante :

Proposition 3.31. *Si* $A = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u)$, *alors* $\mathrm{rg}(A) = \mathrm{rg}(u)$.

Remarquons également que deux matrices équivalentes ont même rang. Montrons la réciproque de cette affirmation. Pour cela, on définit la matrice $J_r \in \mathrm{M}_{n,p}(K)$ par

$$J_r := \left(\begin{array}{c|c} I_r & 0_{r,p-r} \\ \hline 0_{n-r,r} & 0_{n-r,p-r} \end{array}\right).$$

Proposition 3.32. *Soit* $A \in \mathrm{M}_{n,p}(K)$ *de rang* r. *Alors il existe des matrices* $P \in \mathrm{GL}_n(K)$ *et* $Q \in \mathrm{GL}_p(K)$ *telles que*

$$P \cdot A \cdot Q^{-1} = J_r\,.$$

Autrement dit, deux matrices sont équivalentes si et seulement si elles ont même rang.

Démonstration: Il suffit de prendre une base de l'image, une base du noyau, et de compléter... □

Définition 3.33. Soit $A \in \mathrm{M}_{n,p}(K)$. On note $L_1, \ldots, L_n$ les vecteurs lignes de A, dans K^p. Soit (S) un système d'équations linéaires associé à A, i.e. $A \cdot X = B$, pour un certain $B \in K^n$.

Le rang du système (S) est le rang de la famille $(L_1, \ldots, L_n)$ dans K^p, i.e. la dimension de $\mathrm{Vect}(L_1, \ldots, L_n)$.

Le rang du système est donc le nombre maximal d'équations indépendantes dans le système.

Là encore, on voit tout de suite que $\mathrm{rg}(S) \leq \min(n, p)$. En revanche, le lien entre le rang du système d'équations et le rang de la matrice associée n'est pas évident. Il sera expliqué plus bas.

3.4 Dualité

La dualité est un concept fondamental en mathématiques. L'idée générale est que pour étudier un objet, on peut étudier les fonctions naturellement définies sur cet objet. Par exemple, plutôt qu'étudier un groupe abélien fini, on peut étudier ses caractères. Plutôt qu'étudier un groupe (fini), on peut étudier ses représentations linéaires. Ici, pour étudier un espace vectoriel, on va étudier les formes linéaires sur cet espace :

Définition 3.34. Soit E un K-espace vectoriel.

- Une forme linéaire sur E est une application linéaire $E \to K$.
- Le dual de E, noté E^*, est l'espace vectoriel des formes linéaires sur E, i.e. $E^* := \mathcal{L}(E, K)$.

Remarquons qu'en analyse, l'espace vectoriel E est parfois muni d'une topologie ou d'une norme. Dans ce cas, il peut être judicieux de définir un dual plus petit, formé des formes linéaires *continues* sur E.

Remarque 3.35. Supposons E de dimension finie pour simplifier. On rappelle que si $\varphi : E \to K$ est une forme linéaire non nulle, alors $\ker(\varphi)$ est un hyperplan de E. Réciproquement, pour tout hyperplan de E, il existe $\varphi_0 \in E^*$ non nulle telle que $H = \ker(\varphi_0)$, et l'ensemble des $\varphi \in E^*$ telles que $H = \ker(\varphi)$ est exactement l'ensemble des multiples non nuls de φ_0. Par conséquent, on dispose des bijections naturelles suivantes :

$$\{\text{hyperplans de } E\} \xleftarrow{\sim} \{\text{formes linéaires } \neq 0 \text{ sur } E\}/\text{homothéties} = (E^* \backslash \{0\})/K^\times =: \mathbb{P}(E^*)\,,$$

où $\mathbb{P}(F)$ désigne l'espace projectif d'un espace vectoriel F, c'est-à-dire l'ensemble $(F \setminus \{0\})/K^\times$, qui s'identifie à l'ensemble des droites de F. Il y a donc un lien fort entre formes linéaires et hyperplans, que l'on peut résumer via l'identification (canonique) $\mathbb{P}(E^*) \xrightarrow{\sim} \text{hyperplans de } E\}$. Notez par exemple que pour $\dim(E) = 2$, on a une bijection canonique entre les droites de E^* et les droites de E ; pour $\dim(E) = 3$, on a une bijection canonique entre les droites de E^* et les plans de E (et réciproquement) : cette correspondance se traduit en une dualité en géométrie projective qui "échange" droites et plans ...

Notons que l'on dispose, par définition, d'une application bilinéaire cruciale, souvent notée comme un produit scalaire :

$$\langle \cdot, \cdot \rangle : \begin{array}{ccc} E \times E^* & \to & K \\ (x, \varphi) & \mapsto & \langle x, \varphi \rangle := \varphi(x) \end{array} . \tag{1}$$

Exemple 3.36. Considérons l'application $b : \mathrm{M}_{1,n}(K) \times \mathrm{M}_{n,1}(K) \to K$ définie par le produit matriciel, i.e. $b(X, Y) := X \cdot Y$ (on identifie $\mathrm{M}_{1,1}(K)$ et K). Il est clair que cette application est bilinéaire, et qu'elle induit un isomorphisme d'espaces vectoriels

$$\mathrm{M}_{1,n}(K) \xrightarrow{\sim} \mathrm{M}_{n,1}(K)^*\,,$$

identifiant naturellement le dual de l'espace vectoriel des vecteurs colonnes à l'espace vectoriel des vecteurs lignes (de même taille).

Dès lors que l'on choisit d'identifier les vecteurs de K^n avec les vecteurs colonnes, il devient donc naturel d'identifier les vecteurs du dual de K^n (i.e. les formes linéaires sur K^n) aux vecteurs lignes.

Le résultat suivant est immédiat (au moins en dimension finie, en utilisant par exemple la proposition 3.41 qui suit) :

Proposition 3.37. *Soit E un K-vectoriel.*

On dispose d'un morphisme injectif canonique $\varphi : E \to (E^*)^*$ défini par $\varphi(x)(f) = f(x)$ pour tout $x \in E$ et $f \in E^*$.

Si E est de dimension finie, c'est un isomorphisme (dit de bidualité).

Vérifions maintenant que l'opération de dualité, qui à un espace E associe son dual E^*, induit également une opération sur les applications linéaires entre espaces vectoriels : on dit que l'opération de dualité est "fonctorielle".

Définition 3.38. Soit $u \in \mathcal{L}(E, F)$.

L'application linéaire duale (ou transposée) de u, notée ${}^t u$, est l'application ${}^t u \in \mathcal{L}(F^*, E^*)$ définie par ${}^t u(f) := f \circ u$, pour tout $f \in F^*$.

Autrement dit, la dualité associe à une application linéaire $u : E \to F$, une application linéaire duale ${}^t u : F^* \to E^*$.

Lemme 3.39. *Soient $u \in \mathcal{L}(F, G)$ et $v \in \mathcal{L}(E, F)$.*

Alors ${}^t(u \circ v) = {}^t v \circ {}^t u$ dans $\mathcal{L}(G^*, E^*)$.

Cette dualité se comporte bien vis-à-vis des bases :

Définition 3.40. Soit E un K-espace vectoriel muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.

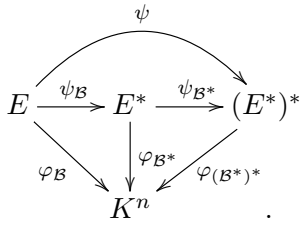
La base duale de $\mathcal{B}$, notée $\mathcal{B}^* = (e_1^*, \ldots, e_n^*)$ est la base de E^* définie par $e_i^*(e_j) = \delta_{i,j}$ pour tout $1 \leq i, j \leq n$.

Il est immédiat de vérifier que $\mathcal{B}^*$ est bien définie, puis que c'est une base de E^*.

Proposition 3.41. *Soit E un K-espace vectoriel de dimension finie.*

Pour toute base $\mathcal{B}$ de E, il existe un isomorphisme naturel $\psi_{\mathcal{B}} : E \xrightarrow{\sim} E^*$, qui envoie $\mathcal{B}$ sur $\mathcal{B}^*$. En particulier, $\dim(E^*) = \dim(E)$.

Là encore, l'isomorphisme entre E et E^* dépend du choix de la base $\mathcal{B}$, il n'est pas canonique - contrairement au morphisme de bidualité. Remarquons que l'on dispose d'un diagramme commutatif :



Remarque 3.42. En utilisant l'isomorphisme de bidualité $\psi_E : E \xrightarrow{\sim} E^{**}$, et le même isomorphisme pour un espace vectoriel F, on vérifie que le diagramme suivant

$$\begin{array}{ccc} E & \xrightarrow{u} & F \\ \downarrow{\scriptstyle \psi_E} & & \downarrow{\scriptstyle \psi_F} \\ E^{**} & \xrightarrow{{}^t({}^t u)} & F^{**} \end{array},$$

ce qui est une version abstraite de l'énoncé matriciel évident ${}^t({}^t A) = A$.

Observons maintenant comment la dualité se comporte vis-à-vis des noyaux et des images :

Proposition 3.43. *Soit $u \in \mathcal{L}(E, F)$. Alors on dispose d'isomorphismes canoniques*

$$\ker({}^t u) \xrightarrow{\sim} (F/\mathrm{Im}\,(u))^*$$

et

$$\mathrm{Im}\,({}^t u) \xrightarrow{\sim} (E/\ker(u))^* \xleftarrow{\sim} \mathrm{Im}\,(u)^*.$$

En particulier, $\mathrm{rg}({}^t u) = \mathrm{rg}(u)$. *Et u est injective (resp. surjective) si et seulement si ${}^t u$ est surjective (resp. injective).*

Traduisons matriciellement cette opération de dualité : soient E et F deux K-espaces vectoriels munis de bases $\mathcal{B}$ et $\mathcal{C}$ respectivement. Soit $u \in \mathcal{L}(E,F)$.

Proposition 3.44. *On a*

$$\mathrm{Mat}_{\mathcal{B}^*,\mathcal{C}^*}({}^t u) = {}^t\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u)\,.$$

Autrement dit, l'application linéaire duale a pour matrice (dans les bases duales) la transposée de la matrice de l'application linéaire initiale. C'est-à-dire que le diagramme suivant (où toutes les applications sont des isomorphismes) commute :

$$\begin{array}{ccc} \mathcal{L}(E,F) & \xrightarrow{{}^t(\cdot)} & \mathcal{L}(F^*,E^*) \\ \Big\downarrow{\scriptstyle \mathrm{Mat}_{\mathcal{C},\mathcal{B}}} & & \Big\downarrow{\scriptstyle \mathrm{Mat}_{\mathcal{B}^*,\mathcal{C}^*}} \\ \mathrm{M}_{n,p}(K) & \xrightarrow{{}^t(\cdot)} & \mathrm{M}_{p,n}(K)\,. \end{array}$$

Ainsi l'opération de transposition des matrices n'est-elle pas une opération arbitraire : elle est la traduction concrète de l'opération très naturelle de dualité.

Une autre façon d'interpréter ce résultat : étant donnée une matrice $A \in \mathrm{M}_{n,p}(K)$, on peut intérpréter les colonnes de cette matrices comme les coordonnées des images par u des vecteurs de la base $\mathcal{B}$ dans la base $\mathcal{C}$, alors que les lignes de la même matrice s'interprètent comme les coordonnées des images par ${}^t u$ des vecteurs de la base $\mathcal{C}^*$ dans la base $\mathcal{B}^*$.

Remarque 3.45. La proposition précédente implique notamment que $P_{\mathcal{B}'^*\to\mathcal{B}^*} = {}^t P_{\mathcal{B}\to\mathcal{B}'}$.

Avec ces traductions, la proposition 3.43 implique le résultat classique suivant :

Corollaire 3.46. *Soit $A \in \mathrm{M}_{n,p}(K)$.*
Alors $\mathrm{rg}({}^t A) = \mathrm{rg}(A)$.

Abordons maintenant la notion d'othogonalité pour la dualité, suggérée notamment par l'accouplement (1).

Définition 3.47. Soit E un K-espace vectoriel et $F \subset E$ une partie de E.
L'orthogonal de F, noté $F^\perp$, est le sous-ensemble de E^* défini par

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

Autrement dit, on a $F^\perp := \{f \in E^* : \forall x \in F\,,\ \langle x, f\rangle = 0\}$, d'où la notation "$\perp$", qui rappelle l'orthogonalité pour un produit scalaire (ou plus généralement pour une forme bilinéaire).

Symétriquement, on peut définir :

Définition 3.48. Soit E un K-espace vectoriel et $G \subset E^*$ une partie de E^*.
L'orthogonal de G, noté G°, est le sous-ensemble de E défini par

$$G^\circ := \{x \in E : \forall f \in G\,,\ f(x) = 0\}\,.$$

Lemme 3.49. *Soient $F_1, F_2 \subset E$ et $G_1, G_2 \subset E^*$.*

1. *Le sous-ensemble $F_1^\perp$ (resp. G_1°) est un sous-espace vectoriel de E^* (resp. de E).*
2. *$(F_1 \cup F_2)^\perp = F_1^\perp \cap F_2^\perp$ et $(G_1 \cup G_2)^\circ = G_1^\circ \cap G_2^\circ$.*

3. $(F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp$ *et* $(G_1 \cap G_2)^\circ = G_1^\circ + G_2^\circ$.
4. $\left(F_1^\perp\right)^\circ = \mathrm{Vect}(F_1)$ *et* $(G_1^\circ)^\perp = \mathrm{Vect}(G_1)$.
5. *Si* $F_1 \subset F_2$ *(resp.* $G_1 \subset G_2$*), alors* $F_2^\perp \subset F_1^\perp$ *(resp.* $G_2^\circ \subset G_1^\circ$*).*

Proposition 3.50. *Si* $F \subset E$ *(resp.* $G \subset E^*$*) est un sous-espace vectoriel, alors on a un isomorphisme canonique* $E^*/F^\perp \xrightarrow{\sim} F^*$ *(resp.* $E/G^\circ \xrightarrow{\sim} G^*$*).*

En particulier, $\dim(F)+\dim(F^\perp) = \dim(E)$ *(resp.* $\dim(G)+\dim(G^\circ) = \dim(E)$*).*

Un outil très utile pour idenfier un espace vectoriel à son dual :

Proposition 3.51. *La donnée d'un morphisme* $\varphi : E \to E^*$ *est équivalente à celle d'une forme bilinéaire sur* E*. Celle-ci est non dégénérée si et seulement si* φ *est un isomorphisme.*

Une application :

Proposition 3.52. *La forme bilinéaire non dégénérée* $\mathrm{M}_n(K) \times \mathrm{M}_n(K) \to K$ *donnée par* $(A, B) \mapsto \mathrm{tr}(A \cdot B)$ *définit un isomorphisme* $\mathrm{M}_n(K) \xrightarrow{\sim} \mathrm{M}_n(K)^*$.

Corollaire 3.53. *Si* $n \geq 2$*, alors tout hyperplan de* $\mathrm{M}_n(K)$ *rencontre* $\mathrm{GL}_n(K)$.

3.5 Pivot de Gauss

L'objectif principal de cette partie est de présenter l'algorithme fondamental du pivot de Gauss, sur un corps. Cet outil permet notamment de déterminer explicitement et efficacement des bases (ou des équations) des noyaux et images d'une matrice, de résoudre des systèmes linéaires, de calculer des déterminants et des inverses de matrices ...

On reviendra plus tard sur le pivot de Gauss dans le cadre plus général des matrices à coefficients dans un anneau commutatif (et surtout dans le cas d'un anneau principal ou euclidien).

On commence par définir les opérations élémentaires sur les matrices. Dans toute cette partie, K est un corps et $A \in \mathrm{M}_{n,p}(K)$ est une matrice dont on note $a_{i,j}$ ses coefficients. On écrira $A = (C_1|\ldots|C_p)$, où les C_i sont des vecteurs colonnes, et $A = \left(\begin{array}{c} L_1 \\ \hline \vdots \\ \hline L_n \end{array}\right)$, où les L_j sont des vecteurs lignes.

Définition 3.54. Soient $1 \leq i \neq j \leq p$, et $\lambda \in K$, $\mu \in K^\times$. On dispose de trois opérations élémentaires sur les colonnes de la matrices A :

1. l'opération $t_{i,j}(\lambda)$, notée $C_i \leftarrow C_i + \lambda \cdot C_j$, qui consiste à remplacer la colone C_i de A par la colonne $C_i + \lambda \cdot C_j$.
2. l'opération $d_i(\mu)$, notée $C_i \leftarrow \mu \cdot C_i$, qui consiste à remplacer la colone C_i de A par la colonne $\mu \cdot C_i$.
3. l'opération $\tau_{i,j}$, notée $C_i \leftrightarrow C_j$, qui consiste à échanger dans la matrice A les colonnes C_i et C_j.

On définit de façon tout à fait analogue les trois opérations élémentaires sur les lignes de A, avec $1 \leq i \neq j \leq n$:

1. l'opération $t'_{i,j}(\lambda)$, notée $L_i \leftarrow L_i + \lambda \cdot L_j$.
2. l'opération $d'_i(\mu)$, notée $L_i \leftarrow \mu \cdot L_i$.
3. l'opération $\tau'_{i,j}$, notée $L_i \leftrightarrow L_j$.

Traduisons immédiatemment ces opérations en un calcul de produit matriciel. Pour cela, on introduit les trois familles de matrices suivantes :

Définition 3.55. Soit $1 \leq i \neq j \leq p$, et $\lambda, \mu \in K$, $\sigma \in \mathfrak{S}_p$.

1. On note $T_{i,j}(\lambda) := I_p + \lambda \cdot E_{i,j} \in \mathrm{M}_p(K)$. Une telle matrice est appelée une matrice de transvection.
2. On note $D_i(\mu) := I_p + (\mu - 1) \cdot E_{i,i} \in \mathrm{M}_p(K)$. Une telle matrice est appelée une matrice de dilatation.
3. On note $P_\sigma \in \mathrm{M}_p(K)$ la matrice de permutation associée à σ, dont le coefficient d'indice (i, j) vaut $\delta_{i,\sigma(j)}$. En particulier, on note $P_{i,j} := P_{(ij)}$.

On dispose alors de la traduction suivante des opérations élémentaires :

1. L'opération $t_{i,j}(\lambda) : C_i \leftarrow C_i + \lambda \cdot C_j$ revient à remplacer la matrice A par la matrice $A \cdot T_{j,i}(\lambda)$.
2. L'opération $d_i(\mu) : C_i \leftarrow \mu \cdot C_i$ revient à remplacer la matrice A par la matrice $A \cdot D_i(\mu)$.
3. L'opération $\tau_{i,j} : C_i \leftrightarrow C_j$ revient à remplacer la matrice A par la matrice $A \cdot P_{i,j}$.

De façon symétrique, quitte à transposer, on obtient :

1. L'opération $t'_{i,j}(\lambda) : L_i \leftarrow L_i + \lambda \cdot L_j$ revient à remplacer la matrice A par la matrice $T_{i,j}(\lambda) \cdot A$.
2. L'opération $d'_i(\mu) : L_i \leftarrow \mu \cdot L_i$ revient à remplacer la matrice A par la matrice $D_i(\mu) \cdot A$.
3. L'opération $\tau'_{i,j} : L_i \leftrightarrow L_j$ revient à remplacer la matrice A par la matrice $P_{i,j} \cdot A$.

En résumé, les opérations élémentaires sur les *colonnes* reviennent à multiplier *à droite* par des matrices élémentaires, les opérations élémentaires sur les *lignes* reviennent à multiplier *à gauche* par des matrices élémentaires.

On peut ainsi transformer la matrice A par une suite d'opérations élémentaires, afin de la simplifier au maximum : on cherche à décrire à la fois la matrice la plus simple que l'on peut obtenir ainsi à partir de A, ainsi qu'un algorithme efficace pour réaliser ces opérations élémentaires successives en pratique. En termes plus savants, on dipose ainsi de plusieurs actions du sous-groupe de $\mathrm{GL}_n(K)$ (ou $\mathrm{GL}_p(K)$) engendré par les matrices élémentaires, et on cherche à décrire les orbites sous cette action, en donnant un représentant "le plus simple possible" de chaque orbite.

En notant $E_n(K)$ le sous-groupe de $\mathrm{M}_n(K)$ engendré par les matrices élémentaires, les trois actions de groupes que l'on sera amené à regarder sont donc les suivantes :

— l'action par multiplication à droite $E_p(K) \times M_{n,p}(K) \to \mathrm{M}_{n,p}(K)$, définie par $(Q, A) \mapsto A \cdot Q^{-1}$, qui décrit les opérations élémentaires sur les colonnes.

— l'action par multiplication à gauche $E_n(K) \times \mathrm{M}_{n,p}(K) \to \mathrm{M}_{n,p}(K)$, définie par $(P, A) \mapsto P \cdot A$, qui décrit les opérations élémentaires sur les lignes.

— l'action par multiplication à droite et à gauche (dite action par équivalence) $(E_n(K) \times E_p(K)) \times \mathrm{M}_{n,p}(K) \to \mathrm{M}_{n,p}(K)$, définie par $((P, Q), A) \mapsto P \cdot A \cdot Q^{-1}$, qui décrit les opérations élémentaires sur les lignes et les colonnes.

Décrivons maintenant les orbites pour chacune de ces actions, ainsi que l'algorithme permettant de "simplifier" une matrice donnée par une suite d'opérations élémentaires.

Définition 3.56. La matrice A est dite échelonnée réduite en colonnes (resp. en lignes) si elle vérifie :

- Si une colonne C_i (resp. ligne L_i) est nulle, alors toutes les colonnes (resp. lignes) suivantes C_j (resp. L_j), avec $j > i$, sont nulles.
- Dans une colonne (resp. ligne) non nulle, le premier terme non nul, appelé pivot, est égal à 1, et c'est le seul coefficient non nul de sa ligne (resp. colonne).
- Si $j > i$, le pivot de la colonne C_j (resp. ligne L_j) est situé strictement en dessous (resp. à droite) du pivot de la colonne C_i (resp. ligne L_i).

Théorème 3.57. *Soit $A \in \mathrm{M}_{n,p}(K)$.*

1. *Il existe une unique matrice E échelonnée réduite en colonnes dans l'orbite de A pour l'action de $E_p(K)$ par multiplication à droite, i.e. E est l'unique telle matrice qui peut être obtenue à partir de A par opérations élémentaires sur les colonnes. En outre, deux matrices sont dans la même orbite si et seulement si elles ont même image.*
2. *Il existe une unique matrice E' échelonnée réduite en lignes dans l'orbite de A pour l'action de $E_n(K)$ par multiplication à gauche, i.e. E' est l'unique telle matrice qui peut être obtenue à partir de A par opérations élémentaires sur les lignes. En outre, deux matrices sont dans la même orbite si et seulement si elles ont même noyau.*
3. *Il existe un unique entier $r \leq \min(n,p)$ tel que la matrice J_r soit dans l'orbite de A pour l'action de $E_n(K) \times E_p(K)$ par équivalence, i.e. J_r est l'unique matrice de ce type qui peut être obtenue à partir de A par opérations élémentaires sur les lignes et les colonnes.*

En outre, dans les trois cas, on dispose d'un algorithme permettant d'obtenir explicitement les opérations élémentaires nécessaires.

Remarque 3.58. On peut interpréter le troisième point du théorème comme une version effective de la proposition 3.32, où l'on ne s'autorise que des opérations élémentaires.

Démonstration: Les deux premiers points se déduisent l'un de l'autre par transposition. Il suffit donc de prouver l'un des deux, par exemple le premier.

- Existence : la preuve de l'existence dans ce théorème est essentiellement la description de l'algorithme de Gauss. Si la matrice A est nulle ou vide, il n'y a rien à faire (A est déjà échelonnée). Supposons donc A non nulle. On cherche la première ligne non nulle, disons L_i, de A. Sur la ligne L_i, on cherche le premier coefficient non nul, disons $a_{i,j}$. Puis on fait les opérations élémentaires suivantes :
  1. $C_j \leftarrow \frac{1}{a_{i,j}} C_j$.
  2. pour tout $k \neq j$, $C_k \leftarrow C_k - a_{i,k} \cdot C_j$.
  3. $C_1 \leftrightarrow C_j$.

  Après ces opérations, on arrive à une matrice A' de la forme

$$A' = \left(\begin{array}{c|c} 0_{i-1,1} & 0_{i-1,p-1} \\ \hline 1 & 0_{1,p-1} \\ \hline \star & A'_1 \end{array}\right),$$

où $A'_1 \in \mathrm{M}_{n-i,p-1}(K)$. On recommence l'étape précédente, en remplaçant A par la matrice $A_1 := \left(\begin{array}{c} 0_{i-1,p-1} \\ \hline 0_{1,p-1} \\ \hline A'_1 \end{array}\right)$.
On s'arrête quand on arrive à une matrice nulle ou vide. Il est évident que cet algorithme termine (la taille de la matrice diminue strictement à chaque étape), et on obtient à la fin une matrice A_∞ de la forme :

$$A_\infty = \left(\begin{array}{c|c|c|c|c} & 0 & & & \\ \hline 1 & & & & \\ \vdots & & & & \\ \vdots & 0 & & & \\ \hline \star & 1 & & & \\ \vdots & \vdots & & 0 & \\ \vdots & \vdots & & & \\ \hline \star & \star & 1 & & \\ \vdots & \vdots & \vdots & & \\ \vdots & \vdots & \vdots & 0 & \\ \hline \star & \star & \star & 1 & \\ \vdots & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right).$$

Enfin, la dernière étape consiste à remplacer les termes non nuls $\star$ sur les lignes des pivots par des 0. Pour cela, on fait les opérations suivantes : pour tout i, pour tout $j < i$, $C_j \leftarrow C_j - a_{p(i),j} \cdot C_i$, où $p(i)$ désigne la position (l'indice) du pivot

sur la colonne C_i. On arrive alors à la matrice échelonnée réduite souhaitée :

$$E_\infty = \left(\begin{array}{c|ccc|cc} & & 0 & & & \\ \hline 1 & & & & & \\ \vdots & & 0 & & & \\ \vdots & & & & & \\ \hline 0 & 1 & & & & \\ \vdots & \vdots & & 0 & & \\ \vdots & \vdots & & & & \\ \hline 0 & 0 & 1 & & & \\ \vdots & \vdots & \vdots & & 0 & \\ \vdots & \vdots & \vdots & & & \\ \hline 0 & 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{array}\right).$$

— Unicité : soient E, E' deux matrices échelonnées réduites en colonnes, et $Q \in \mathrm{GL}_p(K)$ tels que $E' = E \cdot Q$. Comme Q est inversible, on voit que $\mathrm{Im}\,(E') = \mathrm{Im}\,(E)$. Or il est facile de voir qu'une matrice échelonnée réduite est complétement déterminée par l'image de la base canonique, i.e. on en déduit que $E = E'$, d'où l'unicité souhaitée.

Pour démontrer le troisième point, il suffit d'appliquer successivement le premier point, puis le deuxième : partant en effet de la matrice E_∞ précédemment obtenue, on peut la réduire suivant les lignes, de la façon suivante : on peut d'abord permuter les lignes de manière à ramener les pivots sur la diagonale, et obtenir une matrice triangulaire inférieure avec d'abord des 1, puis éventuellement des 0, sur la diagonale. Et enfin, on applique l'algorithme du pivot de Gauss sur les lignes, de la façon suivante : pour i croissant, pour $j > i$ croissant, on fait $L_j \leftarrow L_j - a_{j,i} \cdot L_i$. Il est clair que l'on obtient alors une matrice J_r à la fin. □

Voici quelques conséquences immédiates du théorème 3.57 et de sa preuve :

Corollaire 3.59.
1. *Deux matrices sont équivalentes si et seulement si elles ont même rang.*
2. *Le rang d'une matrice est égale à celui de sa transposée, i.e. le rang d'une matrice coïncide avec le rang du système associé.*
3. $\mathrm{GL}_n(K) = E_n(K)$, *i.e. toute matrice inversible est produit de matrices élémentaires, ou encore : les matrices élémentaires engendrent* $\mathrm{GL}_n(K)$.

Corollaire 3.60. *Si* $K = \mathbf{R}$ *ou* $\mathbf{C}$, *alors* $\mathrm{GL}_n(K)$ *est dense dans* $\mathrm{M}_n(K)$.

Corollaire 3.61. *Par* 5 *points du plan passe une conique.*

On peut raffiner légérement le pivot de Gauss pour la décomposition LU et la décomposition de Bruhat :

Théorème 3.62. *Soit $A \in \mathrm{GL}_n(K)$. Il existe une unique permutation $\sigma \in \mathfrak{S}_n$, et deux uniques matrices L et U telle que $P_\sigma \cdot A = L \cdot U$, avec L triangulaire inférieure avec des 1 sur la diagonale, et U triangulaire supérieure.*

Application : résolution de système $A \cdot X = B$ pour plusieurs vecteurs B (et la même matrice A).

Théorème 3.63. *On note $T_n(K)^*$ l'ensemble des matrices triangulaires supérieures dans $\mathrm{GL}_n(K)$ et $U_n(K) \subset T_n(K)^*$ l'ensemble des matrices triangulaires supérieures avec des 1 sur la diagonale.*

Alors

$$\mathrm{GL}_n(K) = \bigsqcup_{\sigma \in \mathfrak{S}_n} T_n(K)^* \cdot P_\sigma \cdot U_n(K)\,.$$

Corollaire 3.64. *L'action de $\mathrm{GL}_n(K)$ sur les couples de drapeaux a exactement $n!$ orbites. Plus précisément, si $\mathcal{D}$ désigne le drapeau canonique de K^n, alors un système de représentants des orbites est donné par les couples $(\mathcal{D}, P_\sigma \cdot \mathcal{D})$, pour $\sigma \in \mathfrak{S}_n$.*

Étudions rapidement la complexité de l'algorithme de Gauss décrit dans la preuve. Pour obtenir la forme échelonnée réduite associée à une matrice $A \in \mathrm{M}_n(K)$, l'algorithme de Gauss procède à au plus $\frac{n(n+1)}{2}$ opérations élémentaires. Une opération élémentaire consistant en n opérations sur les coefficients de la matrice (somme ou produit d'éléments de K), on obtient donc une complexité de l'ordre de n^3 opérations dans K pour réduire une matrice sous une forme échelonnée réduite (ou sous sa forme J_r).

Mentionnons maintenant quelques premières applications pratiques de l'algorithme du pivot de Gauss. Soit $A \in \mathrm{M}_{n,p}(K)$. Notons au passage que tous les algorithmes proposés ci-dessous se ramènent in fine à un pivot de Gauss, ce qui assure que leur complexité est en $\mathcal{O}(n^3)$ opérations élémentaires dans K.

1. Calcul du rang d'une matrice :
   l'algorithme aboutit à une matrice échelonnée à r lignes (ou r colonnes), et le rang de la matrice A est exactement le nombre de pivots, c'est-à-dire le nombre de lignes (resp. colonnes) non nulles.
2. Détermination d'une base de l'image et du noyau d'une matrice :
   Il s'agit d'appliquer l'algorithme de Gauss (suivant les colonnes) à la matrice "augmentée" suivante $A' := \left(\begin{array}{c} A \\ \hline I_p \end{array}\right)$ dans $\mathrm{M}_{n+p,p}(K)$. On obtient à la fin une matrice de la forme $A'' := \left(\begin{array}{c} E \\ \hline B \end{array}\right)$, avec E échelonnée stricte en colonnes, donc de la forme $A'' := \left(\begin{array}{c|c} E_r & 0 \\ \hline B_r & N \end{array}\right)$, avec E_r formée de r colonnes non nulles. Alors les r vecteurs colonnes de E_r forment une base de $\mathrm{Im}\,(A)$ dans K^n, alors que les $p - r$ vecteurs colonnes de N forment une base de $\ker(A)$ dans K^p.
3. Détermination de systèmes minimaux d'équations décrivant le noyau ou l'image d'une matrice :
   La méthode est duale du point précédent. Il s'agit d'appliquer l'algorithme de Gauss (suivant les lignes) à la matrice "augmentée" suivante $A' := \left(\begin{array}{c|c} A & I_n \end{array}\right)$ dans $\mathrm{M}_{n,p+n}(K)$. On obtient à la fin une matrice de la forme $A'' := \left(\begin{array}{c|c} F & C \end{array}\right)$,

avec F échelonnée stricte en lignes, donc de la forme $A'' = \left(\begin{array}{c|c} F_r & C_r \\ \hline 0 & M \end{array}\right)$, avec F_r formée de r lignes non nulles. Alors les r vecteurs lignes de F_r forment une base du dual de $\ker(A)$, c'est-à-dire que ces vecteurs lignes donnent un système minimal d'équations décrivant $\ker(A)$ dans K^p (le système minimal, formé de r équations indépendantes, est $F_r \cdot X = 0$), alors que les $n - r$ vecteurs lignes de M forment une base du dual de $\mathrm{Im}\,(A)$ dans K^n, donnant le système minimal de $n - r$ équations indépendantes $M \cdot Y = 0$ dans K^n.

4. Résolution de systèmes linéaires :

   Pour résoudre le système (S) homogène donné par $A \cdot X = 0$, d'inconnue $X \in K^p$, il suffit de réduire A à une matrice échelonnée :
   — en lignes, si l'on souhaite obtenir un système minimal de $p - r$ équations linéaires indépendantes décrivant l'ensemble des solutions, système qui plus est triangulaire et donc immédiat à résoudre complétement.
   — en colonnes si l'on souhaite obtenir une base de l'espace vectoriel $\ker(A)$ des solutions.

   Dans le cas d'un système non homogène $A \cdot X = B$, l'échelonnement en lignes de la matrice augmentée $M := \left(\begin{array}{c|c} A & B \end{array}\right)$ fournit une matrice échelonnée en lignes de la forme $M' := \left(\begin{array}{c|c} A' & B' \\ \hline 0 & B'' \end{array}\right)$. Alors le système initial a des solutions si et seulement si $B'' = 0$ (conditions de compatibilité du système), ce qui équivaut à $B \in \mathrm{Im}\,(A)$.

   Si cette condition est vérifiée, le système initial $A \cdot X = B$ équivaut au système $A' \cdot X = B'$, qui est un sytème échelonné sans ligne nulle. En particulier, $A' \cdot X = B'$ est un système minimal d'équations affines décrivant l'ensemble des solutions du système initial. On peut également obtenir un repère affine du sous-espace des solutions en déterminant une solution particulière X_0 (une fois le systéme échelonné, c'est immédiat en prenant par exemple toutes les coordonnées hors des pivots de A' égales à 0), ainsi qu'une base de $\ker(A)$ (ou de $\ker(A')$).

5. Calcul de l'inverse d'une matrice :

   Si $A \in \mathrm{M}_n(K)$, l'idée est d'appliquer l'algorithme d'échelonnement (en lignes) à la matrice augmentée $\left(\begin{array}{c|c} A & I_n \end{array}\right)$. On aboutit à une matrice échelonnée réduite en lignes, de la forme $\left(\begin{array}{c|c} E & A' \end{array}\right)$. Alors la matrice A est inversible si et seulement si $E = I_n$, et dans ce cas $A' = A^{-1}$.

3.6 Déterminant

Un outil théorique très important pour l'étude des endomorphismes et de leur inversibilité est le déterminant. Dans cette partie, nous parlerons du déterminant d'un endomorphisme, d'une matrice, d'une famille de vecteurs, et nous allons donner plusieurs descriptions du déterminant : une description théorique via les applications multilinéaires alternées, une formule par récurrence sur la taille de la matrice (développement par rapport à une ligne ou à une colonne), une formule explicite à l'aide d'une somme indicées par les permutations dans $\mathfrak{S}_n$. La description à l'aide des puissances extérieures n'est pas au programme de l'agrégation.

3.6.1 Applications multilinéaires alternées

Définition 3.65. Soient $E_1, \ldots, E_n$ et F des K-espaces vectoriels. Une application $\varphi : E_1 \times E_n \to F$ est dite n-linéaire si pour tout $1 \leq i \leq n$, pour tout $(x_j)_{j\neq i} \in \prod_{j\neq i} E_j$, l'application $\varphi(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n) : E_i \to F$ est linéaire. Autrement dit, pour tout $1 \leq i \leq n$, pour tout $(x_j) \in \prod_{j=1}^n E_j$, $y_i \in E_i$, $\lambda \in K$, on a

$$\varphi(x_1, \ldots, x_{i-1}, \lambda \cdot x_i + y_i, x_{i+1}, \ldots, x_n) = \lambda \cdot \varphi(x_1, \ldots, x_i, \ldots, x_n) + \varphi(x_1, \ldots, y_i, \ldots, x_n)\,.$$

Exemples 3.66.
1. Le produit $K \times K \to K$ est une application 2-linéaire (on dit plutôt bilinéaire).
2. Le produit scalaire euclidien $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ est bilinéaire.
3. Le produit vectoriel $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$ est bilinéaire.
4. L'accouplement de dualité $E \times E^* \to K$ (où E est un K-espace vectoriel) est bilinéaire.

Définition 3.67. Une application n linéaire $E^n \to F$ est dite
- symétrique si pour tout $\sigma \in \mathfrak{S}_n$, tout $(x_1, \ldots, x_n) \in E^n$, $\varphi((x_{\sigma(i)})) = \varphi((x_i))$.
- antisymétrique si pour tout $\sigma \in \mathfrak{S}_n$, tout $(x_1, \ldots, x_n) \in E^n$, $\varphi((x_{\sigma(i)})) = \varepsilon(\sigma) \cdot \varphi((x_i))$.
- alternée si pour tout $(x_1, \ldots, x_n) \in E^n$, $\varphi((x_i)) = 0$ s'il existe $i \neq j$ tel que $x_i = x_j$.

Exemples 3.68.
1. Le produit $K \times K \to K$ est symétrique.
2. Le produit scalaire euclidien sur $\mathbf{R}^n$ est symétrique.
3. Le produit vectoriel sur $\mathbf{R}^3$ est alterné et antisymétrique.

Proposition 3.69. *Une application n-linéaire alternée est antisymétrique.*

Démonstration: Soit $\varphi : E^n \to F$ une application n-linéaire alternée. Il suffit de calculer, pour $1 \leq i \leq n-1$, et $(x_k) \in E^n$,

$$0 = \varphi(x_1, \ldots, x_{i-1}, x_i + x_{i+1}, x_{i+1} + x_i, x_{i+1}; \ldots, x_n) = \varphi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \ldots, x_n) + \varphi(x_1, \ldots, x_{i-1}$$

ce qui implique que $\varphi(x_1, \ldots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \ldots, x_n) = -\varphi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \ldots, x_n)$. Or $\mathfrak{S}_n$ est engendré par les transpositions $(i, i+1)$, donc on déduit de la formule précédente que pour tout $\sigma \in \mathfrak{S}_n$, $\varphi((x_{\sigma(i)})) = \varepsilon(\sigma) \cdot \varphi((x_i))$, donc φ est antisymétrique. □

Proposition 3.70. *Si K est de caractéristique différente de 2, une application n-linéaire antisymétrique est alternée.*

Démonstration: Soit $x = (x_k) \in E^n$ tel que $x_i = x_j$, avec $i \neq j$. Alors en appliquant la transposition (i, j), on a $\varphi(x) = -\varphi(x)$, donc $2 \cdot \varphi(x) = 0$. Comme K est de caractéristique différente de 2, cela implique que $\varphi(x) = 0$, donc φ est alternée. □

Remarque 3.71. En caractéristique 2, une forme n-linéaire est symétrique si et seulement si elle est antisymétrique, et il existe des formes n-linéaires antisymétriques non alternées. Par exemple, le produit $\mathbf{F}_2 \times \mathbf{F}_2 \to \mathbf{F}_2$ dans le corps à deux éléments $\mathbf{F}_2$ est bien symétrique, donc antisymétrique, mais il n'est pas alterné : $1 \cdot 1 \neq 0$.

Soit E un K-espace vectoriel de dimension n. On note $\mathcal{L}_k(E)$ l'espace vectoriel des formes k-linéaires sur E, $\mathcal{S}_k(E)$ le sous-espace vectoriel des formes k-linéaires symétriques sur E et $\mathcal{A}_k(E)$ le sous-espace vectoriel des formes k-linéaires alternées sur E.

Définition 3.72. Soit E un K-espace vectoriel et $f_1, \dots, f_k \in E^*$ des formes linéaires. On définit les applications suivantes :

— $f_1 \times \cdots \times f_k : E^k \to K$ par $(x_1, \dots, x_k) \mapsto f_1(x_1) \cdot \cdots \cdot f_k(x_k)$.
— $f_1 \cdot \cdots \cdot f_k : E^k \to K$ par $f_1 \wedge \cdots \wedge f_k := \sum_{\sigma \in \mathfrak{S}_k} f_{\sigma(1)} \times \cdots \times f_{\sigma(k)}$.
— $f_1 \wedge \cdots \wedge f_k : E^k \to K$ par $f_1 \wedge \cdots \wedge f_k := \sum_{\sigma \in \mathfrak{S}_k} \varepsilon(\sigma) f_{\sigma(1)} \times \cdots \times f_{\sigma(k)}$.

Lemme 3.73. *Avec les notations précédentes, $f_1 \times \cdots \times f_k \in \mathcal{L}_k(E)$, $f_1 \cdot \cdots \cdot f_k \in \mathcal{S}_k(E)$ et $f_1 \wedge \cdots \wedge f_k \in \mathcal{A}_k(E)$.*

Démonstration: C'est une simple vérification. □

On dispose donc d'un moyen de fabriquer des formes k-linéaires (symétriques ou alternées) à l'aide de formes linéaires sur E. L'intérêt principal est le suivant :

Proposition 3.74. *Soit E un K-espace vectoriel de dimension n, muni d'une base $(e_1, \dots, e_n)$. Alors*

1. *la famille $(e_{i_1}^* \times \cdots \times e_{i_k}^*)_{1 \le i_1, \dots, i_k \le n}$ est une base de $\mathcal{L}_k(E)$.*
2. *(On suppose $\mathrm{car}(K) = 0$ ou $\mathrm{car}(K) > k$) la famille $(e_{i_1}^* \cdot \cdots \cdot e_{i_k}^*)_{1 \le i_1 \le \cdots \le i_k \le n}$ est une base de $\mathcal{S}_k(E)$.*
3. *la famille $(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*)_{1 \le i_1 < \cdots < i_k \le n}$ est une base de $\mathcal{A}_k(E)$.*

Démonstration:

1. Soit $f \in \mathcal{L}_k(E)$. Pour tout $(x_1, \dots, x_k)$ de E^k, on écrit, pour $1 \le i \le k$, $x_i = \sum_{j=1}^n x_{i,j} \cdot e_j$.
   Alors
   $$f(x_1, \dots, x_k) = f(\sum_{j=1}^n x_{1,j} \cdot e_j, \dots, \sum_{j=1}^n x_{k,j} \cdot e_i)$$
   et en utilisant la multilinéarité, on trouve
   $$f(x_1, \dots, x_k) = \sum_{1 \le i_1, \dots, i_k \le n} f(e_{i_1}, \dots, e_{i_k}) \cdot x_{1,i_1} \dots x_{k,i_k} \,,$$
   ce qui se réécrit :
   $$f = \sum_{1 \le i_1, \dots, i_k \le n} f(e_{i_1}, \dots, e_{i_k}) \cdot e_{i_1}^* \times \cdots \times e_{i_k}^* \,.$$
   Donc la famille est génératrice.
   La liberté a été démontrée en cours.
2. Les deux autres cas sont laissés à la sagacité de la lectrice.

□

Corollaire 3.75. *On suppose que* $\mathrm{car}(K) = 0$ *ou* $\mathrm{car}(K) > k$ *pour la deuxième égalité. Alors*

$$\dim\left(\mathcal{L}_k(E)\right) = n^k\,,\ \dim\left(\mathcal{S}_k(E)\right) = \binom{n+k-1}{k}\,,\ \dim\left(\mathcal{A}_k(E)\right) = \binom{n}{k}\,.$$

Remarque 3.76. Comme mentionné dans la preuve, pour tout $\varphi \in \mathcal{L}_k(E)$, pour tout $(x_1, \ldots, x_k) \in E^k$, en écrivant $x_j = \sum_i a_{i,j} \cdot e_i$, on a

$$\varphi(x_1, \ldots, x_k) = \sum_{1 \leq i_1, \ldots, i_k \leq n} a_{i_1,1} \cdot \cdots \cdot a_{i_k,k} \varphi(e_{i_1}, \ldots, e_{i_k})\,.$$

En particulier, φ est complétement déterminée par les valeurs $\varphi(e_{i_1}, \ldots, e_{i_k})$.

Si φ est symétrique, alors on peut regrouper certains termes dans la somme précédente, et on obtient une écriture pour $\varphi(x_1, \ldots, x_k)$ comme une somme indicée par les $1 \leq i_1 \leq \cdots \leq i_k \leq n$, et donc φ est déterminée par les valeurs $\varphi(e_{i_1}, \ldots, e_{i_k})$ avec $1 \leq i_1 \leq \cdots \leq i_k \leq n$, et ces valeurs peuvent être choisies arbitrairement. On retrouve donc que $\dim\left(\mathcal{S}_k(E)\right) = \binom{n+k-1}{k}$ sans restriction sur la caractéristique. On peut même donner une base de $\mathcal{S}_k(E)$: pour tout $1 \leq i_1 \leq \cdots \leq i_k \leq n$, on définit $r_1, \ldots r_s \geq 1$ tels que $i_1 = \cdots = i_{r_1} < i_{r_1+1} = \cdots = i_{r_1+r_2} < \cdots < i_{r_1+\cdots+r_{s-1}+1} = \ldots i_{r_1+\cdots+r_n}$, avec $r_1 + \cdots + r_s = k$, et alors on pose

$$e^*_{i_1} \bullet \cdots \bullet e^*_{i_k} := \sum_{\sigma \in \mathfrak{S}_k / \prod_{i=1}^s \mathfrak{S}_{r_i}} e^*_{i_{\sigma(1)}} \times \cdots \times e_{i_{\sigma(k)}}$$

Alors les $(e^*_{i_1} \bullet \cdots \bullet e^*_{i_k})_{1 \leq i_1 \leq \cdots \leq i_k \leq n}$ forment une base de $\mathcal{S}_k(E)$ en tout caractéristique.

Le cas fondamental pour l'étude du déterminant est le cas $k = n$. Les calculs précédents assurent que $\dim\left(\mathcal{A}_n(K)\right) = 1$. En effet, pour tout forme φ n-linéaire alternée dans E, pour tout $x \in E^n$, on a (avec les notations précédentes) :

$$\varphi(x_1, \ldots, x_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdot \cdots \cdot a_{\sigma(n),n} \right) \varphi(e_1, \ldots, e_n)\,.$$

Cette formule assure immédiatement que deux formes n-linéaires alternées sur E sont proportionnelles, et il est facile de vérifier que la formule

$$\varphi : (x_1, \ldots, x_n) \mapsto \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdot \cdots \cdot a_{\sigma(n),n}$$

définit une forme n-linéaire alternée non nulle. On retrouve donc que $\dim\left(\mathcal{A}_n(K)\right) = 1$, et qu'une telle forme est déterminée par sa valeur sur une base de E.

Cela justifie la définition suivante :

Définition 3.77. Soit E un K-espace vectoriel de dimension n, muni d'une base $\mathcal{B}$.

Le déterminant dans la base $\mathcal{B}$, noté $\det_{\mathcal{B}}$, est l'unique application n-linéaire alternée sur E qui prend la valeur 1 sur $\mathcal{B}$.

La formule suivante est une conséquence immédiate des remarques qui précèdent la définition :

Proposition 3.78. *Soit $(x_1, \dots, x_n) \in E^n$. On écrit, pour tout j, $x_j = \sum_i a_{i,j} \cdot e_i$.*
Alors

$$\det_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n} \cdot$$

Remarque 3.79. Cette formule est en fait très peu utile pour calculer un déterminant en pratique, sauf si la dimension n est petite ($n = 2$ ou éventuellement $n = 3$). En effet, pour n plus grand, on a une somme de $n!$ termes, ce qui est beaucoup trop.

Observons ce qui se passe si l'on change de base :

Proposition 3.80. *Soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de E.*
Alors $\det_{\mathcal{C}} = \det_{\mathcal{C}}(\mathcal{B}) \cdot \det_{\mathcal{B}}$. *En particulier,* $\det_{\mathcal{B}}(\mathcal{C}) \cdot \det_{\mathcal{C}}(\mathcal{B}) = 1$.

Démonstration: c'est immédiat. □

Proposition 3.81. *Soit $\mathcal{B}$ une base de E et $x \in E^n$.*
La famille x est libre si et seulement si $\det_{\mathcal{B}}(x) \neq 0$.

Démonstration: C'est évident. □

On est désormais à même de définir le déterminant d'un endomorphisme :

Définition 3.82. Soit E un K-espace vectoriel de dimension n, muni d'une base $\mathcal{B}$.
Le déterminant de u, noté $\det(u)$, est $\det_{\mathcal{B}}(u(\mathcal{B}))$, et il est indépendant de la base $\mathcal{B}$ choisie.

Démonstration: Si $\mathcal{C}$ est une base de E, on a
- soit u n'est pas inversible, donc $u(\mathcal{B})$ et $u(\mathcal{C})$ ne sont pas des bases, donc $\det_{\mathcal{B}}(u(\mathcal{B})) = 0 = \det_{\mathcal{C}}(u(\mathcal{C}))$.
- soit u est inversible, donc $u(\mathcal{B})$ et $u(\mathcal{C})$ sont des bases de E, et $\det_{\mathcal{C}}(u(\mathcal{C})) = \det_{\mathcal{C}}(\mathcal{B}) \cdot \det_{\mathcal{B}}(u(\mathcal{B})) \cdot \det_{u(\mathcal{B})}(u(\mathcal{C}))$.
  Or $\det_{u(\mathcal{B})}(u(\mathcal{C})) = \det_{\mathcal{B}}(\mathcal{C})$, car les applications n-linéaires alternées $\det_{u(\mathcal{B})} \circ (u, \dots, u)$ et $\det_{\mathcal{B}}$ prennent la valeur 1 en $\mathcal{B}$ donc sont égales. Donc finalement $\det_{\mathcal{C}}(u(\mathcal{C})) = \det_{\mathcal{B}}(u(\mathcal{B}))$.

□

Les propriétés suivantes sont immédiates :

Proposition 3.83. *Soient $u, v \in \mathcal{L}(E)$.*

1. $\det(\mathrm{id}_E) = 1$.
2. $\det(u \circ v) = \det(u) \cdot \det(v)$.
3. *u est inversible si et seulement si* $\det(u) \neq 0$.

On peut enfin d'intéresser au déterminant d'une matrice :

Définition 3.84. Soit $A \in \mathrm{M}_n(K)$.
Le déterminant de A, noté $\det(A)$, est le déterminant de la famille des vecteurs colonnes de A dans la base canonique de K^n. De façon équivalente, c'est le déterminant de l'endomorphisme $\varphi_A : K^n \to K^n$ associé à A par la formule suivante $\varphi_A(X) = A \cdot X$.

Corollaire 3.85. *Soient* $A, B \in \mathrm{M}_n(K)$.

1. $\det(I_n) = 1$.
2. $\det(A \cdot B) = \det(A) \cdot \det(B)$.
3. $A \in \mathrm{GL}_n(K)$ *si et seulement si* $\det(A) \neq 0$.
4. *Le rang de* $C \in \mathrm{M}_{n,p}(K)$ *est le plus grand entier* r *tel qu'il existe une sous-matrice carrée* Δ *de taille* r *extraite de* C *telle que* $\det(\Delta) \neq 0$.

Exemples 3.86. Si $K = \mathbf{R}$ ou $\mathbf{C}$,

1. la fonction $\mathrm{rg} : \mathrm{M}_n(K) \to \mathbf{R}$ est semi-continue inférieurement, i.e. pour tout $A \in \mathrm{M}_{n,p}(K)$, il existe un voisinage U de A dans $\mathrm{M}_{n,p}(K)$ tel que pour tout $M \in U$, $\mathrm{rg}(M) \geq \mathrm{rg}(A)$. Autrement dit, localement, le rang des matrices ne peut qu'augmenter.
2. L'adhérence de l'ensemble des matrices de rang r dans $\mathrm{M}_{n,p}(K)$ est exactement l'ensemble des matrices de rang $\leq r$.

Une application du déterminant :

Définition 3.87. Soit E un $\mathbf{R}$-espace vectoriel de dimension finie. On dit que deux bases $\mathcal{B}$ et $\mathcal{C}$ de E sont équivalentes si $\det_{\mathcal{B}}(\mathcal{C}) > 0$, ou de façon équivalente, l'automorphisme u de E défini par $u(\mathcal{B}) = \mathcal{C}$ est de déterminant 1.

Alors cette relation d'équivalence partitionne l'ensemble des bases de E en deux classes d'équivalence, qui sont des orbites sous l'action (libre) de $\mathrm{GL}(E)^+$ sur l'ensemble des bases. Une classe d'équivalence est appelée une orientation de E.

Une interprétation importante du déterminant : sur $\mathbf{R}$, le déterminant est un volume.

Proposition 3.88. *Soit* E *un espace vectoriel euclidien de dimension* n*, muni d'une base orthonormée* $\mathcal{B}$*. Alors pour tout famille* $\mathcal{C}$ *de* n *vecteurs de* E, $|\det_{\mathcal{B}}(\mathcal{C})|$ *est le volume du parallélotope de* E *engendré par la famille* $\mathcal{C}$.

Démonstration: On oriente l'espace E par la base $\mathcal{B}$. On définit $\varphi : E^n \to \mathbf{R}$ de la façon suivante : $\varphi(\mathcal{C})$ est le volume du parallélotope engendré par $\mathcal{C}$ si $\mathcal{C}$ est une base directe ; $\varphi(\mathcal{C})$ est l'opposé du volume du parallélotope engendré par $\mathcal{C}$ sinon.

Alors on vérifie que φ est une application n-linéaire alternée sur E (on utilise notamment l'additivité du volume), et que $\varphi(\mathcal{B}) = 1$ car $\mathcal{B}$ est orthonormée. D'où le résultat. □

Exemple 3.89. Soit K un corps infini, et L une extension de K. Soient $A, B \in \mathrm{M}_n(K)$.

Alors A et B sont semblables dans $\mathrm{M}_n(L)$ si et seulement si elles sont semblables dans $\mathrm{M}_n(K)$.

Proposition 3.90. *Notons* $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathrm{M}_n(K)$.
Alors $\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdot \cdots \cdot a_{\sigma(n),n}$.

Corollaire 3.91. $\det({}^tA) = \det(A)$.

Exemple 3.92. Le déterminant d'une matrice de permutation $P_\sigma \in \mathrm{M}_n(K)$ est égal à la signature de σ (ou plus exactement à son image dans le corps K).

Nous allons maintenant présenter deux façons de calculer le déterminant d'une matrice, plus efficaces que la formule précédente dont la complexité est factorielle. Il s'agit de la formule de développement par rapport à une ligne ou à une colonne, et du lien entre déterminant et algorithme du pivot de Gauss.

Définition 3.93. Soit $A \in \mathrm{M}_n(K)$. Pour tout $1 \leq i, j \leq n$, on note $A_{i,j}$ la matrice dans $\mathrm{M}_{n-1}(K)$ obtenue à partir de A en enlevant sa i-ième ligne et sa j-ième colonne.

Ces matrices sont appelées les mineures de A.

Proposition 3.94. *Soit $A = (a_{i,j}) \in \mathrm{M}_n(K)$.*

Alors pour tout $1 \leq i \leq n$ (resp. pour tout $1 \leq j \leq n$), on a

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{i,j} \cdot \det(A_{i,j})$$

(resp.

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{i,j} \cdot \det(A_{i,j})\,.)$$

Démonstration: On vérifie que la première formule est n-linéaire et alternée (si $j < k$ et $C_j = C_k$ dans A, alors $A_{i,k}$ se déduit de $A_{i,j}$ en faisant une permutation circulaire des colonnes d'indices compris entre j et $k-1$, de signature $(-1)^{k-j-1}$ ...) en les colonnes de la matrice A. En outre, cette première formule prend clairement la valeur 1 en la matrice I_n. Donc par définition du déterminant, la première formule est vérifiée.

On en déduit la seconde en transposant. □

Ces formules de développement permettent un calcul récursif du déterminant, et elles peuvent s'avérer efficaces si la matrice contient beaucoup de 0. On peut également la combiner avec les opérations élémentaires comme expliqué plus bas.

Définition 3.95. Soit $A \in \mathrm{M}_n(K)$. La comatrice (ou matrice des cofacteurs) de A, notée $\mathrm{Com}(A)$, est la matrice de $\mathrm{M}_n(K)$ dont le coefficient d'indice (i, j) est défini par

$$(-1)^{i+j} \cdot \det(A_{i,j})\,.$$

Ces coefficients sont appelés les cofacteurs de A.

Corollaire 3.96. *Soit $A \in \mathrm{M}_n(K)$.*

Alors

$$^t\mathrm{Com}(A) \cdot A = \det(A) \cdot I_n\,.$$

Démonstration: Soit $1 \leq i, j \leq n$. Calculons le coefficient d'indice (i, j) de $^t\mathrm{Com}(A) \cdot A$: celui-ci vaut par définition $\sum_{k=1}^{n} (-1)^{i+k} \cdot a_{k,j} \cdot \det(A_{k,i})$.

- Si $i \neq j$, ce coefficient est le développement par rapport à la i-ième colonne de la matrice obtenue à partir de A en remplaçant sa colonne C_i par la colonne C_j. Cette matrice ayant deux colonnes identiques, son déterminant est nul, donc le coefficient en question est 0.
- Si $i = j$, ce coefficient est le développement de A par rapport à la i-ème colonne, donc il vaut $\det(A)$ par la proposition 3.94.

On a donc bien montré que le coefficient d'indice (i,j) de ${}^t\mathrm{Com}(A)\cdot A$ vaut $\det(A)\cdot\delta_{i,j}$.
□

Faisons maintenant le lien entre déterminant et pivot de Gauss, afin de calculer efficacement un déterminant.

Soit $A \in \mathrm{M}_n(K)$. On applique l'algorithme du pivot de Gauss (selon les colonnes par exemple) sans s'autoriser de multiplier les colonnes par un scalaire, afin d'obtenir, via des opérations élémentaires sur A (ajout d'un mutliple d'une colonne à une autre colonne et permutation de colonnes), une matrice A' échelonnée en colonnes (les pivots peuvent être différents de 1), avec r colonnes non nulles. En particulier, la matrice A' est triangulaire.

— Si $r < n$, alors $\det(A) = 0$.
— Sinon, le déterminant de A est le produit des coefficients diagonaux de A', multiplié par un signe donné par la parité du nombre de permutations effectuées lors des opérations élémentaires.

Cet algorithme a une complexité de l'ordre de $\mathcal{O}(n^3)$ opérations élementaires dans K.

On peut également combiner le pivot de Gauss avec la formule de développement selon les lignes, en faisant en sorte d'abord (via des opérations $C_i \leftarrow C_i + \lambda C_j$) de mettre des zéros partout sur la première ligne de la matrice, sauf le pivot, puis en développant ensuite selon la première ligne, afin de se ramener à un déterminant de taille inférieure.

Définition 3.97. Le groupe spécial linéaire, noté $\mathrm{SL}_n(K)$, est le noyau du morphisme de groupes $\det : \mathrm{GL}_n(K) \to K^\times$.

Lemme 3.98. *Le sous-groupe $\mathrm{SL}_n(K)$ est distingué dans $\mathrm{GL}_n(K)$, et le déterminant induit un isomorphisme $\overline{\det} : \mathrm{GL}_n(K)/\mathrm{SL}_n(K) \xrightarrow{\sim} K^\times$.*

Abordons maintenant quelques applications de la notion de déterminant.

Les résultats suivants découlent essentiellement du pivot de Gauss :

Corollaire 3.99.
— *Le groupe $\mathrm{SL}_n(K)$ est engendré par les matrices de transvections.*
— *Si $n \geq 3$ ou $|K| \geq 3$, alors $D(\mathrm{GL}_n(K)) = \mathrm{SL}_n(K)$.*
— *Si $n \geq 3$ ou $|K| \geq 4$, alors $D(\mathrm{SL}_n(K)) = \mathrm{SL}_n(K)$.*

Démonstration: Pour le premier point, on a besoin d'une variante du pivot de Gauss, pour montrer que toute matrice de $\mathrm{GL}_n(K)$ est produit de matrices de transvection et d'*une* matrice de dilatation. Cela se fait par récurrence sur n...

Pour les deux autres points, on montre que toute transvection est un commutateur, via les formules suivantes :
— Si $n \geq 3$, alors $T_{i,j}(\lambda) = [T_{i,k}(\lambda); T_{k,j}(1)]$ pour tout $k \notin \{i; j\}$.
— Si $n = 2$ et $|K| \geq 4$, alors il existe $\alpha \in K\backslash\{-1; 0; 1\}$, et $T_{1,2}(\lambda) = [\mathrm{diag}(\alpha, \alpha^{-1}); T_{1,2}(\frac{\lambda}{\alpha^2-1})]$.
— Le cas $n = 2$ et $|K| = 3$ est laissé à la lectrice curieuse.

□

Remarque 3.100. On peut étudier à la main les cas $n = 2$ et ($K = \mathbf{F}_2$ ou $\mathbf{F}_3$), en utilisant par exemple les isomorphismes exceptionnels du théorème 3.106.

Une première application du fait que $\mathrm{SL}_n(K)$ est engendré par les matrices de transvections :

Proposition 3.101. *Pour tout $n \geq 2$, et pour tout nombre premier p, la réduction modulo p induit un morphisme de groupes surjectif*

$$\mathrm{SL}_n(\mathbf{Z}) \to \mathrm{SL}_n(\mathbf{Z}/p\mathbf{Z})\,.$$

Démonstration: Il suffit de relever les matrices de transvection $T_{i,j}(\overline{k})$ dans $\mathrm{SL}_n(\mathbf{Z}/p\mathbf{Z})$ en des matrices de transvection (de déterminant 1) $T_{i,j}(k)$ dans $\mathrm{SL}_n(\mathbf{Z})$, et d'utiliser le premier point du théorème. □

En utilisant le lemme chinois, et une version du pivot de Gauss pour les matrices à coefficients dans $\mathbf{Z}/p^\alpha\mathbf{Z}$, on peut généraliser le résultat précédent :

Proposition 3.102 (Serre)**.** *Pour tout $n \geq 2$, et pour tout entier $N \geq 2$, la réduction modulo N induit un morphisme de groupes surjectif*

$$\mathrm{SL}_n(\mathbf{Z}) \to \mathrm{SL}_n(\mathbf{Z}/N\mathbf{Z})\,.$$

En outre, si $N \geq 3$, son noyau est sans torsion.

Une conséquence importante de ce résultat concerne les sous-groupes finis de $\mathrm{SL}_n(\mathbf{Z})$: pour tout nombre premier p impair, le cardinal d'un sous-groupe fini de $\mathrm{SL}_n(\mathbf{Z})$ divise $\frac{(p^n-1)\ldots(p^n-p^{n-1})}{p-1}$. On prend souvent $p = 3$ dans les applications.

Donnons maintenant une caractérisation géométrique des matrices de transvections, afin de l'utiliser dans le théorème qui suit :

Proposition 3.103. *Soit $u \in \mathcal{L}(E)$, $u \neq \mathrm{id}_E$. Les assertions suivantes sont équivalentes :*

1. $\det(u) = 1$ *et il existe un hyperplan H de E tel que $u_{|H} = \mathrm{id}_H$.*
2. *Il existe $\varphi \in E^* \setminus \{0\}$ et $v \in \ker(\varphi) \setminus \{0\}$ tels que pour tout $x \in E$, $u(x) = x + \varphi(x) \cdot v$.*
3. *Il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit une matrice de transvection.*

On déduit alors (avec du travail) du théorème précédent les conséquences suivantes :

Théorème 3.104. *Si $n \geq 3$ ou $|K| \geq 4$, alors $\mathrm{PSL}_n(K)$ est un groupe simple*

Démonstration: On se limite à la preuve dans le cas $n \geq 3$. Le cas $n = 2$ est pénible, mais pas difficile.

On montre d'abord que les transvections sont conjuguées dans $\mathrm{SL}_n(K)$. Si $n \geq 2$, on voit facilement que $T_{i,j}(\lambda)$ est conjuguée à $T_{1,2}(1)$ dans $\mathrm{GL}_n(K)$, puis si $n \geq 3$ dans $\mathrm{SL}_n(K)$.

Donc pour $n \geq 3$, les transvections engendrent $\mathrm{SL}_n(K)$ et sont toutes conjuguées dans $\mathrm{SL}_n(K)$.

Soit alors $\overline{N} \neq \{\mathrm{id}\}$ un sous-groupe distingué non trivial de $\mathrm{PSL}_n(K)$. Son image réciproque N dans $\mathrm{SL}_n(K)$ est un sous-groupe distingué contenant strictement les matrices scalaires de déterminant 1.

Il suffit alors de montrer que N contient une transvection. Il existe $A \in N$ non scalaire. En particulier, A n'est pas une homothétie, donc il existe $x \in K^n$ tel que $A \cdot x$ n'est pas proportionnel à x.

On note T une transvection de droite $K \cdot x$.

Alors $A' := [A; T] \in N$ et il existe H hyperplan tel que $\mathrm{Im}\,(A' - I_n) \subset H$. Donc en particulier H est stable par A'.

Deux cas :

1. soit il existe une transvection T' d'hyperplan H ne commutant pas avec A', auquel cas $A'' := [A', T']$ est une transvection dans N.
2. soit A' commute avec toutes les transvections d'hyperplan H, et on en déduit que la restriction de A' à H est l'identité, donc A' est une transvection dans N.

□

Remarque 3.105. On a vu dans la preuve que pour $n \geq 3$, les matrices de transvections sont conjuguées dans $\mathrm{SL}_n(K)$.

En revanche, on montre que pour $n = 2$, les matrices $T_{1,2}(a)$ et $T_{1,2}(b)$ sont conjuguées dans $\mathrm{SL}_2(K)$ si et seulement si $\frac{a}{b}$ est un carré dans $K^\times$.

On pourra s'intéresser aux cas exclus par les hypothèses du théorème précédent, et étudier les isomorphismes exceptionnels entre groupes simples :

Théorème 3.106. *L'action de* $\mathrm{PGL}_n(K)$ *sur* $\mathbf{P}^{n-1}(K)$ *induit des isomorphismes "exceptionnels" :*

1. $\mathrm{GL}_2(\mathbf{F}_2) = \mathrm{SL}_2(\mathbf{F}_2) = \mathrm{PSL}_2(\mathbf{F}_2) = \mathrm{PGL}_2(\mathbf{F}_2) \xrightarrow{\sim} \mathfrak{S}_3$.
2. $\mathrm{PGL}_2(\mathbf{F}_3) \xrightarrow{\sim} \mathfrak{S}_4$ *et* $\mathrm{PSL}_2(\mathbf{F}_3) \xrightarrow{\sim} \mathfrak{A}_4$.
3. $\mathrm{PSL}_2(\mathbf{F}_4) = \mathrm{PGL}_2(\mathbf{F}_4) \xrightarrow{\sim} \mathfrak{A}_5$.
4. $\mathrm{PGL}_2(\mathbf{F}_5) \xrightarrow{\sim} \mathfrak{S}_5$ *et* $\mathrm{PSL}_2(\mathbf{F}_5) \xrightarrow{\sim} \mathfrak{A}_5$.

Remarque 3.107.
- Plus délicat, on peut montrer en utilisant les théorèmes de Sylow, que l'on a un isomorphisme $\mathrm{PSL}_2(\mathbf{F}_9) \xrightarrow{\sim} \mathfrak{A}_6$.
- Dans un registre légèrement différent, on peut montrer d'autres isomorphismes exceptionnels : $\mathrm{PSL}_2(\mathbf{F}_5) \xrightarrow{\sim} \mathrm{PSL}_2(\mathbf{F}_4)$, et en utilisant les théorèmes de Sylow, $\mathrm{PSL}_2(\mathbf{F}_7) \xrightarrow{\sim} \mathrm{PSL}_3(\mathbf{F}_2)$ - l'unique groupe simple d'ordre 168. On peut également montrer que ce groupe simple est le "deuxième" groupe fini simple non abélien après $\mathfrak{A}_5$.
- En revanche, les groupes simples $\mathrm{PSL}_3(\mathbf{F}_4)$ et $\mathrm{PSL}_4(\mathbf{F}_2)$ ont même cardinal 20160, mais ne sont pas isomorphes : on peut par exemple comparer les classes de conjugaison d'éléments d'ordre 2 dans les deux groupes.

Un application importante de l'isomorphisme $\mathrm{PGL}_2(\mathbf{F}_5) \xrightarrow{\sim} \mathfrak{S}_5$: la construction d'un automorphisme non intérieur de $\mathfrak{S}_6$.

Énonçons enfin un développement désormais classique :

Théorème 3.108 (Frobenius-Zolotarev)**.** *Soit* $K = \mathbf{F}_q$ *un corps fini à* q *éléments. Si* $a \in \mathbf{F}_q^\times$*, on note* $\left(\frac{a}{q}\right) = 1$ *si* a *est un carré dans* K *et* -1 *sinon. On rappelle l'injection naturelle* $\iota : \mathrm{GL}_n(\mathbf{F}_q) \to \mathfrak{S}((\mathbf{F}_q)^n)$*. On suppose* $n \neq 2$ *ou* $q \neq 2$*.*

Alors pour tout $A \in \mathrm{GL}_n(\mathbf{F}_q)$*,* $\varepsilon(\iota(A)) = \left(\frac{\det(A)}{q}\right)$.

Une application : calcul du symbole de Legendre $\left(\frac{a}{p}\right)$, pour p premier impair et $a \in \mathbf{F}_p^\times$, comme la signature de la permutation de $\mathbf{F}_p$ donnée par $x \mapsto a \cdot x$. Par exemple, on peut retrouver $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

4 Réduction des endomorphismes

L'objectif principal de cette partie est le suivant : étant donné un endomorphisme $u \in \mathcal{L}(E)$, trouver une base $\mathcal{B}$ de E de sorte que la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ soit la plus simple possible, afin d'étudier u.

4.1 Sous-espaces stables, polynômes d'endomorphismes

Soit E un K-espace vectoriel de dimension n et $u \in mathcalL(E)$.

Définition 4.1. Un sous-espace vectoriel $F \subset E$ est stable par u (ou u-stable) si $u(F) \subset F$.

Si F est stable par u, alors la restriction de u à F définit un endomorphisme $u_{|F} \in \mathcal{L}(F)$ de F.

Traduction matricielle :

Proposition 4.2. *Soit $F \subset E$. Alors F est stable par u si et seulement s'il existe une base $\mathcal{C}$ de F telle que pour toute base $\mathcal{B} = \mathcal{C} \cup \mathcal{C}'$ de E complétant $\mathcal{C}$, la matrice de u dans la base $\mathcal{B}$ est de la forme :*

$$\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$$

avec $A = \mathrm{Mat}_{\mathcal{C}}(u_{|F})$.

Un moyen particulièrement utile de fabriquer des sous-espaces stables en vue de la réduction des endomorphismes est la notion de polynôme d'endomorphisme.

Définition 4.3. Soit $u \in \mathcal{L}(E)$ et $P = \sum_{i=0}^{n} a_i \cdot X^i \in K[X]$.

On définit l'endomorphisme $P(u) := \sum_{i=0}^{n} a_i \cdot u^i$, où u^i désigne la composée de u avec lui-même i fois.

On note également $K[u] \subset \mathcal{L}(E)$ la sous-algèbre des polynômes en u, i.e. $K[u] := \{P(u) : P \in K[X]\}$.

Voici un exemple important de sous-espace stable :

Proposition 4.4. *Soit $v \in \mathcal{L}(E)$ tel que $u \circ v = v \circ u$.*
Alors $\ker(v)$ est stable par u. En particulier, pour tout $P \in K[X]$, $\ker(P(u))$ est stable par u.

Les polynômes les plus utiles pour l'étude d'un endomorphisme u sont les polynômes annulateurs de u :

Définition 4.5.
- Un polynôme $P \in K[X]$ est dit annulateur de u si $P(u) = 0$.
- Un polynôme minimal de u est un polynôme annulateur non nul de degré minimal.

Le résultat suivant est immédiat :

Proposition 4.6. *Les polynômes annulateurs non nuls existent, et donc le polynôme minimal est bien défini.*

Le polynôme minimal de u est unique à un scalaire non nul près, et l'ensemble des polynômes annulateurs de u est un idéal de $K[X]$ engendré par un polynôme minimal.

Démonstration: L'espace vectoriel $\mathcal{L}(E)$ est de dimension finie n^2, donc la famille $(\mathrm{id}, u, u^2, \ldots, u^{n^2})$ est liée. Cela assure l'existence d'un polynôme non nul, de degré au plus n^2, qui annule u.

Le morphisme de K-algèbre $K[X] \to \mathcal{L}(E)$ défini par $P \mapsto P(u)$ est bien défini, et son noyau I, qui est un idéal, est exactement l'ensemble des polynômes annulateurs de u. Comme $K[X]$ est euclidien, et ses inversibles sont les scalaires non nuls, on en déduit que les générateurs de l'idéal I sont exactement les polynômes minimaux de u (division euclidienne), et que deux tels polynômes minimaux diffèrent d'un inversible, c'est-à-dire d'un scalaire non nul. □

À partir de la notion de sous-espace stable, et afin de procéder à la réduction d'un endomorphisme par récurrence sur la dimension de l'espace, il serait idéal de pouvoir décomposer E en une somme directe (non triviale) de sous-espaces stables. En particulier, on aimerait être en mesure de construire un supplémentaire stable à un sous-espace stable. Malheureusement, ce n'est pas toujours possible. Mais le résultat crucial suivant est un cas particulier fondamental où on obtient une telle décomposition en sous-espaces stables.

Lemme 4.7 (lemme des noyaux)**.** *Soient $P, Q \in K[X]$ premiers entre eux.*
Alors $\ker((PQ)(u)) = \ker(P(u)) \bigoplus \ker(Q(u))$.

Démonstration: L'anneau $K[X]$ étant principal, on écrit une relation de Bézout entre P et Q : il existe deux polynômes $U, V \in K[X]$ tels que $U \cdot P + V \cdot Q = 1$. En appliquant cette égalité à l'endomorphisme u, on obtient

$$U(u) \circ P(u) + V(u) \circ Q(u) = \mathrm{id}_E \,.$$

On en déduit immédiatement que $\ker(P(u)) \cap \ker(Q(u)) = \{0\}$.

Soit maintenant $x \in ker((PQ)(u))$. On a $x = (UP)(u)(x) + (VQ)(u)(x)$. On affirme alors que $(VQ)(u)(x) \in \ker(P(u))$ et $(UP)(u)(x) \in \ker(Q(u))$. Vérifions la première propriété, la seconde étant similaire. On calcule $P(u)(VQ)(u)(x))$. On a

$$P(u)(VQ)(u)(x)) = (PVQ)(u)(x) = V((PQ)(u)(x)) = V(0) = 0$$

car $(PQ)(u)(x) = 0$ par hypothèse. Donc $(VQ)(u)(x) \in \ker(P(u))$.

Cela conclut la preuve. □

Une récurrence simple implique la généralisation suivante :

Corollaire 4.8. *Soit $u \in \mathcal{L}(E)$, et $P_1, \ldots, P_r \in K[X]$ des polynômes deux-à-deux premiers entre eux. Alors*

$$\ker((P_1 \ldots P_r)(u)) = \bigoplus_{i=1}^{r} \ker(P_i(u)) \,.$$

Une application fondamentale - en vue de la réduction de l'endomorphisme u - du lemme des noyaux est la suivante :

Corollaire 4.9. *Soit $P \in K[X]$ un polynôme annulateur non nul de u, et notons $P = \prod_i P_i^{m_i}$ sa décomposition en polynômes irréductibles, avec P_i et P_j non proportionnels si $i \neq j$.*
Alors $E = \bigoplus_i \ker(P_i^{m_i})$.

Ce corollaire sera particulièrement utile lorsque le polynôme P sera scindé ($\deg P_i = 1$) : si le polynôme P est scindé, alors E est somme directe des sous-espaces caractéristiques de u.

4.2 Polynôme caractéristique et théorème de Cayley-Hamilton

Dans cette sous-partie, on démontre un résultat central concernant les polynômes annulateurs d'un endomorphisme.

Définition 4.10. Soit $A \in \mathrm{M}_n(K)$.
Le polynôme caractéristique de A, noté $\chi_A(X)$, est défini par

$$\chi_A(X) := \det(A - X \cdot I_n) \in K[X]\,.$$

Remarquons que dans cette définition, la matrice $A - X \cdot I_n$ est a priori à coefficients dans $K[X]$, qui n'est pas un corps. Si on veut lui appliquer les résultats précédents, et notamment donner un sens à ce déterminant, on peut par exemple voir $K[X]$ comme un sous-anneau du corps $K(X)$, ce qui permet bien de définir le déterminant précédent comme une fraction rationnelle dans $K(X)$. Puis, en utilisant par exemple les propositions 3.90 ou 3.94, on voit que le déterminant d'une matrice de $K(X)$ dont les coefficients sont dans $K[X]$ est en fait dans $K[X]$, puisque le déterminant est une application polynomiale (à coefficients dans $\mathbf{Z}$) en les coefficients de la matrice. On peut aussi utiliser le début de la section 5.

Exemples 4.11.
1. La fonction qui à une matrice $A \in \mathrm{M}_n(\mathbf{R})$ associe son polynôme caractéristique $\chi_A \in \mathbf{R}_n[X]$ est continue. Qu'en est-il du polynôme minimal ?
2. $\mathrm{GL}_n(K)$ est dense dans $\mathrm{M}_n(K)$ si $K = \mathbf{R}$ ou $\mathbf{C}$ (nouvelle preuve), et $\mathrm{SL}_n(K)$ est d'interieur vide.

Cherchons maintenant à étendre cette définition pour obtenir le polynôme caractéristique d'un endomorphisme :

Définition 4.12. Soit $u \in \mathcal{L}(E)$.
Le polynôme caractéristique de u, noté $\chi_u(X)$, est défini par

$$\chi_u := \chi_{\mathrm{Mat}_{\mathcal{B}}(u)} \in K[X]\,,$$

où $\mathcal{B}$ est une base de E. En particulier, ce polynôme ne dépend pas de la base $\mathcal{B}$ choisie.

Il est clair que le polynôme ne dépend pas de la base, en utilisant la formule de changement de bases dans $\mathrm{M}_n(K(X))$.

La motivation principal pour l'introduction de ce polynôme est notamment le fait élémentaire suivant :

Lemme 4.13. *Soit $\lambda \in K$.*
Alors $\chi_u(\lambda) = 0$ si et seulement si λ est valeur propre de u.
Autrement dit, les racines de χ_u dans K sont exactement les valeurs propres de u.

Démonstration: Le morphisme $K[X] \to K$ d'évaluation en λ étant un morphisme d'anneaux, et le déterminant d'une matrice étant ponynomial en les coefficients, on voit que $\chi_u(\lambda) = \det(u - \lambda \cdot \mathrm{id}_E)$ dans K. Ce scalaire vaut 0 si et seulement si $u - \lambda \cdot \mathrm{id}_E$ n'est pas inversible si et seulement l'espace propre $\ker(u - \lambda \cdot \mathrm{id}_E)$ n'est pas réduit à 0. □

Définition 4.14. Soit λ une valeur propre de u, de multiplicité algébrique m (i.e. λ est une racine de χ_u de multiplicité m).

Le sous-espace caractéristique associé à λ est $\ker((u - \lambda \cdot \mathrm{id}_E)^m)$.

Remarquons que le sous-espace caractéristique est un espace stable par u, qui contient le sous-espace propre correspondant.

Rappelons qu'un calcul simple de dimension assure que u admet un polynôme annulateur de degré $\leq n^2$. Le théorème suivant permet de montrer que u admet un polynôme annulateur de degré n - le polynôme caractéristique en l'occurence.

Théorème 4.15 (Cayley-Hamilton)**.** *Pour tout $u \in \mathcal{L}(E)$,*

$$\chi_u(u) = 0\,.$$

Autrement dit, le polynôme minimal π_u divise le polynôme caractéristique χ_u.

Démonstration: On propose plusieurs preuves :

1. Via la formule de la comatrice : on admet (voir section 5 pour la preuve) que la formule de la comatrice (proposition 3.96) est valable pour des matrices à coefficients dans un anneau commutatif. En particulier, on en déduit l'égalité suivante dans $\mathrm{M}_n(K[X])$:

$$^t\mathrm{Com}(A - X \cdot I_n) \cdot (A - X \cdot I_n) = \chi_A(X) \cdot I_n\,.$$

   On voit désormais les trois matrices $P := {}^t\mathrm{Com}(A - X \cdot I_n)$, $Q := A - X \cdot I_n$ et $R := \chi_A(X) \cdot I_n$ comme des polynômes à coefficients dans l'anneau (non commutatif) $\mathrm{M}_n(K)$, via la bijection évidente $\mathrm{M}_n(K[X]) \cong \mathrm{M}_n(K)[X]$. Remarquons le point particulier (crucial) suivant : le polynôme $Q = A - X \cdot I_n$ a ses coefficients qui commutent avec A. Cela assure (le vérifier !) que $(P \cdot Q)(A) = P(A) \cdot Q(A)$. Or de façon évidente $Q(A) = 0$ et $R(A) = \chi_A(A) \cdot I_n$, ce qui assure que $\chi_A(A) \cdot I_n = 0$, donc $\chi_A(A) = 0$.
2. Via la matrice universelle et le corps de décomposition : on considère la matrice universelle de taille n, i.e. la matrice $A := (X_{i,j})_{1 \leq i,j \leq n}$, à coefficients dans $\mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n]$ (où les $X_{i,j}$ sont des indéterminées). Notons k le corps $\mathbf{Q}(X_{i,j})$ contenant $\mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n]$. Alors $\chi_A(X) := \det(A - X \cdot I_n) \in (\mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n])[X]$. Introduisons maintenant ℓ un corps de décomposition de $\chi_A(X)$ sur k : ℓ/k est une extension finie engendrée par les racines de $\chi_A(X)$. En choisissant une matrice diagonale à valeurs propres distinctes $A_0 \in \mathrm{M}_n(\mathbf{C})$, on considère l'unique morphisme d'anneaux $\varphi : \mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n] \to \mathbf{C}$ envoyant les indéterminées $X_{i,j}$ sur les coefficients de A_0. Alors $\chi_{A_0}(X)$ s'obtient en remplaçant les variables $X_{i,j}$ dans $\chi_A(X)$ par les coefficients de A_0. Or χ_{A_0} a n racines distinctes, donc χ_A également (cette dernière propriété peut se montrer en utilisant le discriminant du polynôme χ_A, qui se spécialise sur celui de χ_{A_0}, donc est non nul).

En particulier, A est diagonalisable (via le lemme des noyaux) dans ℓ, et il est clair qu'alors $\chi_A(A) = 0$ dans ℓ, donc dans k, donc dans $\mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n]$.
Soit alors K un corps et $B \in \mathrm{M}_n(K)$. Alors il existe un unique morphisme d'anneaux $\psi : \mathbf{Z}[X_{i,j}; 1 \leq i, j \leq n] \to K$ envoyant A sur B. Alors $\chi_B(B) = \varphi(\chi_A(A)) = \varphi(0) = 0$.

□

Une troisième preuve, plus élémentaire, du théorème de Cayley-Hamilton, sera donnée plus loin, au corollaire 4.62.

4.3 Premiers résultats de réduction : diagonalisation, trigonalisation

On définit les notions de diagonalisation et trigonalisation des endomorphismes, et on établit des conditions nécessaires et suffisantes simples pour décider si un endomorphisme est diagonalisable (resp. trigonalisable).

Définition 4.16. Soit $u \in \mathcal{L}(E)$.
On dit que u est diagonalisable (resp. trigonalisable) s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale (resp. triangulaire supérieure).

On dira de même qu'une matrice est diagonalisable (resp. trigonalisable) si l'endomorphisme associé à la matrice l'est. Autrement dit, $A \in \mathrm{M}_n(K)$ est diagonalisable (resp. trigonalisable) si et seulement s'il existe $P \in \mathrm{GL}_n(K)$ telle que $P \cdot A \cdot P^{-1}$ est diagonale (resp. triangulaire supérieure).

Exemples 4.17.

On dispose de deux critères très utiles (conditions nécessaires et suffisantes) pour décider si un endomorphisme est diagonalisable (resp. trigonalisable). Avant d'énoncer ces critères, formulons d'abord quelques propriétés simples :

Proposition 4.18. *Soit $u \in \mathcal{L}(E)$. Alors les conditions suivantes sont équivalentes :*

1. *u est diagonalisable.*
2. *E est somme des sous-espaces propres de u.*
3. *E est somme directe des sous-espaces propres de u.*
4. *E admet une base formée de vecteus propres de u.*

Démonstration:

— Montrons (1) $\implies$ (2) : il existe une base $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale. Alors les vecteurs de $\mathcal{B}$ sont des vecteurs propres de u, donc E est engendré par des vecteurs propres de u.

— Montrons (2) $\implies$ (3) : il suffit de montrer que les espaces propres de u sont en somme directe. C'est par exemple une conséquence du lemme des noyaux : si l'on note $\lambda_1, \ldots, \lambda_r \in K$ les valeurs propres deux-à-deux distinctes de u, alors les polynômes $X - \lambda_i$ sont deux-à-deux premiers entre eux, donc les espaces propres $\ker(u - \lambda_i \cdot \mathrm{id}_E)$ sont en somme directe.

— Montrons (3) $\implies$ (4) : on choisit des bases $\mathcal{B}_i$ de chacun des sous-espaces propres de u. On concatène ces bases pour obtenir une base $\mathcal{B}$ de E. Alors $\mathcal{B}$ est la base souhaitée.

— Montrons (4) $\implies$ (1) : c'est évident.

□

Définition 4.19. Soit E un K-espace vectoriel de dimension n.

Un drapeau de E est une suite $(F_0, \dots, F_n)$ de sous-espaces vectoriels de E, tels que pour tout $0 \leq i \leq n-1$, $F_i \subset F_{i+1}$, et pour tout $0 \leq i \leq n$, $\dim(E_i) = i$.

Proposition 4.20. *Soit $u \in \mathcal{L}(E)$.*

Alors u est trigonalisable si et seulement si u stabilise un drapeau de E, i.e. il existe un drapeau de E formé de sous-espaces stables par u.

Démonstration:

— On suppose u trigonalisable. On dispose d'une base $\mathcal{B} = (e_1, \dots, e_n)$ dans laquelle u est triangulaire supérieure. On note $F_i := \text{Vect}(e_1, \dots, e_i)$. Alors il est clair que (F_i) est un drapeau de E stable par u.

— On suppose qu'il existe un drapeau (F_i) stable par u. On choisit, pour tout $1 \leq i \leq n$, un vecteur $e_i \in F_i \setminus F_{i-1}$. Une récurrence simple assure que $\mathcal{B} := (e_1, \dots, e_n)$ est libre, donc c'est une base de E. Alors par construction, $\text{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure.

□

Passons maintenant au lien entre diagonalisabilité (resp. trigonalisabilité) et polynômes annulateurs, afin d'obtenir les critères les plus utiles en pratique.

Théorème 4.21. *Soit $u \in \mathcal{L}(E)$.*

Alors u est diagonalisable si et seulement si u admet un polynôme annulateur scindé à racines simples si et seulement si π_u est scindé à racines simples.

Par exemple, Cayley-Hamilton assure que si χ_u est scindé à racines simples, alors u est diagonalisable.

Exemples 4.22. 1. Soit u et $n \geq 1$ tel que $u^n = \text{id}_E$. Si K contient les racines n-ièmes de l'unité et si n est premier à la caractéristique de k, alors u est diagonalisable. Par exemple, en charactéristique différente de 2, les symétries (endomorphismes vérifiant $u^2 = \text{id}_E$) sont diagonalisables.

Démonstration:

— On suppose u diagonalisable. Il existe une base $\mathcal{B}$ telle que $\text{Mat}_{\mathcal{B}}(u) = \text{diag}(\lambda_1, \dots, \lambda_n)$, où les λ_i sont exactement les valeures propres de u dans K. Alors si $\lambda_1, \dots, \lambda_r$ désigne les valeurs propres deux-à-deux distinctes de u, il est clair que le polynôme $\prod_{i=1}^{r}(X - \lambda_i)$ est scindé à racines simples dans K, et il annule la matrice $\text{Mat}_{\mathcal{B}}(u)$, donc u.

— On suppose que u admet un polynôme annulateurs P scindé à racines simples. Alors π_u divise P, donc π_u est scindé à racines simples.

— On suppose que π_u est scindé à racines simples. On écrit $\pi_u = (X - \lambda_1) \cdots (X - \lambda_r)$, avec $\lambda_i \in K$. Comme les λ_i sont exactement les valeurs propres de K, le lemme des noyaux assure que E est somme directe des espaces propres de U. Donc u est diagonalisable.

□

Corollaire 4.23. *Soit $K = \mathbf{F}_q$ un corps fini de cardinal q, et E un K-espace vectoriel de dimension n.*

1. *Un endomorphisme $u \in \mathcal{L}(E)$ est diagonalisable si et seulement si $u^q = u$.*
2. *Le nombre d'endomorphismes diagonalisables dans $\mathrm{M}_n(\mathbf{F}_q)$ est égal à*

$$\sum_{\substack{(m_1,\dots,m_q) \in \mathbf{N}^q \\ \sum_i m_i = n}} \frac{|\mathrm{GL}_n(\mathbf{F}_q)|}{|\mathrm{GL}_{m_1}(\mathbf{F}_q)| \dots |\mathrm{GL}_{m_q}(\mathbf{F}_q)|}.$$

Démonstration:

1. Montrons d'abord que les polynômes scindés à racines simples sur K sont exactement les diviseurs de $X^q - X$. En effet, $X^q - X = \prod_{x \in K}(X - x)$, ce qui assure le résultat. Ensuite, u est diagonalisable si et seulement s'il existe $P \in K[X]$ scindé à racines simples tel que $P(u) = 0$ si et seulement si u est annulé par un diviseur de $X^q - X$ si et seulement si $u^q = u$.
2. On fait agir le groupe $\mathrm{GL}_n(\mathbf{F}_q)$ sur l'ensemble $D_n(\mathbf{F}_q)$ des matrices diagonalisables dans $\mathrm{M}_n(\mathbf{F}_q)$, par conjugaison. Montrons que deux matrices $A, B \in D_n(\mathbf{F}_q)$ sont dans la même orbite si et seulement si pour tout $\lambda \in \mathbf{F}_q$, $\dim \ker(A - \lambda \cdot I_n) = \dim \ker(B - \lambda \cdot I_n)$. Il est clair que deux matrices dans la même orbite vérifie cette propriété. Réciproquement, si A et B vérifie cette propriété, comme A et B sont diagonalisables, on a $\mathbf{F}_p^n = \bigoplus_{\lambda \in \mathbf{F}_q} E_\lambda(A) = \bigoplus_{\lambda \in \mathbf{F}_q} E_\lambda(B)$; donc en utilisant l'égalité des dimensions, pour tout $\lambda \in \mathbf{F}_q$, il existe un isomorphisme $_\lambda : E_\lambda(A) \xrightarrow{\sim} E_\lambda(B)$. Comme les $E_\lambda(A)$ sont en somme directe, cela définit un automorphisme φ de $\mathbf{F}_q^n$ envoyant les espaces propres de A sur ceux de B. Par conséquent, on en déduit que $\varphi \circ A \circ \varphi^{-1} = B$, donc A et B sont dans la même orbite. Par conséquent, l'ensemble des orbites de cette action est en bijection avec l'ensemble des $(m_\lambda)_{\lambda \in \mathbf{F}_q} \in \mathbf{N}^{\mathbf{F}_q}$ tels que $\sum_\lambda m_\lambda = n$.
   Calculons maintenant le stabilisateur d'une orbite donnée par $(m_\lambda)_{\lambda \in \mathbf{F}_q}$. Fixons une décomposition (par exemple celle construite naturellement à l'aide de la base canonique) $\mathbf{F}_q^n = \bigoplus_{\lambda \in \mathbf{F}_q} E_\lambda$ et une matrice $A \in D_n(\mathbf{F}_q)$ compatible à cette décomposition, i.e. telle que $A_{|E_\lambda} = \lambda \mathrm{id}$ pour tout λ. Calculons le stabilisateur de A. Pour tout $M \in \mathrm{GL}_n(\mathbf{F}_q)$, on a $M \cdot A \cdot M^{-1} = A$ si et seulement si M stabilise les espaces propres de A si et seulement si pour tout λ, $M(E_\lambda) \subset E_\lambda$ si et seulement si pour tout λ, $M(E_\lambda) = E_\lambda$. Par conséquent, en fixant une base de $\mathbf{F}_q^n$ compatible à la décomposition $\mathbf{F}_q^n = \bigoplus_{\lambda \in \mathbf{F}_q} E_\lambda$ (par exemple la base canonique), on a un isomorphisme entre le stabilisateur de A et le groupe $\prod_{\lambda \in \mathbf{F}_q} \mathrm{GL}_{m_\lambda}(\mathbf{F}_q)$ (qui envoie M sur la matrice de $M_{|E_\lambda}$ pour tout λ).
   On écrit alors l'équation aux classes pour conclure.

□

Étudions maintenant les endomorphismes trigonalisables :

Théorème 4.24. *Soit $u \in \mathcal{L}(E)$.*

Alors u est trigonalisable si et seulement si u admet un polynôme annulateur scindé si et seulement si π_u est scindé si et seulement si χ_u est scindé.

Exemples 4.25.
1. Si K est algébriquement clos, tout endomorphisme est trigonalisable.

Démonstration: On vérifie d'abord que χ_u scindé équivaut à π_u scindé équivaut à l'existence d'un polynôme annulateur scindé. Il est clair que l'existence d'un polynôme annulateur scindé implique π_u scindé. Montrons que π_u et χ_u ont les mêmes facteurs irréductibles. Fixons une base de E et notons A la matrice de u dans cette base. Soit P irréductible divisant χ_u. On note L un corps de rupture de P sur K. Alors P a une racine $\lambda \in L$. Donc λ est valeur propre de A dans $\mathrm{M}_n(L)$, avec un vecteur propre $x \in L^n$ associé. Donc $\pi_u(A)(x) = \pi_u(\lambda) \cdot x$. Or $\pi_u(A) = 0$, donc $\pi_u(\lambda) = 0$, donc λ est racine de π_u dans L. Or P est le polynôme minimal de λ sur K, donc P divise π_u. Cela assure les équivalences souhaitées.

Intéressons-nous maintenant à l'équivalence entre ces propriétés et la trigonalisabilité. Si u est trigonalisable, il est clair que χ_u est scindé. Pour la réciproque, on propose deux démonstrations :

— Un argument de dualité : on raisonne par récurrence sur la dimension n de E (le cas $n = 0$ étant évident). On considère l'endomorphisme dual ${}^t u : E^* \to E^*$. On sait que $\chi_{u^*} = \chi_u$. Par conséquent, χ_{u^*} est scindé. En particulier, ce polynôme admet une racine $\lambda \in K$, qui est donc valeur propre de u^*. Soit $D \subset E^*$ une droite propre pour u^* associée à la valeur propre λ. Alors le sous-espace $F := D^\circ \subset E$ est un hyperplan de E, stable par u (le vérifier !). On peut donc considérer $u_F \in \mathcal{L}(F)$ la restriction de u à F. Puisque F est stable, on voit que χ_{u_F} divise χ_u. Donc χ_{u_F} est scindé, donc par hypothèse de récurrence, u_F est trigonalisable. Il existe donc une base $\mathcal{B}'$ de F telle que $\mathrm{Mat}_{\mathcal{B}'}(u_F)$ soit triangulaire supérieure. On choisit un vecteur quelconque $x \in E \setminus F$, et on note $\mathcal{B} := \mathcal{B}' \cup \{x\}$. Alors $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure.

— Un calcul matriciel par blocs (i.e. un argument de quotient) : on raisonne par récurrence sur la dimension n de E (le cas $n = 0$ étant évident). χ_u est scindé, donc u admet une valeur propre $\lambda \in K$. Il existe un vecteur propre $v_1 \in E$ associé à λ. On complète v_1 en une base $\mathcal{B}$ de E, de sorte que

$$\mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c} \lambda & L \\ \hline 0 & A \end{array}\right),$$

avec $A \in \mathrm{M}_{n-1}(K)$.
Alors $\chi_u = (\lambda - X) \cdot \chi_A$, donc χ_A est scindé. Par hypothèse de récurrence, A est trigonalisable, donc il existe $P \in \mathrm{GL}_{n-1}(K)$ telle que $P \cdot A \cdot P^{-1}$ est triangulaire supérieure. On note alors

$$Q := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right) \in \mathrm{GL}_n(K),$$

et on vérifie par un simple calcul par blocs que $Q \cdot \mathrm{Mat}_{\mathcal{B}}(u) \cdot Q^{-1}$ est triangulaire supérieure.

□

Corollaire 4.26. *1. l'adhérence de $\mathbf{D}_n(\mathbf{R})$ dans $\mathrm{M}_n(\mathbf{R})$ est $\mathbf{T}_n(\mathbf{R})$, et l'intérieur de $D_n(\mathbf{R})$ est l'ensemble des matrices réelles diagonalisables à valeurs propres distinctes (i.e. dont le polynôme caractéristique est scindé à racines simples).*

2. $\mathbf{D}_n(\mathbf{C})$ est dense dans $\mathbf{T}_n(\mathbf{C}) = \mathrm{M}_n(\mathbf{C})$. Son intérieur est l'ensemble des matrices complexes diagonalisables à valeurs propres distinctes (i.e dont le polynôme caractéristique est à racines simples).

Démonstration:

1. pour l'intérieur, on peut utiliser le résultant (i.e. le discriminant).
2. plus facile.

□

Quelques autres propriétés concernant la topologie des classes de similitude :

Proposition 4.27. *Si $K = \mathbf{R}$ ou $\mathbf{C}$, et $A \in \mathrm{M}_n(K)$, on a*

1. *La classe de similitude de A est bornée si et seulement si A est une matrice scalaire (i.e. une homothétie).*
2. *Si $K = \mathbf{C}$, la classe de similitude de A est fermée si et seulement si A est diagonalisable.*

Démonstration:

1. On suppose A non diagonale. Alors il existe $i \neq j$ tel que $a_{i,j} \neq 0$. Alors pour tout $\lambda \in K^\times$, $D_i(\lambda) \cdot A \cdot D_i(\lambda)^{-1}$ a un coefficient d'indice (i,j) égal à $\lambda \cdot a_{i,j}$, qui n'est pas borné quand λ décrit $K^\times$. Supposons maintenant $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ diagonale non scalaire : il existe $i \neq j$ tel que $\lambda_i \neq \lambda_j$. Le coefficient d'indice (i,j) de $T_{i,j}(t) \cdot A \cdot T_{i,j}(t)^{-1}$ est égal à $t \cdot (\lambda_j - \lambda_i)$, qui n'est pas borné quand t décrit K.

   Donc : si la classe de similitude est bornée, alors A est scalaire. Réciproquement, une matrice scalaire est centrale, donc sa classe de similitude est un singleton, donc bornée.
2. Soit A diagonalisable, et soit (B_n) une suite de matrices conjugués à A, i.e. $B_n = P_n \cdot A \cdot P_n^{-1}$ pour tout n, telle que (B_n) converge vers B dans $\mathrm{M}_n(K)$. Comme le polynôme caractéristique est un invariant de similitude, et puisqu'il dépend continuement de la matrice, $\chi_B = \chi_A$. En outre, $\pi_A(B_n) = 0$ pour tout n, donc par passage à la limite, $\pi_A(B) = 0$, donc π_B divise π_A, donc π_B est scindé à racines simples. Donc B est diagonalisable, et est semblable à A (puisqu'elle a le même polynôme caractéristique que A). Donc la classe de similitude de A est fermée.

   Réciproquement, supposons la classe de similitude de A fermée. Comme $K = \mathbf{C}$, on peut supposer A triangulaire supérieure. Notons $P := \mathrm{diag}(1, 2, \ldots, n)$. Alors pour tout $k \geq 1$, le coefficient d'indice (i,j) de $P^k \cdot A \cdot P^{-k}$ vaut $\left(\frac{i}{j}\right)^k \cdot a_{i,j}$, donc pour $i < j$, ce coefficient tend vers 0. Donc la suite de matrices $P^k \cdot A \cdot P^{-k}$ (dans la classe de similitude de A) converge vers une matrice diagonale. Comme la classe de similitude est fermée, cette matrice est semblable à A, donc A est diagonalisable.

□

Parlons désormais un peu de réduction simultannée des endomorphismes.

Proposition 4.28. *Soient $u, v \in \mathcal{L}(E)$ tels que $u \circ v = v \circ u$.*
Pour tout $P \in K[X]$, $\ker(P(v))$ est stable par u.

Remarque 4.29. En revanche, il n'est pas vrai que tout sous-espace stable par v est stable par u.

En particulier, les sous-espaces propres de v sont stables par u.

Proposition 4.30. *Soit $(u_i)_{i\in I}$ une famille de d'endomorphismes de E. Supposons que les u_i commutent entre eux.*

Si tous les u_i sont diagonalisables (resp. trigonalisables), alors ils sont co-diagonalisables (resp. co-trigonalisables), i.e. il existe une base $\mathcal{B}$ de E telle que pour tout $i \in I$, $\mathrm{Mat}_{\mathcal{B}}(u_i)$ est diagonale (resp. triangulaire supérieure).

Démonstration: On raisonne par récurrence sur la dimension n de E (les cas $n = 0$ et $n = 1$ sont évidents).

— Cas diagonalisable. Si tous les u_i sont des homothéties, le résultat est clair. On suppose donc qu'il existe i_0 tel que u_{i_0} n'est pas une homothétie.
Alors $E = \bigoplus_{\lambda\in \mathrm{Sp}(u_{i_0})} \ker(u-\lambda\cdot \mathrm{id}_E)$, et tous les $\ker(u-\lambda\cdot\mathrm{id}_E)$ sont stables par tous les u_i grâce à la proposition 4.28. Puisque u_{i_0} n'est pas une homothétie, il existe au moins deux valeurs propres distinctes dans $\mathrm{Sp}(u_{i_0})$, et donc pour toute telle valeur propre λ, $\dim(\ker(u_{i_0} - \lambda\cdot\mathrm{id}_E)) < n$. On applique donc l'hypothèse de récurrence aux restrictions des u_i à chacun des sous-espaces propres de u_{i_0} (on rappelle que la restriction d'un endomorphisme diagonalisable à un sous-espace stable est diagonalisable (par exemple, le polynôme minimal de la restriction divise celui de l'endomorphisme initial). L'hypothèse de récurrence assure que pour tout $\lambda \in \mathrm{Sp}(u_{i_0})$, il existe une base $\mathcal{B}_\lambda$ de $\ker(u_{i_0} - \lambda\cdot\mathrm{id}_E)$ telle que pour tout i, la restriction de u_i à $\ker(u_{i_0} - \lambda\cdot\mathrm{id}_E)$ a une matrice diagonale dans la base $\mathcal{B}_\lambda$. Alors la base $\mathcal{B}$ obtenue en concaténant les bases $\mathcal{B}_\lambda$ est une base de diagonalisation commune des u_i.

— Cas trigonalisable (ce second cas peut-être adapté pour contenir également une preuve du premier cas). On note $V \subset \mathcal{L}(E)$ le sous K-espace vectoriel de $\mathcal{L}(\mathcal{E})$ engendré par les u_i. Ce sous-espace V est de dimension finie sur K, et on peut extraire de la famille (u_i) une base $(v_1, \ldots, v_k)$ de V. Il est alors clair que si les v_i sont simultanément trigonalisables, alors les u_i le sont également. Par hypothèse, v_1 admet une valeur propre λ_1. L'espace propre $\ker(v_1 - \lambda_1\cdot\mathrm{id}_E)$ est stable par v_2, et la restriction de v_2 à ce sous-espace est trigonalisable, donc admet une valeur propre λ_2. En particulier, $\ker(v_1 - \lambda_1\cdot\mathrm{id}_E) \cap \ker(v_2 - \lambda_2\cdot\mathrm{id}_E)$ est un sous-espace non nul de E stable par tous les u_i. Une récurrence simple assure ainsi que E admet un vecteur $x \neq 0$, qui est vecteur propre pour v_1, …, v_k. Alors les endomorphismes v_i induisent des endomorphismes $\overline{v_i}$ de $E/K\cdot x$. Clairement, les $\overline{v_i}$ commutent deux-à-deux, et ils sont trigonalisables (leur polynôme caractéristique divise celui des v_i). Or $\dim(E/K\cdot x) < n$, donc par hypohtèse de récurrence, il existe une base $\overline{B}$ de $E/K\cdot x$ qui trigonalise simultanément les $\overline{v_i}$. On relève $\overline{B}$ en une famille libre $\mathcal{B}'$ de E, et la base $\mathcal{B} := \mathcal{B}' \cup \{x\}$ est une base de trigonalisation simultanée des v_i, donc des u_i.

□

Exemple 4.31. Si $\mathrm{car}(K) \neq 2$, les groupes $\mathrm{GL}_n(K)$ et $\mathrm{GL}_m(K)$ sont isomorphes si et seulement si $n = m$.

Remarque : cette propriété reste vraie en caractéristique 2, mais la preuve est différente. En caractéristique $p > 0$, on peut utiliser le sous-groupe des matrices triangulaires supérieures avec des 1 sur la diagonale, et sa suite central descendante, qui caractérisent n.

Si des endomorphismes trigonalisables qui commutent sont simultanément trigonalisables, la réciproque est fausse. On dispose d'une généralisation de la proposition 4.30, appelé théorème de Lie-Kolchin :

Théorème 4.32. *Soit $G \subset \mathrm{GL}_n(\mathbf{C})$ un sous-groupe résoluble connexe.*

Alors il existe $P \in \mathrm{GL}_n(\mathbf{C})$ tel que $P \cdot G \cdot P^{-1} \subset \mathbf{T}_n(\mathbf{C}) \cap \mathrm{GL}_n(\mathbf{C})$. Autrement dit, les matrices de G sont co-trigonalisables.

Démonstration:

— Si G est abélien, c'est une conséquence de la proposition 4.30. On suppose donc désormais G non abélien.

— On montre que si G est un groupe topologique connexe, alors $D(G)$ est connexe.

— Montrons maintenant qu'il existe un sous-espace non trivial stable par tous les éléments de G.

Rappelons que puisque G est résoluble, il existe $m \geq 2$ minimal tel que $D^m(G) = \{I_n\}$. On introduit alors le sous-groupe non trivial $H := D^{m-1}(G) \subset G$. Puisque $D(H) = \{I_n\}$, alors H est abélien. Par la proposition 4.30, H stabilise un drapeau, donc en particulier, les éléments de H ont un vecteur propre $x \neq 0$ commun.

Notons alors $V := (G \cdot x)$. Par construction, $V \neq \{0\}$ est un sous-espace stable par tous les éléments de G. Supposons que $V = \mathbf{C}^n$.

On remarque d'abord que pour tout $g \in G$, pour tout $h \in H$, $h(g(x)) = g(g^{-1} \circ h \circ g(x))$, or H est distingué dans G, donc $g^{-1} \circ h \circ g \in H$, donc x est un vecteur propre pour $g^{-1} \circ h \circ g$, donc $g(x)$ est un vecteur propre (non nul) pour h.

Par conséquent, tous les $g(x)$, $g \in G$, sont des vecteurs propres par tous les éléments de H. En particulier, il existe une base $\mathcal{B}$ de $V = \mathbf{C}^n$ dans laquelle tous les éléments de H sont diagonaux. En outre, pour tout $h \in H$ et $g \in G$, $g \cdot h \cdot g^{-1} \in H$, donc est diagonale dans la base $\mathcal{B}$ et a le même spectre que h. Par conséquent, pour tout $h \in H$, la classe de conjugaison de h dans G est finie. Or G est connexe, donc cette classe est connexe. Par conséquent, la classe de conjugaison de h dans G est réduite à $\{h\}$, i.e. $H \subset Z(G)$.

En outre, pour tout $h \in H$ notons λ la valeur propre de h associée à x. Alors $x \in E_\lambda(h)$ et $E_\lambda(h)$ est stable par G (car $h \in Z(G)$), donc $E_\lambda = V = \mathbf{C}^n$. Donc h est une homothétie. Donc les éléments de H sont des homothéties. Or $H \subset D(G) \subset \mathrm{SL}_n(\mathbf{C})$, donc $\det(H) = \{1\}$, donc les éléments de H sont des homothéties de rapport une racine n-ième de l'unité, donc H est fini. Or le point précédent assure que H est connexe, donc $H = \{I_n\}$, ce qui est contradictoire.

On a donc montré par l'absurde que $V \neq \mathbf{C}^n$, i.e. il existe un sous-espace vectoriel $\{0\} \neq V \neq \mathbf{C}^n$ de $\mathbf{C}^n$ stable par G. Par conséquent, tout élement $g \in G$ induit un élément $\overline{G} \in \mathrm{GL}(\mathbf{C}^n/V)$. On dispose donc de morphismes de groupes $\varphi_1 : G \to \mathrm{GL}(V)$ et $\varphi_2 : G \to \mathrm{GL}(\mathbf{C}^n/V)$, définis par $\varphi_1(g) = g_{|V}$ et $\varphi_2(g) = \overline{g}$. Notons G_i l'image de G par φ_i.

— On conclut alors par récurrence sur la dimension de V : les sous-groupes G_1 et G_2 sont connexes et résolubles, donc par récurrence il existe des bases $\mathcal{B}_1$ de V et $\mathcal{B}_2$ de $\mathbf{C}^n/V$ dans lesquelles G_1 et G_2 sont triangulaires supérieurs. Alors la base $\mathcal{B}$ de $\mathbf{C}^n$ formée de la concaténation de $\mathcal{B}_1$ et d'une famille relevant $\mathcal{B}_2$ dans $\mathbf{C}^n$, permet de trigonaliser le groupe G dans $\mathrm{GL}_n(\mathbf{C})$.

□

4.4 Endomorphismes semi-simples

Dans cette sous-partie, on s'intéresse à une généralisation de la notion d'endomorphisme diagonalisable. Cette notion est vraiment utile quand le corps de base n'est pas algébriquement clos.

Définition 4.33. Soit $u \in \mathcal{L}(E)$. On dit que u est semi-simple si pour tout sous-espace $F \subset E$ stable par u, il existe un sous-espace $G \subset E$, stable par u, tel que $E = F \bigoplus G$.

On dit que u est simple si les seuls sous-espaces stables par u sont $\{0\}$ et E.

Autrement dit, un endomorphisme est semi-simple si et seulement si tout sous-espace stable admet un supplémentaire stable.

Par définition, un tel endomorphisme est diagonalisable par blocs, et chaque bloc est un endomorphisme simple. Si tous les blocs sont de dimension 1, on retrouve le cas des ensomorphismes diagonsalisables.

De façon analogue, on définit une matrice semi-simple comme une matrice dont l'endomorphisme associé est semi-simple.

On dispose d'un critère pour la semi-simplicité, en lien avec les polynômes annulateurs de l'endomorphisme.

Théorème 4.34. *Soit $u \in \mathcal{L}(E)$.*

- *Les assertions suivantes sont équivalentes :*
  1. *u est simple.*
  2. *χ_u est irréductible.*
- *Les assertions suivantes sont équivalentes :*
  1. *u est semi-simple.*
  2. *Le polynôme π_u est sans facteur carré.*
  3. *u admet un polynôme annulateur sans facteur carré.*

Démonstration: □

On va maintenant s'intéresser au lien entre semi-simplicité et diagonalisabilité. Rappelons qu'un corps K est dit parfait si toute extension finie de K est séparable (i.e. tout polynôme irréductible dans $K[X]$ est scindé à racines simples dans une extension finie de K - un corps de décomposition de P par exemple). Les corps de caractéristique nulle sont parfaits, ainsi que les corps finis. En caractéristique $p > 0$, le corps K est parfait si et seulement si le morphisme (de corps) de Frobenius : $K \to K$, défini par $x \mapsto x^p$, est surjectif.

Théorème 4.35. *Soit $A \in \mathrm{M}_n(K)$.*

S'il existe une extension finie L/K telle que la matrice A est diagonalisable dans $\mathrm{M}_n(L)$, alors A est semi-simple.

Si le corps K est parfait, alors la réciproque est vérifiée.

On propose maintenant une classification des endomorphismes semi-simples sur $\mathbf{R}$:

Théorème 4.36. *On suppose $K = \mathbf{R}$. Soit $u \in \mathcal{L}(E)$.*

u est semi-simple si et seulement s'il existe une base $\mathcal{B}$ de E, des scalaires $\lambda_1, \ldots, \lambda_r \in \mathbf{R}$, $a_1, b_1, \ldots, a_s, b_s \in \mathbf{R}$, avec $b_i > 0$, tels que $n = r + 2s$ et

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ldots & \lambda_r & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & a_1 & -b_1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & b_1 & a_1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & \ldots & 0 & 0 & 0 & \ldots & a_s & -b_s \\ 0 & \ldots & 0 & 0 & 0 & \ldots & b_s & a_s \end{pmatrix}.$$

Autrement dit, les endomorphismes semi-simples réels sont diagonalisables par blocs, avec des blocs de taille 1 ou 2.

Démonstration: □

Exemple 4.37. Les matrices semi-simples dans $\mathrm{M}_n(\mathbf{R})$ sont exactement les matrices diagonalisables dans $\mathbf{C}$.

4.5 Décomposition de Dunford ; exponentielle de matrices

Dans cette partie, on va s'intéresser à la décomposition de Dunford des endomorphismes (et donc des matrices). Il s'agit de décomposer un endomorphisme en une somme d'un endomorphisme diagonalisable et d'un endomorphisme nilpotent, de façon canonique.

Définition 4.38. Soit $u \in \mathcal{L}(E)$.

On dit que u est nilpotent s'il existe un entier $m \geq 1$ tel que $u^m = 0$. Dans ce cas, l'entier m minimal tel que $u^m = 0$ est appelé indice de nilpotence de u.

Proposition 4.39. *Soit $u \in \mathcal{L}(E)$. Les conditions suivantes sont équivalentes :*

1. *u est nilpotent*
2. *u est trigonalisable et 0 est sa seule valeur propre.*
3. *Il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure stricte (avec des zéros sur la diagonale).*
4. *Le polynôme minimal de u est de la forme X^m.*
5. *Le polynôme caractéristique de u est $(-1)^n \cdot X^n$.*

Exemple 4.40. Soit $K = \mathbf{F}_q$ un corps fini de cardinal q. Le nombre de matrices nilpotentes dans $\mathrm{M}_n(\mathbf{F}_q)$ est exactement q^{n^2-n}. On peut donc dire que la probabilité qu'une matrice aléatoire (loi uniforme) dans $\mathrm{M}_n(\mathbf{F}_q)$ soit nilpotente est $\frac{1}{q^n}$.

La preuve passe par le lemme de Fitting qui permet d'associer à tout endomorphisme u de $E = K^n$, une décomposition de E en somme directe $E = F \bigoplus G$ de sous-esapces stables, telle que $u_F \in \mathrm{GL}(F)$ et U_G est nilpotente. ? ?

Lemme 4.41. *Soit $u, v \in \mathcal{L}(E)$ telles que $u \circ v = v \circ u$.*

Si u et v sont nilpotentes, alors $u + v$ et $u \circ v$ sont nilpotentes

Proposition 4.42. *Un endomorphisme nilpotent et diagonalisable est nul.*

On peut maintenant démontrer le théorème de décomposition de Dunford :

Théorème 4.43. *Soit $u \in \mathcal{L}(E)$ tel que χ_u est scindé.*
Il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$ tel que
— *$u = d + n$*
— *d est diagonalisable et u est nilpotent.*
— *$d \circ n = n \circ d$.*
En outre, d et n sont des polynômes en u.

Démonstration:

1. Existence : puisque χ_u est scindé, le théorème de Cayley-Hamilton (voir théorème 4.15) et le lemme des noyaux (cf corollaire 4.9) assurent que E somme directe des sous-espaces caractéristique de u, à savoir, si on note $\lambda_1, \dots, \lambda_r \in K$ les valeurs propres deux-à-deux distinctes de u, et $m_1, \dots, m_r$ leur multiplicité respective, on a
$$E = \bigoplus_{i=1}^{r} \ker\left((u - \lambda_i \cdot \mathrm{id})^{m_i}\right) .$$
Chacun de ces sous-espaces caractéristiques $E_i := \ker\left((u - \lambda_i \cdot \mathrm{id})^{m_i}\right)$ est stable par u, et on note $u_i \in \mathcal{L}(E_i)$ la restriction de u à E_i. Il suffit de définir d_i et n_i convenables sur chacun des sous-espaces E_i.
Pour ce faire, on pose $d_i := \lambda_i \cdot \mathrm{id}_{E_i}$ et $n_i := u_i - d_i$. Enfin, on note $d := \bigoplus_i d_i$ et $n := \bigoplus_i n_i$.
Alors d est clairement diagonalisable, $u = d + n$ et d et n commutent (vérifier ces propriétés sur chacun de E_i). Il reste à montrer que n est nilpotente, et que d et n sont des polynômes en u.
Pour tout i, $(X - \lambda_i)^{m_i}$ annule u_i, donc X^{m_i} annule n_i, donc $X^{\dim E}$ annule n, donc n est nilpotente.
Si $r = 1$, on constate que $n = (X - \lambda_1)(u)$ et $d = u - n$, donc d et n sont des polynômes en u. Si $r \geq 2$, les polynômes $P_1, \dots, P_r$, définis par $P_i := \prod_{j \neq i} \frac{\chi_u}{(X-\lambda_i)^{m_i}}$ sont premiers entre eux dans leur ensemble, donc le théorème de Bézout assure qu'il existe $Q_1, \dots, Q_r \in K[X]$ tels que $\sum_i Q_i \cdot P_i = 1$. On définit alors $P := \sum_i \lambda_i \cdot Q_i \cdot P_i$. Calculons alors $P(u)$. Il suffit de calculer $P(u_i)$, et comme $P_j(u_i) = 0$ pour tout $j \neq i$, on obtient que $P(u_i) = \lambda_i \cdot (Q_i \cdot P_i)(u_i) = \lambda_i \left(\sum_j Q_i \cdot P_i\right)(u_i) = \lambda_i \cdot \mathrm{id}_{E_i}$. Donc $P(u_i) = d_i$, pour tout i, donc $d = P(u)$, et $n = (X - P)(u)$.
2. Unicité : soit (d', n') une décomposition convenable de u. Alors $d + n = d' + n'$, donc $d - d' = n' - n$. Or d' et n' commutent, donc commutent à u, donc commutent à tout polynôme en u. En particulier, d et d' commutent, et n et n' commutent. Donc $d - d'$ est diagonalisable et $n' - n$ est nilpotente. Par conséquent, $d - d' = n' - n$ est diagonalisable et nilpotent, donc nul, donc $d = d'$ et $n = n'$.

□

Exemples 4.44. 1. Calcul de puissances de matrices.
2. Applications aux systèmes dynamiques discrets linéaires.

Proposition 4.45. *Si $K = \mathbf{R}$ ou $\mathbf{C}$, et $A \in \mathrm{M}_n(K)$, on a*

1. *L'adhérence de la classe de similitude de A contient 0 si et seulement si A est nilpotente.*
2. *Si χ_A est scindé et $A = D + N$ est la décomposition de Dunford de A, alors D est dans l'adhérence de la classe de similitude de A.*

Une variante : on peut étendre la décomposition de Dunford en supprimant l'hypothèse que le polynôme caractéristique est scindé, et en remplaçant "diagonalisable" par "semi-simple" :

Théorème 4.46. *Soit K un corps parfait et $u \in \mathcal{L}(E)$.*
Il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$ tel que
— $u = d + n$
— *d est semi-simple et u est nilpotent.*
— $d \circ n = n \circ d$.

Démonstration: Pour l'existence, on étend d'abord les scalaires à un corps de décomposition L du polynôme minimal de u. Alors le polynôme minimal de u (ou disons d'uune matrice A représentant u) est scindé dans L, donc on peut écrire la décomposition de Dunford de u à coefficients dans L : $u = d + n$, avec d diagonalisable sur L et n nilpotente. Comme la décomposition est unique, un petit argument de théorie de Galois (on utilise ici que K est parfait) assure que d et n sont en fait à coefficients dans K. Enfin, comme d est diagonalisable sur L, on en déduit que d est semi-simple sur K.

Pour l'unicité, c'est une conséquence immédiate de l'unicité sur L. □

Remarque 4.47. La décomposition de Dunford effective permet de démontrer la version semi-simple sans recours à la théorie de Galois. Nous reviendrons dessus si le temps le permet. ? ?

L'une des principales applications de la décomposition de Dunford, outre le calcul de puissances de matrices, est le calcul d'exponentielle de matrices.

Définition 4.48. Soit $K = \mathbf{R}$ ou $\mathbf{C}$ et $A \in \mathrm{M}_n(K)$.
On définit $\exp(A) := \sum_{k \in \mathbf{N}} \frac{A^k}{k!} \in \mathrm{M}_n(K)$.

Tout d'abord, $\exp(A)$ est bien définie car, en munissant $\mathrm{M}_n(K)$ d'une norme sous-multiplicative, on voit que la série $\sum_{k \in \mathbf{N}} \frac{A^k}{k!}$ converge normalement (et même uniformément sur tout compact) - on rappelle que $\mathrm{M}_n(K) \cong K^{n^2}$ est complet.

Proposition 4.49. *Soient $A, B \in \mathrm{M}_n(K)$.*
Si $A \cdot B = B \cdot A$, alors $\exp(A + B) = \exp(A) \cdot \exp(B)$.

Corollaire 4.50. *Pour tout $A \in \mathrm{M}_n(K)$, $\exp(A) \in \mathrm{GL}_n(K)$ et $exp(A)^{-1} = \exp(-A)$.*

Exemples 4.51.
1. Pour calculer $\exp(A)$, on peut écrire la décomposition de Dunford $A = D+N$ de A, avec D diagonalisable et N nilpotente, D et N commutant. Alors $\exp(A) = \exp(D) \cdot \exp(N)$. En outre, si l'on connaît les valeurs propres de A, il est aisé de calculer $\exp(D)$ via un polynôme interpolateur par exemple, et on a $\exp(N) = \sum_{k=0}^{n-1} \frac{A^k}{k!}$.

2. Application : résolution de systèmes d'équations différentielles linéaires à coefficients constants.

La décomposition de Dunford implique par exemple la propriété suivante :

Théorème 4.52. *1. L'application* $\exp : \mathrm{M}_n(\mathbf{C}) \to \mathrm{GL}_n(\mathbf{C})$ *est surjective.*

2. L'image de $\exp : \mathrm{M}_n(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ *est l'ensemble des* M^2*, pour* $M \in \mathrm{GL}_n(\mathbf{R})$*.*

Démonstration: □

4.6 Réduction de Jordan

On peut voir la réduction de Jordan des matrices comme un raffinement de la décomposition de Dunford. Cela va notamment permettre de décrire précisément les classes de similitude de matrices dont le polynôme caractéristique est scindé.

On définit d'abord les matrices $J_{\underline{p}}(\lambda)$. Tout d'abord, pour tout $p \geq 1$, on note $N_p \in \mathrm{M}_p(K)$ la matrice triangulaire supérieure avec des zéros sur la diagonale et des 1 juste au-dessus de la diagonale :

$$N_p := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Définition 4.53. Soit $n \in \mathbf{N}$, $\lambda \in K$ et $\underline{p} = (p_1, \dots, p_k)$ une partition de n, i.e. une suite décroissante d'entiers $p_i \geq 1$ telle que $p_1 + \cdots + p_k = n$. On définit la matrice $J_{\underline{p}}(\lambda)$, diagonale par blocs, de la façon suivante :

$$\begin{pmatrix} \lambda \cdot I_{p_1} + N_{p_1} & 0 & \dots & 0 \\ 0 & \lambda \cdot I_{p_2} + N_{p_2} & (0) & 0 \\ \vdots & (0) & \ddots & \vdots \\ 0 & \dots & 0 & \lambda \cdot I_{p_k} + N_{p_k} \end{pmatrix}.$$

C'est donc une matrice formée de blocs carrés diagonaux, de taille variable, de la forme :

$$\lambda \cdot I_p + N_p := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Théorème 4.54. *Soit* $u \in \mathcal{L}(E)$ *tel que* χ_u *est scindé. On note* $\lambda_1, \dots, \lambda_r \in K$ *les valeurs propres deux-à-deux distinctes de* u *et* $\underline{\lambda} := (\lambda_1, \dots, \lambda_r)$*.*

Alors pour tout $1 \leq i \leq r$*, il existe une partition* $\underline{p}_i$ *de* $\dim(E_{(\lambda_i)})$*, et une base* $\mathcal{B}$ *de* E *telle que* $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{p}_1, \dots, \underline{p}_r}(\underline{\lambda})$*.*

En outre, la famille $(\underline{p}_1, \lambda_1), \dots, (\underline{p}_r, \lambda_r)$ *est unique à l'ordre près.*

Démonstration: Comme χ_u est scindé, on peut décomposer E en somme directe des sous-espaces caractéristiques de u. On peut donc se restreindre à un sous-espace caractéristique, et donc supposer que le polynôme caractéristique de u est de la forme $\pm(X-\lambda)^n$. On note $1 \leq m \leq n$ le degré du polynome minimal de u, i.e. $\pi_u = (X-\lambda)^m$. Alors $u = \lambda \cdot \mathrm{id} + v$, avec v nilpotent d'indice m (c'est la décomposition de Dunford de u). On raisonne maintenant uniquement avec l'endomorphisme nilpotent v.

On propose ici deux preuves de ce théorème :

1. Méthode des noyaux itérés :

   Pour $0 \leq i \leq m$, on note $K_i := \ker(v^i)$. En particulier, on a $K_0 = \{0\}$ et $K_m = E$, et pour tout $0 \leq i \leq m-1$, $K_i \subset K_{i+1}$. Pour tout i, on note $q_i := \dim(K_i) - \dim(K_{i-1})$.

   Montrons que la suite $\underline{q} := (q_1, \ldots, q_m)$ est décroissante. On remarque que $v(K_{i+1}) \subset K_i$ pour tout i. En particulier, v induit une application linéaire $K_{i+1} \to K_i$, d'où en passant au quotient $v_i : K_{i+1} \to K_i/K_{i-1}$. Pour tout $x \in K_{i+1}$, on voit que $x \in \ker(v_i)$ si et seulement si $v(x) \in K_{i-1}$ si et seulement si $v^{i-1}(v(x)) = 0$ si et seulement si $x \in K_i$. Donc $\ker(v_i) = K_i$. Donc v_i induit une application linéaire injective $\overline{v}_i : K_{i+1}/K_i \to K_i/K_{i-1}$. Donc $\dim(K_{i+1}/K_i) \leq \dim(K_i/K_{i-1})$, i.e. $q_{i+1} \leq q_i$. Donc la suite $\underline{q}$ est décroissante. On a montré au passage que la restriction de v à tout supplémentaire de K_i dans K_{i+1} était une injection de ce supplémentaire dans un supplémentaire de K_{i-1} dans K_i.

   En outre, on a $\sum_{i=1}^m q_i = \dim(K_m) - \dim(K_0) = n$, donc $\underline{q}$ est une partition de n.

   On construit alors une base $\mathcal{B}$ adapté à v, de la façon suivante : on choisit une base $\mathcal{B}_m$ d'un supplémentaire de K_{m-1} dans K_m, puis on complète $v(\mathcal{B}_m)$ en une base $\mathcal{B}_{m-1}$ d'un supplémentaire de K_{m-2} dans K_{m-1}. On continue, en construisant par récurrence, pour tout $1 \leq i \leq m-1$, une base $\mathcal{B}_i$ d'un supplémentaire de K_{i-1} dans K_i, de sorte que $\mathcal{B}_i$ complète $v(\mathcal{B}_{i+1})$.

   La base $\mathcal{B}$ de E est alors obtenue en concaténant les familles $\mathcal{B}_i = (x_1^i, \ldots, x_{q_i}^i)$, dans l'ordre suivant :

   — on commence par prendre le premier vecteur de $\mathcal{B}_m$, le premier vecteur de $\mathcal{B}_{m-1}$, …, le premier vecteur de $\mathcal{B}_1$.
   — ensuite, on poursuit avec le deuxième vecteur de $\mathcal{B}_m$, …, le deuxième vecteur de $\mathcal{B}_1$.
   — on continue jusqu'à épuiser ainsi toutes les bases $\mathcal{B}_i$.

   Alors par construction la base $\mathcal{B}$ est de la forme

   $$(x_1^m, \ldots, v^{m-1}(x_1^m), x_2^m, \ldots, v^{m-1}(x_2^m), \ldots, x_{q_m}^m, \ldots, v^{m-1}(x_{q_m}^m), x_1^{m-1}, \ldots, v^{m-2}(x_1^{m-1}), \ldots, x_{q_{m-1}}^{m-1}, .$$

   ce qui assure que $\mathrm{Mat}_{\mathcal{B}}(v) = J_{\underline{p}}$, où $\underline{p}$ est la partition duale de $\underline{q}$.

   On en déduit alors que $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{p}}(\underline{\lambda})$.

2. Sous-espaces cycliques et argument de dualité :

   On raisonne par récurrence sur la dimension de E. Le cas où la dimension est nulle est évident.

   Tout d'abord, il est clair qu'il existe $x \in E$ tel qu'un générateur $\pi_{v,x} \in K[X]$ de l'idéal $\{P \in K[X] : P(v)(x) = 0\}$ est égal au polynôme minimal $\pi_v = X^m$ de v. On note E_x le sous-espace vectoriel $\mathrm{Vect}(x, v(x), \ldots, v^{m-1}(x))$, qui est

clairement stable par v et de dimension m. En outre, si $\mathcal{B}_1$ désigne la base $(x, v(x), \dots, v^{m-1}(x))$ de E_x, on a bien $\mathrm{Mat}_{\mathcal{B}_1}(v_x) = J_m$, $v_x \in \mathcal{L}(E_x)$ désignant la restriction de v à E_x.

Il s'agit maintenant de construire un supplémentaire F_x de E_x qui soit stable par u, et on appliquera l'hypothèse de récurrence à la restriction de v à F_x, qui reste nilpotent d'indice inférieur ou égal à m. Pour construire F_x, on utilise la dualité : puisque $v^{m-1}(x) \neq 0$, il existe une forme linéaire $\varphi \in E^*$ telle que $\varphi(v^{m-1}(x)) \neq 0$. Alors en particulier, on voit que ${}^t v^{m-1}(\varphi) \neq 0$, alors que ${}^t v^m = 0$ (puisque $v^m = 0$). En particulier, le sous-espace vectoriel $G_\varphi := (\varphi, {}^t u(\varphi), \dots, {}^t u^{m-1}(\varphi))$ de E^* est stable par ${}^t u$ et de dimension m. On en déduit facilement que son orthogonal $F_x := {G_\varphi}^\circ$ est un sous-espace vectoriel de E, stable de u, et de dimension $n - m$. Montrons que F_x est un supplémentaire de E_x : il suffit de montrer que $E_x \cap F_x = \{0\}$. Soit $y \in E_x \cap F_x$. Alors $y = \sum_{i=0}^{m-1} a_i \cdot v^i(x)$ et pour tout $0 \leq k \leq m-1$, ${}^t v^k(\varphi)(y) = 0$, donc pour tout $0 \leq k \leq m-1$, $\sum_{i=0}^{m-1} a_i \cdot \varphi(v^{i+k}(x)) = 0$. Or pour tout $j \geq m$, $v^j = 0$, donc l'égalité précédente donne

$$\begin{cases} a_0 \cdot \varphi(x) + \dots a_1 \cdot \varphi(v(x)) + \dots + a_{m-1} \cdot \varphi(v^{m-1}(x)) & = 0 \\ a_0 \cdot \varphi(v(x)) + \dots + a_{m-2} \cdot \varphi(v^{m-1}(x)) & = 0 \\ \dots & \\ a_0 \cdot \varphi(v^{m-2}(x)) + a_1 \cdot \varphi(v^{m-1}(x)) & = 0 \\ a_0 \cdot \varphi(v^{m-1}(x)) & = 0 \end{cases}.$$

C'est un système linéaire triangulaire en les a_i, dont tous les coefficients diagonaux sont égaux à $\varphi(v^{m-1}(x))$. Or par construction $\varphi(v^{m-1}(x)) \neq 0$, donc $a_i = 0$ pour tout i, donc $y = 0$ et $E_x \cap F_x = \{0\}$. Donc on a montré que F_x est un supplémentaire stable de E_x.

On peut donc conclure la preuve de l'existence par récurrence sur la dimension de E : il existe une base $\mathcal{B}_2$ de F_x dans laquelle la matrice de v est de la forme $J_{\underline{p}'}$, et dans la base $\mathcal{B}$ obtenue en concaténant $\mathcal{B}_1$ et $\mathcal{B}_2$, on obtient bien une matrice de la forme $J_{\underline{p}}$.

On en déduit alors que $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{p}}(\underline{\lambda})$.

Reste à discuter l'unicité... □

Remarque 4.55. Il est clair que le théorème 4.54 implique la décomposition de Dunford (théorème 4.43)

Corollaire 4.56. *Les classes de similitude de matrices nilpotentes dans* $\mathrm{M}_n(K)$ *sont en bijection avec l'ensemble des partitions de* n *; un système de représentants de ces classes de similitude est donné par les matrices* $J_{\mathbf{p}}$*, où* $\mathbf{p}$ *décrit l'ensemble des partitions de* n*.*

Corollaire 4.57. *Soit* $A \in \mathrm{M}_n(K)$*. Alors* A *et* ${}^t A$ *sont semblables.*

4.7 Réduction de Frobenius ; invariants de similitude

Si le théorème 4.54 décrit exactement certaines classes de similitude d'endomorphismes de E, il souffre de l'hypothèse (nécessaire) χ_u scindé. En particulier, il ne décrit pas toutes les classes de similitude, dès que le corps K n'est pas algébriquement

clos. Le résultat principal de cette partie permet de décrire toutes les classes de similitude d'endomorphismes, sans hypothèse supplémentaire (quel que soit le corps de base).

4.7.1 Endomorphismes cycliques ; matrices compagnon

Définition 4.58. Un endomorphisme $u \in \mathcal{L}(E)$ est dit cyclique s'il existe $x \in E$ tel que $\mathrm{Vect}(u^k(x), k \in \mathbf{N}) = E$.

Exemple 4.59. Un endomorphisme simple est cyclique.

Définition 4.60. Soit $P = \sum_{k=0}^{n} a_k \cdot X^k$ un polynôme unitaire de $K[K]$ (on a donc $a_n = 1$).

La matrice compagnon associée à P est la matrice $C(P) \in \mathrm{M}_n(K)$ définie par

$$C(P) := \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}.$$

Lemme 4.61. *Soit $P \in K[X]$. Alors*

$$\chi_{C(P)} = (-1)^n \cdot P\,.$$

Démonstration: En développant par rapport à la première ligne, on voit que $\chi_{C(P)} = (-1)^n \cdot a_0 - X\chi_{C(Q)}$, où $Q = \sum_{i=0}^{n-1} a_{i+1} \cdot X^i = \frac{P - a_0}{X}$. Une récurrence simple assure alors le résultat. Il est en fait plus rapide de développer directement par rapport à la dernière colonne. □

Corollaire 4.62 (Cayley-Hamilton)**.** *Pour tout $u \in \mathcal{L}(E)$, $\chi_u(u) = 0$.*

Démonstration: Soit $x \in E \setminus 0\}$. Notons $E_x := \mathrm{Vect}(u^k(x), k \in \mathbf{N})$. Ce sous-espace vectoriel de E est stable par u. Si k est le plus petit entier ≥ 1 tel que $(x, u(x), \dots, u^k(x))$ est liée, alors E_x est de dimension k, il existe $a_0, \dots, a_{k-1} \in K$ tels que $u^k(x) = -\sum_{i=0}^{k-1} a_i \cdot u^i(x)$, et une base de E_x est $(x, u(x), \dots, u^{k-1}(x))$. Dans cette base, la matrice de la restriction de u est $C(P)$, avec $P = X^k + \sum_{i=0}^{k-1} a_i X^i$. Donc le lemme assure que $\chi_{u_{|E_x}} = P$, et comme E_x est stable, $\chi_{u_{|E_x}}$ divise χ_u. Donc $P | \chi_u$, et par construction, $P(u)(x) = 0$, donc $\chi_u(u)(x) = 0$. Ceci étant valable pour tout $x \in E$, on a $\chi_u(u) = 0$. □

Proposition 4.63. *Soit $P \in K[X]$. Alors*

$$\pi_{C(P)} = P\,.$$

Démonstration: Si on note $(X_1, \dots, X_n)$ les vecteurs colonnes de la base canonique de K^n, alors $C(P)^{k-1} \cdot X_1 = X_k$ pour tout $1 \leq k \leq n$. En particulier, la famille $(X_1, C(P) \cdot X_1, \dots, C(P)^{n-1} \cdot X_1)$ est libre. Cela implique que le polynôme minimal de $C(P)$ est de degré au moins n. Or $\pi_{C(P)}$ divise $\chi_{C(P)}$ par Cayley-Hamilton, donc il divise P. Puisqu'ils ont même degré et qu'ils sont unitaires, on en déduit que $\pi_{C(P)} = P$. □

Proposition 4.64. *$u \in \mathcal{L}(E)$ est cyclique si et seulement s'il existe un polynôme $P \in K[X]$ et une base $\mathcal{B}$ de E tels que $\mathrm{Mat}_{\mathcal{B}}(u) = C(P)$.*

Démonstration: C'est évident. □

Le lemme suivant est utile pour la suite :

Lemme 4.65. *Soit $u \in \mathcal{L}(E)$. Pour tout $x \in E$, on note $\pi_{u,x} \in K[X]$ un générateur de l'idéal $\{P \in K[X] : P(u)(x) = 0\}$.*

Alors il existe $x \in E$ tel que $\pi_{u,x} = \pi_u$.

Démonstration: Raisonner avec chaque facteur irréductible de π_u (avec multiplicité), puis faire la somme des vecteurs obtenus. □

Théorème 4.66. *Soit $u \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes :*

1. *u est cyclique.*
2. *$\pi_u = \pm\chi_u$.*
3. *L'ensemble des endomorphismes de E commutant avec u est $K[u]$.*

Si le corps K est infini, ces conditions équivalent à :

4. *E a un nombre fini de sous-espaces stables par u.*

Démonstration:

— (1) $\implies$ (2) : cela résulte des propositions 4.63 et 4.64.

— (2) $\implies$ (3) : il est clair que tout élément de $K[u]$ commute à u. Réciproquement, le lemme 4.65 assure l'existence de $x \in E$ tel que $\pi_{u,x} = \pi_u$. Soit alors $v \in \mathcal{L}(E)$ commutant avec u. Alors $\{P(u)(x) : P \in K[X]\}$ est isomorphe à $K[X]/(\pi_{u,x})$, donc est de dimension n par hypothèse. Donc tout élément de E s'écrit $P(u)(x)$ pour un certain $P \in K[X]$. En particulier, $v(x) = P(u)(x)$, pour un $P \in K[X]$. Alors pour tout $y \in E$, il existe $Q \in K[X]$ tel que $y = Q(u)(x)$, et donc

$$v(y) = (v \circ Q(u))(x) = (Q(u) \circ v)(x) = (Q(u) \circ P(u))(x) = P(u)(Q(u)(x)) = P(u)(y)\,.$$

Donc on a bien $v = P(u)$, d'où l'implication.

— (3) $\implies$ (1) : Par le lemme 4.65, il existe $x \in E$ tel que $\pi_{u,x} = \pi_u$. Notons $E_x := \mathrm{Vect}(u^k(x), k \in \mathbf{N})$. Alors E_x est stable par u, et il admet un supplémentaire stable F_x, comme dans la preuve du théorème 4.54. Considérons alors la projection $p : E \to E$ sur F_x parallèlement à E_x. Alors p commute à u, donc par hypothèse, il existe $Q \in K[X]$ tel que $p = Q(u)$. En particulier, on en déduit que $Q(u_{|E_x}) = 0$, donc $Q(u)(x) = 0$, donc $\pi_{u,x}$ divise Q, donc π_u divise Q, donc $Q(u) = 0$, donc $p = 0$, donc $F_x = 0$, donc $E = E_x$ et u est cyclique.

□

Corollaire 4.67. *Soit $K = \mathbf{F}_q$ un corps fini à q éléments.*

Alors le nombre de matrices nilpotentes d'indice n dans $\mathrm{M}_n(\mathbf{F}_q)$ est exactement

$$(q^n - 1) \cdot (q^n - q) \dots (q^n - q^{n-2})\,.$$

Démonstration: Il suffit d'écrire l'équation aux classes pour l'action transitive de $\mathrm{GL}_n(\mathbf{F}_q)$ sur l'ensemble de ces matrices, et d'utiliser la propriété du commutateur des endomorphismes cycliques.

? ? □

4.7.2 Réduction de Frobenius

Définition 4.68. Soit $u \in \mathcal{L}(E)$. Un sous-espace cyclique de E (relativement à u) est un sous-espace de la forme $E_x := \{P(u)(x) : P \in K[X]\}$ pour un certain $x \in E$.

La proposition suivante a déjà été utilisée dans un cas particulier dans la preuve de la réduction de Jordan. Elle sera utile pour la preuve du théorème de Frobenius :

Proposition 4.69. *Soit $u \in \mathcal{L}(E)$ et $x \in E$ tel que $\pi_{u,x} = \pi_u$.*

Alors le sous-espace cyclique $E_x := \mathrm{Vect}(u^i(x) : i \in \mathbf{N})$ admet un supplémentaire stable.

Démonstration: Notons $d := \deg(\pi_u)$.

Pour construire un supplémentaire stable F_x, on utilise la dualité : puisque $(x, u(x), \ldots, u^{d-1}(x))$ est une base de E_x, il existe une forme linéaire $\varphi \in E^*$ telle que $\varphi(u^{d-1}(x)) = 1$ et $\varphi(u^i(x)) = 0$ pour $0 \leq i < d-1$. Puisque $\pi_{u,x} = \pi_u$, on voit que ${}^tu^d \in \mathrm{Vect}({}^tu^i : 0 \leq i \leq d-1)$. En particulier, le sous-espace vectoriel $G_\varphi := \mathrm{Vect}(\varphi, {}^tu(\varphi), \ldots, {}^tu^{m-1}(\varphi))$ de E^* est stable par ${}^tu$, et de dimension d (calculer ${}^tu^i(\varphi)$ sur la famille $(x, u(x), \ldots, u^{d-1}(x))$). On en déduit donc que son orthogonal $F_x := G_\varphi{}^\circ$ est un sous-espace vectoriel de E, stable de u, et de dimension $n - d$. Montrons que F_x est un supplémentaire de E_x : il suffit de montrer que $E_x \cap F_x = \{0\}$. Soit $y \in E_x \cap F_x$. Alors $y = \sum_{i=0}^{d-1} a_i \cdot u^i(x)$ et pour tout $0 \leq k \leq d-1$, $(\varphi(u^k(x)) = 0$, donc $\varphi(y) = 0$ implique $a_{d-1} = 0$. On a ensuite $u(y) = \sum_{i=0}^{d-2} a_i \cdot u^{i+1}(x)$, donc en appliquant φ et en utilisant le fait que ${}^tu(\varphi)(y) = 0$, on obtient $a_{d-2} = 0$. Par récurrence, on montre ainsi que pour tout i, $a_i = 0$, donc $y = 0$ et $E_x \cap F_x = \{0\}$. Donc on a montré que F_x est un supplémentaire stable de E_x. □

Le théorème de Frobenius affirme notamment que tout endomorphisme admet une décomposition en somme directe de sous-espaces cycliques.

Théorème 4.70. *Soit $u \in \mathcal{L}(E)$.*

Alors il existe une base $\mathcal{B}$ de E et des polynômes unitaires $P_1, \ldots, P_r \in K[X]$ tels que $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{diag}(C(P_1), \ldots, C(P_r))$ et $P_1 \mid P_r \mid \ldots \mid P_r$. En outre, les P_i sont uniques.

Remarque 4.71. De façon équivalente, il existe une décomposition en somme directe $E = E_1 \oplus \cdots \oplus E_r$, avec E_i stable par u, $u_{|E_i}$ cyclique de polynôme minimal P_i, avec $P_1 | \ldots | P_r$. Les polynômes P_i sont alors uniques.

Démonstration:

— Pour l'existence, on raisonne par récurrence sur la dimension n de E. Si $n = 1$, c'est évident. On suppose $n > 1$, et on utilise la proposition 4.69 : il existe $x \in E$ tel que $\pi_{u,x} = \pi_u$. Si $d = \deg(\pi_u)$, la famille $\mathcal{B}' := (x, u(x), \ldots, u^{d-1}(x))$ est une base de $E_x = \mathrm{Vect}(u^k(x), k \in \mathbf{N})$, et la restriction de u à E_x a pour matrice dans la base précédente la matrice $C(P_r)$, avec $P_r := \pi_{u,x}$. En outre, le lemme **??** assure que E_x admet un supplémentaire stable F_x. On applique

l'hypothèse de récurrence à la restriction de u à F_x : il existe des polynômes $P_1, \dots, P_{r-1}$ tels que $P_1 | \dots | P_{r-1}$ et une base $\mathcal{B}''$ de F_x telle que $\mathrm{Mat}_{\mathcal{B}''}(u) = \mathrm{diag}(C(P_1), \dots, C(P_{r-1}))$. En posant $\mathcal{B} = \mathcal{B}'' \cup \mathcal{B}'$, on obtient que $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{diag}(C(P_1), \dots, C(P_r))$. Il reste à montrer que P_{r-1} divise P_r. Pour cela, on remarque que $P_r = \pi_u$, donc $P_r(u) = 0$, donc $P_r(C(P_{r-1})) = 0$, donc par la proposition 4.63, P_{r-1} divise P_r, ce qui conclut la preuve de l'existence.

— Pour l'unicité : on a construit une suite de sous-espaces E_i cycliques pour u, tels que $E = E_1 \oplus \dots \oplus E_r$, et pour tout i, P_i est le polynôme minimal de $u_{|E_i}$. Soit $E = F_1 \oplus \dots \oplus F_s$ une autre décomposition, avec $u_{|F_j}$ cyclique de polynôme Q_j pour tout j, et $Q_1 | \dots | Q_s$. D'abord, $P_r = Q_s$ est le polynôme minimal de u. Montrons que les familles (Q_j) et (P_i) coïncident. Soit $k \geq 1$ tel que pour tout $0 \leq i < k$, on a $P_{r-i} \neq Q_{s-i}$. Calculons $P_{r-k}(u)$: on a $P_{r-k}(u)(E) = P_{r-k}(u)(E_{r-k+1}) \oplus \dots P_{r-k}(u)(E_r)$ et $P_{r-k}(u)(E) = P_{r-k}(u)(F_1) \oplus \dots P_{r-k}(u)(F_{r-k}) \oplus P_{r-k}(u)(F_{r-k+1}) \oplus \dots P_{r-k}(u)(F_s)$. Or par hypothèse, pour tout $0 \leq i < k$, $\dim(P_{r-k}(u)(E_{r-i})) = \dim(P_{r-k}(u)(F_{s-i}))$, donc on en déduit que $P_{r-k}(u)(F_{s-k}) = 0$, donc Q_{s-k} divise P_{r-k}. Par symétrique, $Q_{s-k} = P_{r-k}$. On montre ainsi par récurrence que $r = s$ et pour tout i, $P_i = Q_i$, d'où l'unicité.

□

Remarque 4.72. Ce théorème décrit donc en toute généralité les classes de similitude dans $\mathcal{L}(E)$ et les met en correspondance avec les familles de polynômes $P_1, \dots, P_r \in K[X]$ tels que $P_1 \,|\, P_2 \,|\, \dots \,|\, P_r$ et $\sum_i (P_i) = n$.

Définition 4.73. Soit $u \in \mathcal{L}(E)$. Les polynômes P_i associés à u dans le théorème 4.70 sont appelés les invariants de similitude de u. Ce sont des invariants totaux, au sens où deux endomorphismes de E sont semblables si et seulement s'ils ont les mêmes invariants de similitude.

Exemple 4.74. Soit K un corps (pas forcément infini), et L une extension de K. Soient $A, B \in \mathrm{M}_n(K)$.

Alors A et B sont semblables dans $\mathrm{M}_n(L)$ si et seulement si elles sont semblables dans $\mathrm{M}_n(K)$.

Exemple 4.75. On peut calculer le nombre de classes de similitudes dans $\mathrm{M}_n(\mathbf{F}_q)$, pour de petites valeurs de n.

Proposition 4.76. *Soit $u \in \mathcal{L}(E)$ d'invariants de similitude $P_1, \dots, P_r$.*

1. $P_r = \pi_u$.
2. $P_1 \dots P_r = \chi_u$.

Corollaire 4.77. *Si $n \leq 3$, la donnée des polynômes minimaux et caractéristiques forme un invariant total : deux endomorphismes sont semblables si et seulement s'ils ont même polynôme minimal et même polynome caractéristique.*

Exemple 4.78. Sur tout corps K, il existe des matrices non semblables dans $\mathrm{M}_4(K)$ ayant même polynôme minimal, même polynôme caractéristique. Par exemple, les matrices

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

On a bien $\chi_A = \chi_B = X^4$, $\pi_A = \pi_B = X^2$, mais $\mathrm{rg}(A) = 1$ et $\mathrm{rg}(B) = 2$, donc elles ne sont pas semblables. Les invariants de similitudes de A sont (X, X, X^2) et ceux de B sont (X^2, X^2).

Quelques applications : si K est un corps fini, et L/K une extension de corps, alors deux matrices de $\mathrm{M}_n(K)$ sembables sur L sont semblables sur K. Une matrice est semblable à sa transposée.

5 Matrices à coefficients dans un anneau commutatif

Il s'agit notamment d'essayer d'étendre certains résultats valables sur un corps aux matrices à coefficients dans certains anneaux commutatifs.

On rappelle que si A est un anneau commutatif (unitaire), $\mathrm{M}_{n,p}(A)$ est bien défini et muni de l'addition naturelle des matrices, et que $\mathrm{M}_n(A)$ est naturellement une A-algèbre

5.1 Déterminant

On peut définir le déterminant d'une matrice de $\mathrm{M}_n(A)$ de plusieurs façons équivalentes :

1. si A est intègre, on dispose du corps des fractions K et A, et l'inclusion $A \subset K$ induit une inclusion $\mathrm{M}_n(A) \subset \mathrm{M}_n(K)$. On peut donc voir une matrice de $\mathrm{M}_n(A)$ comme une matrice de $\mathrm{M}_n(K)$ et définir son déterminant comme plus haut. On remarque ensuite que les formules (somme sur les permutations, ou développement par rapport à une ligne ou une colonne) assurent que le déterminant est un polynôme à coefficients entiers en les coefficients de la matrice, donc si $M \in \mathrm{M}_n(A)$, alors $\det(M) \in A$.
2. dans le cas général, on peut définir le déterminant via la formule de la proposition 3.90. Par définition on voit alors que $\det(M) \in A$ pour tout $M \in \mathrm{M}_n(A)$.
3. dans le cas général, on peut aussi définir le déterminant, par récurrence sur n, via la proposition 3.94. Là encore, on a $\det(M) \in A$ pour tout $M \in \mathrm{M}_n(A)$.

Avec ces définitions, dont on peut vérifier qu'elles coïncident, il faut vérifier un certain nombre de propriétés du déterminant. Le caractère multilinéaire alterné est facile, tout comme la formule de la comatrice. Mais par exemple, dans le cas non intègre, la formule $\det(M \cdot N) = \det(M) \cdot \det(N)$ n'est pas complétement évidente : on peut la démontrer à la main en utilisant uniquement le caractère multilinéaire alterné du déterminant.

Une fois les définitions précisées, on peut alors démontrer par exemple le résultat suivant :

Théorème 5.1. *Soit A un anneau commutatif et $M \in \mathrm{M}_n(A)$. Notons $\varphi_M : A^n \to A^n$ l'application induite par A sur les vecteurs colonnes.*

1. *φ_M est injective si et seulement si $\det(M)$ n'est pas diviseur de 0 dans A.*
2. *φ_M est surjective si et seulement si φ_M est bijective si et seulement si $\det(M) \in A^\times$.*

Démonstration:

1. Supposons d'abord que $\det(M)$ n'est pas diviseur de 0. Soit $x \in A^n$ tel que $\varphi_M(x) = 0$. Écrivons la formule de la comatrice : ${}^t\mathrm{Com}M \cdot M = \det(M) \cdot I_n$. On applique cette égalité à x : $0 = {}^t\mathrm{Com}M \cdot M \cdot x = \det(M) \cdot x$. Donc $\det(M) \cdot x = 0$, donc $x = 0$ car $\det(M)$ n'est pas diviseur de 0 dans A.

   Supposons maintenant que $\det(M)$ est diviseur de 0. Alors il existe $a \in A \setminus \{0\}$ tel que $a \cdot \det(M) = 0$.
   - Si a annule les déterminants de toutes les matrices carrées extraites de M, alors en particulier $a \cdot m_{i,j} = 0$ pour tout i, j, ce qui assure que le vecteur colonne X_a, dont toutes les coordonnées sont égales à a, est un vecteur non nul du noyau de φ_M.
   - Sinon, il existe une matrice extraite N de M, de taille $d < n$ maximale, telle que $a \cdot \det(N) \neq 0$. Quitte à permuter lignes et colonnes de M (ce qui ne change pas le caractère injectif ou non de M), on peut supposer que N est la matrice obtenue à partir de M en extrayant les d premières lignes et les d premières colonnes. Notons alors N_i la matrice de taille $d+1$ formée à partir de N en ajoutant à droite la $d+1$-ième colonne de M (tronquée aux d premières coordonnées) et en bas la i-ième ligne de M (tronquée aux $d+1$ premières coordonnées). Comme N_i a deux lignes identiques (cas $i \leq d$) ou est une matrice extraite de taille $d+1$ dans M (cas $i > d$), on a pour tout i, $a \cdot \det(N_i) = 0$. On calcule aors le déterminant de N_i en développant par rapport à la dernière ligne de N_i : on obtient $0 = a \cdot \sum_{k=1}^{d+1} (-1)^k m_{i,k} \cdot \Delta_k$. Donc le vecteur colonne $a \cdot (\Delta_1, \ldots, (-1)^d \Delta_{d+1}, 0, \ldots, 0)$ est dans le noyau de φ_M. Il reste à vérifier que ce vecteur est non nul : par construction, $a \cdot \Delta_{d+1} = a \cdot \det(N) \neq 0$, donc ce vecteur est non nul, donc φ_M n'est pas injective.

2. Supposons φ_M surjective. Notant (e_i) la base canonique de A^n, pour tout i, il existe $x_i \in A^n$ tel que $\varphi_M(x_i) = e_i$. Alors l'application A-linéaire $\psi : A^n \to A^n$ définie par $\psi(e_i) = x_i$ vérifie $\varphi \circ \psi = \mathrm{id}$, donc matriciellement, si N désigne la matrice associée à ψ, on a $M \cdot N = I_n$, donc $\det(M) \cdot \det(N) = 1$. Donc $\det(M) \in A^\times$.

   Supposons maintenant que $\det(M) \in A^\times$. Alors la formule de la comatrice assure que M est inversible, d'inverse $(\det(M))^{-1} \cdot {}^t\mathrm{Com}M$, ce qui assure également que φ_M est bijective.

□

5.2 Pivot de Gauss et Forme normale dans le cas anneau euclidien

On s'intéresse ici aux classes d'équivalence de matrices à coefficients dans un anneau euclidien, et on cherche à adapter le pivot de Gauss dans ce contexte. La difficulté réside dans le fait que dans un anneau qui n'est pas un corps, on ne peut pas faire n'importe quelle division. Dans ce contexte, on appelle matrice élémentaire une matrice de transvection ou une matrice de permutation.

Néanmoins, on dispose de la généralisation suivante du théorème **??** :

Théorème 5.2. *Soit A un anneau euclidien, et $M \in \mathrm{M}_{n,p}(A)$.*

Alors il existe des matrices inversibles $P \in \mathrm{GL}_n(A)$, $Q \in \mathrm{GL}_p(A)$, produits de matrices élémentaires, et des éléments $d_1, \ldots, d_r \in A$ tels que

$$P \cdot M \cdot Q^{-1} = \mathrm{diag}(d_1, \ldots, d_r, 0, \ldots, 0)\,,$$

avec $d_1 | \ldots | d_r$. En outre, les d_i sont uniques, à multiplication près par un élément inversible de A.

Les d_i sont appelés les facteurs invariants de la matrice M. Ce théorème décrit en particulier exactement les classes d'équivalence de matrices à coefficients dans A. En outre, la preuve fournit un algorithme pour déterminer les d_i, ainsi que les matrices P et Q.

Remarque 5.3. La description précédente des classes d'équivalence de matrices reste valable quand A est seulement supposé principal. En revanche, dans ce cadre plus général, on a besoin de matrices inversibles qui ne sont pas des matrices élémentaires pour obtenir la forme diagonale souhaitée. En outre, la preuve du théorème dans le cas euclidien est constructive et fournit un algorithme calculant les matrices P, Q et les coefficients d_i, ce qui n'est pas le cas dans le contexte plus général des anneaux principaux.

Démonstration: On note v le stathme euclidien, et on définit $v(M) := \min\{v(m_{i,j}); m_{i,j} \neq 0\}$ si $M \neq 0$.

On s'intéresse d'abord à l'existence de la décomposition.

— Si $M = 0$, le résultat est évident.

— On suppose désormais $M \neq 0$. Il existe (i, j) tel que $v(m_{i,j}) = v(M)$. En faisant des permutations des lignes et des colonnes, on peut supposer $(i, j) = (1, 1)$.

1. Étape 1 : pour tout $2 \leq i \leq n$, on effectue la division euclidienne de $m_{i,1}$ par $m_{1,1}$: $m_{i,1} = m_{1,1} \cdot q_i + r_i$, avec $r_i = 0$ ou $v(r_i) < v(m_{1,1})$. On fait alors l'opération élémentaire $L_i \leftarrow L_i - q_i \cdot L_1$, ce qui remplace le coefficient $m_{i,1}$ par r_i. On a alors deux cas : soit tous les coefficients r_i sont nuls, et la matrice obtenue a une première colonne nulle (hormis $m_{1,1}$) et on passe à l'étape 2 ; soit il existe un coefficient $r_i \neq 0$ et on fait l'échange entre les lignes L_1 et L_i : dans ce second cas, la quantité $v(m_{1,1})$ a strictement diminué, et on recommence l'étape 1. Puisque la suite des $v(m_{1,1})$ est une suite strictement décroissante d'entiers positifs, au terme d'un nombre fini de répétitions de l'étape 1, on est le premier cas et on passe à l'étape 2.

2. Étape 2 : on est arrivé à une matrice de la forme

$$\begin{pmatrix} m_{1,1} & * \\ 0 & M' \end{pmatrix}.$$

Si $m_{1,1}$ divise tous les coefficients de la première ligne, alors en effectuant des opérations élémentaires sur les colonnes, on se ramène à une matrice de la forme

$$\begin{pmatrix} m_{1,1} & 0 \\ 0 & M' \end{pmatrix}.$$

Sinon, on effectue des divisions euclidiennes comme à l'étape 1, mais sur les colonnes, pour se ramener à une matrice M' telle que la valeur $v(m'_{1,1})$ est

strictement inférieure à $v(m_{1,1})$. Puis on retourne à l'étape 1. De nouveau, après un nombre fini de retour à l'étape 1, puis de passages par l'étape 2, on arrive dans le premier cas de l'étape 2, donc à une matrice

$$\begin{pmatrix} m_{1,1} & 0 \\ 0 & M' \end{pmatrix}.$$

3. Étape 3 : si $m_{1,1}$ divise tous les coefficients de M', on conclut facilement par récurrence sur n, en retournant au début avec la matrice M'. Sinon, il existe un coefficient $m_{i,j}$, avec $i, j \geq 2$, qui n'est pas divisible par $m_{1,1}$. Alors on effectue la division euclidienne de $m_{i,j}$ par $m_{1,1}$, et avec des opérations élémentaires on obtient une matrice N avec $v(n_{1,1})$ strictement inférieur à $v(m_{1,1})$. On retourne alors à l'étape 1, avec N. Au bout d'un nombre fini d'étapes, on arrive au premier cas de l'étape 3.

Montrons maintenant l'unicité des coefficients d_i. Pour cela, on peut définir pour tout $1 \leq i \leq \min(n, p)$, le nombre $d_i(M)$ comme étant un PGCD des mineurs (déterminants des matrices carrées extraites de M) de taille i. Un calcul simple - via la multilinéarité du déterminant - assure que la multiplication de M, à droite ou à gauche, par une matrice élémentaire, ne modifie pas $d_i(M)$ (à un inversible près). Or un calcul immédiat assure que, une fois la matrice écrite sous la forme diagonale du théorème, on a $d_i(M) = d_1 \dots d_i$ pour tout i, ce qui assure l'unicité modulo les inversibles.

Exemple 5.4. Si $A = \mathbf{Z}$ et $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors on obtient $(d_1, d_2, d_3) = (1, 3, 0)$.

Corollaire 5.5. *Si A est euclidien, le groupe $\mathrm{SL}_n(A)$ est engendré par les matrices élémentaires (transvections et permutations).*

Remarque 5.6. Ce résultat est faux en général dans un anneau principal. Par exemple, il est faux dans $\mathrm{SL}_2(\mathbf{Z}[\frac{1+i\sqrt{19}}{2}])$.

5.3 Lien avec la classification des groupes abéliens finis et la réduction des endomorphismes

Les deux exemples cruciaux d'anneaux euclidiens (qui ne sont pas des corps) sont $\mathbf{Z}$ et $K[X]$. Nous allons maintenant voir que dans ces deux cas, le théorème décrivant les classes d'équivalence de matrices à des conséquences importantes.

5.3.1 Matrices à coefficients dans Z et groupes abéliens de type fini

Lemme 5.7. *Soit $G = \mathbf{Z}^n$ et H un sous-groupe de G. Alors H est de type fini, isomorphe à $\mathbf{Z}^k$ pour un certain $0 \leq k \leq n$.*

Démonstration: On note $\pi_n : G \to \mathbf{Z}$ la projection sur la n-ième coordonnée. Alors $\pi_n(H)$ est un sous-groupe de $\mathbf{Z}$, donc de la forme $d_n \cdot \mathbf{Z}$. Par définition, il existe $h_n \in H$ tel que $\pi_n(h_n) = d_n$. Notons $H' := \ker(\pi_n) \cap H$. Alors H' est naturellement un sous-groupe de $\ker(\pi_n) = \mathbf{Z}^{n-1}$, et considérons le morphisme $\varphi : H' \times \langle h_n \rangle \to H$ donné par la somme dans G. Comme $H' \cap \langle h_n \rangle = \{0\}$, le morphisme φ est injectif. Soit $h \in H$; alors il existe $m \in \mathbf{Z}$ tel que $\pi_n(h) =$

$md_n = m\pi_n(h_n)$, donc $h - mh_n \in H'$, donc $h = \varphi(h - mh_n, mh_n)$, donc φ est surjective. Donc φ est un isomorphisme entre H et $H' \times \langle h_n \rangle$, avec H' sous-groupe de $\mathbf{Z}^{n-1}$. On conclut par récurrence sur n. □

Théorème 5.8. *Soit G un groupe abélien fini. Alors il existe $n \geq 0$ et $d_1 | \dots | d_r$ des entiers positifs tels que $G \xrightarrow{\sim} \mathbf{Z}^n \times \prod_{i=1}^r (\mathbf{Z}/d_i\mathbf{Z})$.*
Remarque 5.9. On a également un résultat d'unicité des entiers n et d_i, on renvoie au cours de Pierre Charollois pour la preuve de l'unicité.

Démonstration: Par définition, G admet une partie génératrice finie $\{g_1, \dots, g_n\}$. On considère l'unique morphisme de groupes $\pi : \mathbf{Z}^n \to G$ défini par $\pi(e_i) := g_i$, où $(e_1, \dots, e_n)$ désigne la "base canonique" de $\mathbf{Z}^n$. Notons $H := \ker(\pi)$. Le lemme 5.7 assure que $H \xrightarrow{\sim} \mathbf{Z}^p$, avec $0 \leq p \leq n$. Notons $(f_1, \dots, f_p)$ l'image dans H de la base canonique de $\mathbf{Z}^p$. Alors il existe une matrice $M = (m_{i,j}) \in \mathrm{M}_{n,p}(\mathbf{Z})$ définie par $f_j = \sum_{k=1}^n m_{k,j} \cdot e_k$. On applique le théorème 5.2 à la matrice M : il existe des matrices P, Q inversibles telles que $P \cdot M \cdot Q^{-1} = \mathrm{diag}(d_1, \dots, d_r, 0, \dots, 0)$. Or par construction, G est isomorphe à

$$G \cong \mathbf{Z}^n/H = \mathbf{Z}^n/M \cdot \mathbf{Z}^p \cong \mathbf{Z}^n/M \cdot Q^{-1} \cdot \mathbf{Z}^p \,,$$

et la matrice P induit un isomorphisme de groupes $P : \mathbf{Z}^n/(PMQ^{-1}) \cdot \mathbf{Z}^p \xrightarrow{\sim} G$. Or $\mathbf{Z}^n/(PMQ^{-1}) \cdot \mathbf{Z}^p \cong \mathbf{Z}^{n-p} \times \prod_{i=1}^r (\mathbf{Z}/d_i\mathbf{Z})$.
Cela conclut la preuve. □

5.3.2 Matrices à coefficients dans $K[X]$ et réduction des endomorphismes

Dans cette section, nous essayons de retrouver le théorème de décomposition de Frobenius (invariants de similitude) à partir du théorème 5.2 dans l'anneau $K[X]$, sans utiliser explicitement la notion de module sur un anneau.
L'idée est de copier la stratégie de la preuve du théorème 5.8 à partir du théorème 5.2, en remplaçant $\mathbf{Z}$ par $K[X]$ et un groupe abélien par un espace vectoriel muni d'un endomorphisme...
Lemme 5.10. *Soit $H \subset G := K[X]^n$ un sous-K-espace vectoriel, stable par multiplication (diagonale) par X.*
Alors il existe un entier $0 \leq k \leq n$ et un isomorphisme de K-espaces vectoriels $H \cong K[X]^k$, équivariant pour l'action de X par multiplication.

Démonstration: On note $\pi_n : G \to K[X]$ la projection sur la n-ième coordonnée (c'est une application K-linéaire, compatible au produit par X). Alors $\pi_n(H)$ est un idéal de $K[X]$ (le vérifier), donc de la forme (P_n). Par définition, il existe $h_n \in H$ tel que $\pi_n(h_n) = P_n$. Notons $H' := \ker(\pi_n) \cap H$. Alors H' est naturellement un sous-K-espace vectoriel de $\ker(\pi_n) = K[X]^{n-1}$, et considérons le morphisme $\varphi : H' \oplus (h_n) \to H$ donné par la somme dans G. Comme $H' \cap (h_n) = \{0\}$, le morphisme φ est injectif. Soit $h \in H$; alors il existe $Q \in K[X]$ tel que $\pi_n(h) = P_n \cdot Q = Q \cdot \pi_n(h_n)$, donc $h - Q \cdot h_n \in H'$ (où le produit désigne l'action diagonale de $K[X]$ sur G), donc $h = \varphi(h - Q \cdot h_n, Q \cdot h_n)$, donc φ est surjective. Donc φ est un isomorphisme de K-espaces vectoriels entre H et $H' \oplus (h_n)$, qui est compatible à l'action de X, avec H' sous K-espace vectoriel

de $K[X]^{n-1}$ stable par X. On conclut par récurrence sur n, en remarquant que $(h_n) \cong K[X]$ comme K-espace vectoriel, de façon compatible avec l'action de X. □

Théorème 5.11. *Soit E un K-espace vectoriel et $u \in \mathcal{L}(E)$. On suppose qu'il existe une partie fini $F \subset E$ telle que la réunion des orbites sous u des éléments de F engendre E.*
Alors il existe un entier n et des polynômes unitaires $P_1 | \dots | P_r$, et un isomorphisme de K-espaces vectoriels

$$\varphi : E \xrightarrow{\sim} K[X]^n \oplus \bigoplus_{i=1}^{r} K[X]/(P_i)\,,$$

qui est équivariant pour l'action de u à gauche et celle de X à droite, i.e. $\varphi(u(v)) = X \cdot \varphi(v)$ pour tout $v \in E$.

Démonstration: Notons $\{v_1, \dots, v_n\}$ une partie finie vérifiant les hypothèses de l'énoncé. On considère alors l'application K-linéaire surjective : $\pi : K[X]^n \to E$ définie par $\pi((P_i)) := \sum_i P_i(u)(v_i)$. Notons alors $H := \ker(\pi)$, qui est un sous-K-espace vectoriel de $K[X]^n$. En outre, ce sous-espace vectoriel est clairement stable par multiplication par X. Alors le lemme 5.10 assure l'existence d'un isomorphisme $H \cong K[X]^p$, qui est K-linéaire et équivariant pour l'action de X. En notant $(f_1, \dots, f_p)$ les éléments de H correspondant à la base canonique de $K[X]^p$, on construit alors la matrice $M = (m_{i,j}) \in \mathrm{M}_{n,p}(K[X])$, via la formule $f_j = (m_{i,j})_{1 \le i \le n}$.
On applique alors le théorème 5.2 à la matrice M : il existe des matrices P, Q inversibles (à coefficients dans $K[X]$) telles que $P \cdot M \cdot Q^{-1} = \mathrm{diag}(d_1, \dots, d_r, 0, \dots, 0)$.
Or par construction, E est isomorphe (comme K-espace vectoriel avec action de X) à

$$E \cong (K[X])^n / H = (K[X])^n / M \cdot K[X]^p \cong (K[X])^n / M \cdot Q^{-1} \cdot (K[X])^p\,,$$

et la matrice P induit un isomorphisme K-linéaire et X-équivariant $P : (K[X])^n/(PMQ^{-1}) \cdot (K[X])^p \xrightarrow{\sim} E$. Or $(K[X])^n/(PMQ^{-1}) \cdot (K[X])^p \cong (K[X])^{n-p} \oplus \bigoplus_{i=1}^{r} K[X]/(P_i)$.
Cela conclut la preuve. □

On en déduit le théorème de Frobenius :
Corollaire 5.12. *Soit E un K-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Alors il existe une base $\mathcal{B}$ de E et des polynômes unitaires $P_1, \dots, P_r \in K[X]$ tels que $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{diag}(C(P_1), \dots, C(P_r))$ et $P_1 \,|\, P_r \,|\, \dots \,|\, P_r$.*

Démonstration: On applique le théorème précédent (l'hypothèse est clairement vérifiée car E est de dimension finie) : on dispose d'un isomorphisme de K-espaces vectoriels

$$\varphi : E \xrightarrow{\sim} K[X]^n \oplus \bigoplus_{i=1}^{r} K[X]/(P_i)\,,$$

vérifiant $\varphi(u(v)) = X \cdot \varphi(v)$. Puisque E est de dimension finie sur K, nécessairement $n = 0$. L'endomorphisme u de E s'identifie donc à l'endomorphisme de multiplication (diagonale) par X sur $\bigoplus_{i=1}^{r} K[X]/(P_i)$. Or pour tout $1 \le i \le r$, la

restriction de cet endomorphisme au sous-K-espace stable $K[X]/(P_i)$ est clairement cyclique (engendré par le vecteur 1), donc cette restriction a pour matrice la matrice compagnon dans la base $(1, X, \ldots, X^{\deg(P_i)-1})$. Cela assure la preuve du corollaire. □

Remarque 5.13. Soit $A \in \mathrm{M}_n(K)$. On déduit facilement (via des opérations sur les lignes et les colonnes) de la preuve du théorème que les invariants de similitude de A sont exactement les facteurs invariants distincts de 1 de la matrice $A - XI_n \in M_n(K[X])$. Cela fournit un algorithme pour calculer les invariants de similitude d'une matrice.
En outre, on déduit de cette dernière remarque, qu'étant données deux matrices $A, B \in \mathrm{M}_n(K)$, alors A et B sont semblables dans $\mathrm{M}_n(K)$ si et seulement si $A - XI_n$ et $B - XI_n$ sont équivalentes dans $\mathrm{M}_n(K[X])$. On peut démontrer cette dernière propriété directement (via des divisions euclidiennes dans $K[X]$).
Exemple 5.14. Déterminer et dénombrer les classes de similitude de matrices de $\mathrm{M}_5(K)$ de polynôme caractéristique $(X^2 - 1)(X - 1)^3$. Puis traiter le cas des matrices de $M_7(K)$ de polynôme minimal $(X^2 - 2)(X + 1)^2$. Puis le cas des matrices de $\mathrm{M}_8(K)$ de polynôme caractéristique $(X^2 + 1)^2(X^2 - 3)^2$.