Chapitre 1

Réduction et invariants de similitudes

Ces notes ont été rédigées par Mélanie GUENAIS (2020-2022), et reprises par Damien THOMINE (2022-2023).

Ce texte contient de nombreux passages en gris. Ceux-ci pourront-être passé en première lecture. Ces passages contiennent cependant de nombreuses démonstrations ; s'il n'est pas important (ni souhaitable) de les apprendre, cela peut être un exercice instructif d'essayer de les faire soi-même.

I Introduction

Il s'agit d'étudier dans ce chapitre l'ensemble des matrices sur $\mathbb{R}$ ou $\mathbb{C}$ en les considérant comme des familles de représentants d'applications linéaires. Cette étude conduit à identifier des classes d'équivalence pour les matrices qui sont les matrices équivalentes de $\mathscr{M}_{p,n}$ et les matrices semblables dans $\mathscr{M}_n(\mathbb{R})$. Celles-ci permettent de comprendre alors comment on peut choisir les matrices dans ces classes les "plus simples" possibles : on retrouve alors dans un premier temps le théorème du rang, qui définit un invariant complet pour la classe des matrices équivalentes. Dans une seconde partie, on étudie les invariants de similitudes qui permettent de comprendre le principe de la réduction et de la décomposition de Dunford.

I.1 Cadre de travail

- ▷ Dans tout le chapitre, E et F désignent des espaces vectoriels sur $\mathbb{R}$ ou $\mathbb{C}$.
- ▷ E et F sont de dimension finie sur $\mathbb{R}$ ou $\mathbb{C}$. On note $n = \dim(E)$ et $p = \dim(F)$.
- ▷ On note $\mathscr{M}_{p,n}(\mathbb{R})$ l'ensemble des matrices à p lignes et n colonnes à coefficients réels ($\mathscr{M}_{p,n}(\mathbb{C})$ si les coefficients sont complexes).
- ▷ $\mathscr{M}_n(\mathbb{R})$ désigne l'ensemble des matrices carrées de taille n à coefficients réels.

I.2 Motivation générale

Partons d'un problème concret de recherche de données numériques qui peut se traduire sous une forme matricielle : il s'agit de résoudre un système linéaire de la forme

$$AX = B \quad \text{où} \quad X \in \mathscr{M}_{n,1}(\mathbb{R}), \quad B \in \mathscr{M}_{p,1}(\mathbb{R}) \quad \text{et} \quad A \in \mathscr{M}_{p,n}(\mathbb{R}).$$

Pour étudier les solutions de ce système, on peut utiliser l'algorithme du pivot de Gauss. Les questions qui se posent sont d'une part de comprendre pourquoi cet algorithme "marche", et d'autre part d'essayer de trouver d'autres algorithmes plus intéressants (à définir). Pour cela, il est utile d'associer à ce système une représentation algébrique, c'est à dire de considérer X et B comme des représentants de vecteurs dans des espaces vectoriels E et F de dimension égale au nombre de colonnes, et d'interpréter AX comme une fonction du vecteur représenté par X.

II Liens matrices colonnes et vecteurs : coordonnées dans une base

II.1 Matrices de coordonnées

Étant donné $X \in \mathscr{M}_{n,1}(\mathbb{R})$ et un espace vectoriel E, on souhaite associer à X un vecteur de E. Pour cela il y a besoin de choisir une base de E.

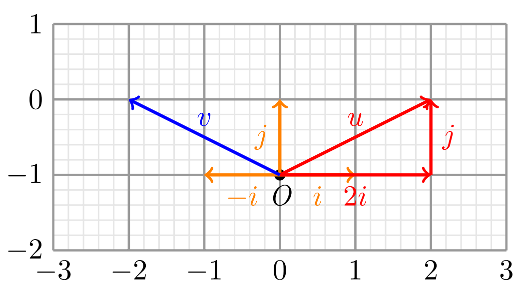
Définition II.1 (Rappel : Coordonnées)**.**
Soit $\mathscr{B} = (e_1, \cdots, e_n)$ une base de E. Alors, pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$, il existe un unique vecteur $u \in E$ dont les coordonnées dans la base $\mathscr{B}$ sont les coefficients de X. X est la matrice des coordonnées[1] de u dans la base $\mathscr{B}$:

$$u = \sum_{i=1}^{n} x_i e_i.$$

Construire géométriquement un vecteur dont ont connaît les coordonnées dans une base précise est facile. Cependant, le vecteur obtenu dépend de la base choisie !

Exemple II.2.
Soit $\mathscr{B} = (i, j)$ la base usuelle de $\mathbb{R}^2$, où $i = (1, 0)$ et $j = (0, 1)$. Soit $u = 2i + j$ le vecteur de coordonnées $X = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ dans $\mathscr{B}$. Pour construire u, il suffit d'additionner[2] les vecteurs $2i$ et j.

Soit $\mathscr{B}' = (-i, j)$ une autre base de $\mathbb{R}^2$, et v le vecteur de coordonnées X dans $\mathscr{B}$. Alors $v = 2(-i) + j = -2i + j$, et on remarque qu'à coordonnées identiques, le vecteur obtenu dépend de la base choisie.



L'opération inverse est plus délicate. Pour construire les coordonnées d'un vecteur selon un vecteur de la base, il faut projeter ce vecteur sur la droite vectorielle engendrée par le vecteur de

1. On pourrait préciser dans ce cas que l'application qui à u associe X constitue un isomorphisme d'espace vectoriel de E sur $\mathscr{M}_{n,1}(\mathbb{R})$. En particulier, on en déduit que tout $\mathbb{R}$-espace vectoriel de dimension n est isomorphe à $\mathbb{R}^n$.

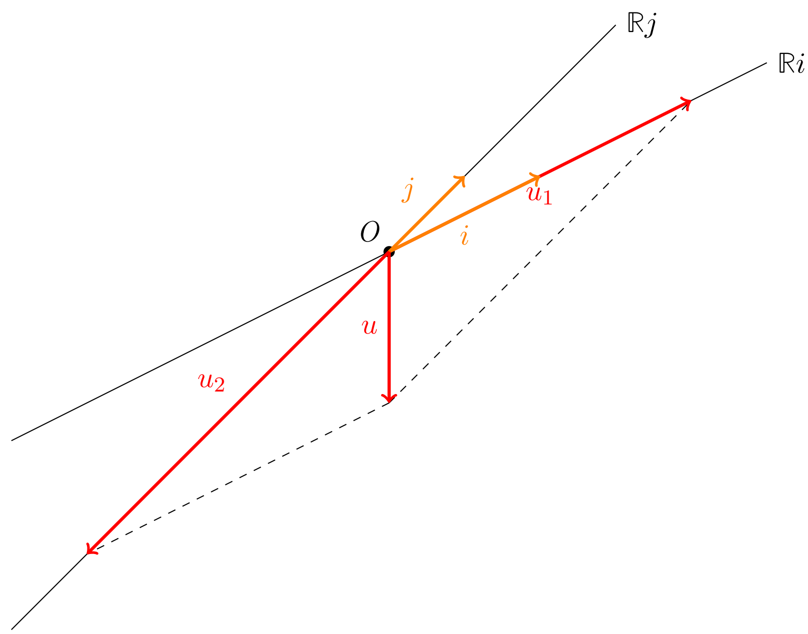
2. DOnc, géométriquement, concaténer.

la base considérée parallèlement à l'espace engendré par les autres vecteurs. Pour bien comprendre cette construction, il vaut mieux l'illustrer avec une base qui n'est pas orthogonale !

Notation : Si u est un vecteur non nul d'un $\mathbb{R}$-espace vectoriel E, alors l'espace vectoriel engendré par u est une droite vectorielle qu'on note $\mathbb{R}u = \{ku; k \in \mathbb{R}\}$. Et si E est un $\mathbb{C}$-espace vectoriel ?

Exemple II.3.

Dans le dessin suivant, $\mathscr{B} = (i, j)$ est une base du plan, et u un vecteur. Le projeté de u sur la droite vectorielle $\mathbb{R}i$ parallèlement à $\mathbb{R}j$ est u_1, et le projeté de u sur la droite vectorielle $\mathbb{R}j$ parallèlement à $\mathbb{R}i$ est u_2.



Par construction, $u = u_1 + u_2$, le vecteur u_1 est proportionnel à i, et le vecteur u_2 est proportionnel à j. On remarque graphiquement (aux erreurs de mesure près) que $u_1 \simeq 2i$ et que $u_2 \simeq -4j$. Par conséquent, $u \simeq 2i - 4j$, c'est-à-dire que les coordonnées du vecteur u dans la base $\mathscr{B}$ sont approximativement $\begin{pmatrix} 2 \\ -4 \end{pmatrix}$.

Exercice 1. Choisir deux bases $\mathscr{B}$, $\mathscr{B}'$ de $\mathbb{R}^2$ et des coordonnées X. Construire le vecteur dont les coordonnées dans $\mathscr{B}$ sont X, et déterminer graphiquement ses coordonnées dans $\mathscr{B}'$.

Exercice 2. Que se passe-t-il si on essaie de mener la construction graphique à bien pour un vecteur de $\mathbb{R}^3$?

Exercice 3. Soit u un vecteur non nul de $\mathbb{R}^2$. Montrer que l'on peut trouver une base $\mathscr{B}$ telle que u soit de coordonnées $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ dans $\mathscr{B}$. Cette base est-elle unique ?

L'énoncé de ce dernier exercice se généralise :

Proposition II.4.

Pour tout vecteur non nul u d'un espace vectoriel E et pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R}) \setminus \{0\}$, il existe une base[3] *telle que X soit la matrice des coordonnées de u dans cette base.*

La démonstration de la Proposition II.4 sera faite plus tard dans le cours.

Exercice 4. Que se passe-t-il si l'on oublie de préciser que les vecteurs u ou X sont non nuls dans la Proposition II.4 ?

II.2 La base canonique

Souvent, on identifie les coordonnées avec le vecteur, et pourtant, ce n'est pas le même objet ! Beaucoup d'espaces vectoriels ont une base "naturelle", qu'on appelle alors ***base canonique***. En voici quelques exemples qui permettent de comprendre la différence entre matrice de coordonnées et vecteur :

- ▷ La base canonique de $\mathbb{R}^2$: on rappelle que $\mathbb{R}^2 = \{(x, y) : x \in \mathbb{R}; y \in \mathbb{R}\}$. Sa base canonique comme $\mathbb{R}$-espace vectoriel est $((1, 0); (0, 1))$. Par conséquent le vecteur $u = (x, y)$ de $\mathbb{R}^2$ admet comme coordonnées dans la base canonique la matrice colonne $\begin{pmatrix} x \\ y \end{pmatrix}$. Ces deux écritures se ressemblent...
- ▷ La base canonique de $\mathbb{R}_n[X]$ est $(1, X, X^2, \cdots, X^n)$. Ainsi, le vecteur (qui est un polynôme) $u = X^n - 3X^2 + X$ a pour matrice de coordonnées dans la base canonique $\begin{pmatrix} 0 \\ 1 \\ -3 \\ \vdots \\ 1 \end{pmatrix}$: dans ce cas le vecteur et sa matrice de coordonnées sont très différents !

Il existe d'autres bases que l'on rencontre fréquemment mais qui ne sont pas les bases canoniques. Par exemple, pour les espaces de polynômes, les polynômes de Tchebychev, de Hermite ou de Bernstein apparaissent fréquemment.

Exercice 5. * On définit 4 polynômes de Bernstein $B_k^3(x) = \binom{3}{k} X^k (1 - X)^{3-k}$ dans $\mathbb{R}_3[X]$. Montrer que ces 4 polynômes forment une base de $\mathbb{R}_3[X]$.

La base canonique n'a pas de meilleures propriétés mathématiques que les autres bases de l'espace vectoriel. Il s'agit avant tout d'une convention. Par exemple, si on travaille dans l'espace vectoriel $\mathbb{R}^2$ et qu'aucune autre base n'est mentionnée, on supposera par défaut que les coordonnées sont prises dans la base canonique.

II.3 Matrices de changements de base : les groupes $GL(E)$ et $GL_n(\mathbb{R})$

Définition II.5 (Rappel : Matrices de passage)**.**

Soient E un espace vectoriel et $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. On appelle **matrice de passage** *de $\mathscr{B}$ à $\mathscr{B}'$ la matrice de l'application linéaire de E dans E qui envoie $\mathscr{B}$ sur $\mathscr{B}'$, exprimée dans la base $\mathscr{B}$.*

Propriété II.6.

3. Est-elle unique ?

1. *Une matrice de $\mathscr{M}_n(\mathbb{R})$ est une matrice de passage si et seulement si elle est inversible.*
2. *L'ensemble des matrices de passage est appelé* **groupe linéaire d'indice** *n et noté $GL_n(\mathbb{R})$. C'est un groupe multiplicatif. Il n'est pas commutatif si $n \geq 2$.*

Preuve : Nous montrons seulement le premier point. Supposons que $P \in \mathscr{M}_n(\mathbb{R})$ est une matrice de passage. Soient $\mathscr{B}$, $\mathscr{B}' = (f_1, \cdots, f_n)$ deux bases d'un espace vectoriel E de dimension n telles que P soit la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.

Nous allons montrer que le noyau de P est trivial. Soit $\lambda \in \operatorname{Ker}(P)$. Notons X_1, $\cdots$, X_n les colonnes de P. Alors
$$0 = P\lambda = \sum_{k=1}^{n} \lambda_k X_k.$$
Or X_k sont les coordonnées de f_k dans la base $\mathscr{B}$, donc
$$0 = \sum_{k=1}^{n} \lambda_k f_k.$$
Or $\mathscr{B}'$ est une base, donc est libre. Donc $\lambda = 0$.

Réciproquement, supposons que P est inversible. Notons X_1, $\cdots$, X_n les colonnes de P. Soit E un espace vectoriel de dimension n, et $\mathscr{B} = (e_1, \cdot, e_n)$ une base de $\mathbb{R}^n$. Soit $\mathscr{B}' = (f_1, \cdots, f_n)$, où f_k est le vecteur de coordonnées X_k dans la base $\mathscr{B}$.

Il faut montrer que $\mathscr{B}'$ est une base ; comme sa cardinalité est $n = \dim(E)$, il suffit de montrer que $\mathscr{B}'$ est une famille libre. Soit $\lambda \in \mathscr{M}_{n,1}(\mathbb{R})$ tel que $\sum_{k=1}^{n} \lambda_k f_k = 0$. En prenant les coordonnées de 0 dans la base $\mathscr{B}$, on obtient
$$\sum_{k=1}^{n} \lambda_k X_k = 0,$$
c'est-à-dire $\lambda \in \operatorname{Ker}(P)$. La matrice P étant inversible, le vecteur λ est nul, donc la famille $\mathscr{B}'$ est libre. □

Exercice 6. Pour montrer le premier point de la Propriété II.6, nous avons utilisé le fait que, sous les bonnes conditions de dimension et de cardinalité, une famille de vecteurs était une base si et seulement si elle était libre, et une matrice inversible si et seulement si son noyau est trivial. Reprendre cette démonstration en utilisant cette fois-ci le fait qu'une base de E engendre E. Quelle caractérisation des matrices inversibles utilise-t-on alors ?

Exercice 7. Pour montrer le seconde point de la Propriété II.6, le plus simple est de montrer que l'ensemble transformations linéaires inversibles de E, muni de la composition, est un groupe. Quels énoncés précis faut-il démontrer ? Le faire.

Exercice 8. $\mathscr{M}_{p,n}(\mathbb{R})$ est-il un groupe multiplicatif ?

Exercice 9. Soient E un espace vectoriel et $\mathscr{B}$, $\mathscr{B}'$, $\mathscr{B}''$ trois bases de E. Soit P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, et Q la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}''$. Quelle est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}''$: est-ce PQ, ou QP ?

Propriété II.7 (Rappel : Formule de changement de base)**.**
Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E et u un élément de E. On note :

▷ *P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$;*

▷ *X la matrice des coordonnées de u dans la base $\mathscr{B}$;*

▷ *X' la matrice des coordonnées de u dans la base $\mathscr{B}'$.*

Alors[4] *:*

$$X = PX'.$$

II.4 Formes linéaires et hyperplans

Nous avons vu comment, étant donnée une base, on peut associer des coordonnées à un vecteur, et comment ces coordonnées changent en fonction de la base. Nous allons ici étudier un objet dual, les formes linéaires.

Définition II.8.

Soit E un $\mathbb{R}$-espace vectoriel.

▷ *Une* **forme linéaire** *sur E est une application linéaire de E à valeurs dans un espace vectoriel de dimension 1 ou, de manière équivalente, dans $\mathbb{R}$.*

▷ *L'ensemble des formes linéaires à valeurs dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel noté E'. On l'appelle* **dual** *de E.*

▷ *Un* **hyperplan** *(vectoriel) de E est un sous-espace vectoriel de E dimension $n-1$.*

Exercice 10. Montrez qu'un sous-espace vectoriel est un hyperplan si et seulement s'il est le noyau d'une forme linéaire non nulle.

L'espace E' est l'espace des applications linéaires du $\mathbb{R}$-espace vectoriel E de dimension n dans un $\mathbb{R}$-espace vectoriel de dimension 1. Étant donnée une base $\mathscr{B} = (e_1, \cdots, e_n)$ de E et une forme linéaire ℓ, on peut donc écrire la matrice L de ℓ dans $\mathscr{B}$. Cette matrice est une matrice à 1 ligne et n colonnes :

$$L = \begin{pmatrix} \ell(e_1) & \cdots & \ell(e_n) \end{pmatrix}.$$

Attention : ne pas confondre une matrice ligne $\begin{pmatrix} x_1 & \cdots x_n \end{pmatrix}$, qui est un élément de $\mathscr{M}_{1,n}(\mathbb{R})$, avec un n-uplet $(x_1, \cdots, x_n)$, qui est un élément de $\mathbb{R}^n$. Le second sera écrit avec des virgules ou des points-virgules séparateurs.

On vérifie alors que, pour tout vecteur $u \in E$ de coordonnées X dans $\mathscr{B}$,

$$\ell(u) = LX.$$

Exemple II.9.

Munissons $\mathbb{R}^2$ d'une base quelconque $\mathscr{B} = (i, j)$. Définission $\ell : \mathbb{R}^2 \to \mathbb{R}$ par :

$$\ell(xi + yj) = 2x + y.$$

Les coordonnées de ℓ dans la base $\mathscr{B}$ sont donc $\begin{pmatrix} 2 & 1 \end{pmatrix}$.

Soient $i' = i + j$ et $j' = 2i + j$, et $\mathscr{B}' = (i', j')$ un autre base de $\mathbb{R}^2$. On calcule :

$$\ell(i') = \ell(1i + 1j) = 2 \times 1 + 1 = 3,$$
$$\ell(j') = \ell(2i + 1j) = 2 \times 2 + 1 = 5.$$

Ainsi, les coordonnées de ℓ dans la base $\mathscr{B}'$ sont $\begin{pmatrix} 3 & 5 \end{pmatrix}$.

4. Autrement dit, on obtient facilement les coordonnées dans l'ancienne base $\mathscr{B}$ à partir des coordonnées dans la nouvelle base $\mathscr{B}'$.

Comme l'exemple précédent le montre, il est très facile de calculer les coordonnées d'une forme linéaire dans la nouvelle base $\mathscr{B}'$ à partir de ses coordonnées dans l'ancienne base. Ce qui est relativement difficile, c'est l'opération inverse : exprimer les coordonnées d'une forme linéaire dans l'ancienne base en fonction des coordonnées dans la nouvelle base. De ce point de vue, les formes linéaires se comportent à l'inverse des vecteurs [5]. Cela est formalisé par la :

Propriété II.10 (Formule de changement de base pour les formes linéaires)**.**

Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E et $\ell \in E'$ une forme linéaire sur E. On note :

- ▷ *P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$;*
- ▷ *L la matrice des coordonnées de ℓ dans la base $\mathscr{B}$;*
- ▷ *L' la matrice des coordonnées de ℓ dans la base $\mathscr{B}'$.*

Alors [6] :

$$L' = LP.$$

Preuve : Soient $\mathscr{B}$, $\mathscr{B}'$, ℓ, L et L' comme dans l'énoncé. Il suffit de montrer que, pour tout $X' \in \mathscr{M}_{n,1}(\mathbb{R})$,

$$L'X' = LPX'.$$

Soit u le vecteur de coordonnées X' dans $\mathscr{B}'$, et X les coordonnées de u dans $\mathscr{B}$. Par les formules de changement de base, $PX' = X$, donc $LPX' = LX$. L'égalité souhaitée découle du fait que LX et $L'X'$ sont tous deux égaux à $\ell(u)$, donc $LX = L'X'$. □

Un deuxième exemple plus abstrait est celui des bases duales [7].

Exemple II.11.

Soit $\mathscr{B} = (e_1, \cdots, e_n)$ une base de E. Tout vecteur $u \in E$ s'écrit en coordonnées dans cette base :

$$u = \sum_{k=1}^{n} a_k e_k.$$

Les coefficients $(a_1, \cdots, a_n)$ dépendent de u. On peut donc les voir comme des fonctions $a_k : E \to \mathbb{R}$. La famille $(a_1, \cdots, a_n)$ est une base de E', que l'on appelle **base duale** [8] *de $\mathscr{B}$.*

Preuve : Montrons que la base duale $(a_1, \cdots, a_n)$ est effectivement une base de E'. Par cardinalité, il suffit de montrer que c'est une famille libre. Soit $\lambda \in \mathbb{R}^n$ tel que $\sum_{k=1}^{n} \lambda_k a_k = 0$. Soit $1 \leq i \leq n$. Alors :

$$0 = \left(\sum_{k=1}^{n} \lambda_k a_k \right)(e_i) = \sum_{k=1}^{n} \lambda_k a_k(e_i).$$

Or $a_k(e_i)$ est, par définition de la base duale, le k-ième coefficient de e_i dans la base $(e_1, \cdot, e_n)$. C'est donc par définition 1 si $k = i$, et 0 si $k \neq i$. Par conséquent, $0 = \lambda_k$. Comme cela est vrai pour tout k, on a $\lambda = 0$, donc la famille $(a_1, \cdots, a_n)$ est bien libre. □

---

5. On dit aussi que les formes linéaires sont contravariantes.
6. Autrement dit, on obtient facilement les coordonnées dans la nouvelle base $\mathscr{B}$ à partir des coordonnées dans l'ancienne base $\mathscr{B}'$.
7. Donné uniquement en exemple. Il ne sera pas exigé de le connaître dans le cadre de ce cours.
8. Souvent notée $(e_1', \cdots, e_n')$ ou $(e^1, \cdots, e^n)$. Nous éviterons la première pour ne pas porter à confusion avec la notaiton $\mathscr{B}'$ que nous utilisons. La seconde n'est intéressant que lorsque l'on manipule les notations d'Einstein, ce qui va bien au-delà des objectifs de ce cours.

Exercice 11. Décrire géométriquement le noyau de a_1. Application : compléter $(1,1,0)$ et $(0,1,1)$ en une base de $\mathbb{R}^3$. Calculer sa base duale. En déduire une écriture implicite (c'est-à-dire comme noyau d'une forme linéaire) du plan vectoriel engendré par les vecteurs $(1,1,0)$ et $(0,1,1)$.

Pour finir, remarquons que, si $(a_1, \cdots, a_n)$ est la base duale de $\mathscr{B}$, alors les coordonnées de a_k dans la base $\mathscr{B}$ sont données par $\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$, où le coeeficient 1 est à la k-ième coordonnée. C'est d'ailleurs l'essence de la démonstration précédente.

Par conséquent, si $(a_1, \cdots, a_n)$ est la base duale de $\mathscr{B}$ et P la matrice de $\mathscr{B}$ à $\mathscr{B}'$, alors les coordonnées de a_k dans $\mathscr{B}$ sont $L = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$. Par la Propriété II.10, les coordonnées L' de a_k dans la base $\mathscr{B}'$ sont LP, c'est-à-dire la k-ième ligne de P. Pour résumer :

1. Les colonnes de P sont les coordonnées des vecteurs de la nouvelle base $\mathscr{B}'$ dans l'ancienne base $\mathscr{B}$;
2. Les lignes de P sont les coordonnées des formes linéaires de la base duale de $\mathscr{B}$ dans la nouvelle base $\mathscr{B}'$.

Intuitivement, ce sont deux façons de dire la même chose : connaître P, c'est pouvoir passer facilement des coordonnées dans $\mathscr{B}'$ aux coordonnées dans $\mathscr{B}$. Or, prendre les coordonnées, c'est par définition appliquer les éléments de la base duale. En inversant P, on a de plus :

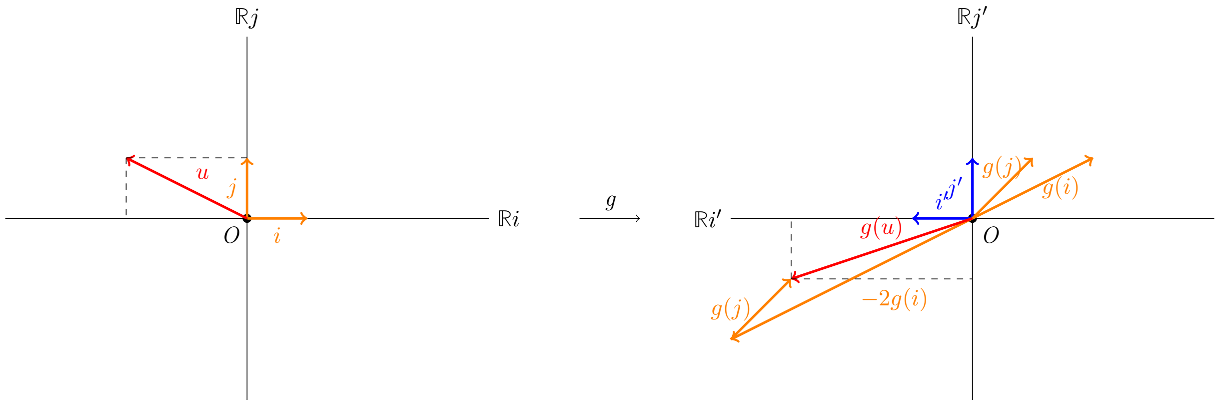
1. Les colonnes de P^{-1} sont les coordonnées des vecteurs de l'ancienne base $\mathscr{B}$ dans la nouvelle base $\mathscr{B}'$;
2. Les lignes de P^{-1} sont les coordonnées des formes linéaires de la base duale de $\mathscr{B}'$ dans l'ancienne base $\mathscr{B}$.

II.5 Matrices et applications linéaires : des bases, obligatoirement !

Nous avons vu qu'en fixant une base $\mathscr{B}_E$ de E, on pouvait associer à tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$ un vecteur u de E. De même, si F est un espace vectoriel de dimension p, et $\mathscr{B}_F$ une base de F, on associe à $B \in \mathscr{M}_{p,1}(\mathbb{R})$ un vecteur $b \in F$. Alors on peut interpréter l'égalité $AX = B$ comme l'image de X par A étant égale à B : comme X est associé à $u \in E$ et B à $b \in F$, A correspond donc à une fonction qui envoie u sur un vecteur de F : cette fonction est une application linéaire de E dans F. On note $\mathscr{L}(E, F)$ l'ensemble de ces applications linéaires.

Théorème 1.
Toute application linéaire est déterminée de manière unique par son action sur une base.

Exemple II.12.
Prenons $E = F = \mathbb{R}^2$, et $\mathscr{B}_E = (i, j)$, $\mathscr{B}_F = (i', j')$. Considérons une transformation $g : E \to F$. Le vecteur u suivant se décompose dans la base $\mathscr{B}_E$ en $u = -2i + j$.



Une transformation $h : E \to F$ est linéaire si pour tous scalaires λ, μ et tous vecteurs u, v,

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

Par linéarité, on observe donc que $g(u) = -2g(i) + g(j)$. Le vecteur $-2g(i) + g(j)$ peut ensuite se décomposer dans la bse $\mathscr{B}_F$. On trouve graphiquement, dans l'exemple ci-dessus, que les coordonnées de $g(u)$ dans $\mathscr{B}'$ sont $X' \simeq \begin{pmatrix} 3 \\ -1 \end{pmatrix}$.

Soit M la matrice de u de $(E, \mathscr{B}_E)$ dans $(F, \mathscr{B}_F)$. Alors $M \begin{pmatrix} 2 \\ 1 \end{pmatrix} \simeq \begin{pmatrix} 3 \\ -1 \end{pmatrix}$.

Remarque II.13.
Ce théorème n'est pas exprimé dans un langage mathématique précis. Une version plus précise serait la suivante. Soient E, F deux espaces vectoriels et $\mathscr{B}_E = (e_1, \cdots, e_n)$ une base de E. Soient $(f_1, \cdots, f_n)$ une famille de $n = \dim(E)$ vecteurs dans F. Alors il existe une unique application linéaire $u : E \to F$ telle que $u(e_k) = f_k$ pour tout $1 \leq k \leq n$.

Preuve : Soient E, F, $\mathscr{B}_E$ et $(f_1, \cdots, f_n)$ comme précédemment.

Soient g, $g' : E \to F$ deux applications linéaires envoyant e_k sur f_k pour tout k. Soit $u \in E$ et X les coordonnées de u dans $\mathscr{B}_E$. Alors, par linéarité de g et g',

$$g(u) = g\left(\sum_{k=1}^{n} X_k e_k\right) = \sum_{k=1}^{n} X_k g(e_k) = \sum_{k=1}^{n} X_k f_k = \sum_{k=1}^{n} X_k g'(e_k) = g'\left(\sum_{k=1}^{n} X_k e_k\right) = g'(u).$$

Comme u est arbitraire, $g = g'$. Il y a donc au plus une telle application.

L'existence d'une telle application est similaire. Pour tout $u \in E$, soient X les coordonnées de u. On pose :

$$g(u) := \sum_{k=1}^{n} X_k f_k.$$

Il reste à montrer que g est linéaire, ce qui découle de la linéarité des coordonnées. □

Définition II.14 (Rappel : Représentation d'une application linéaire par une matrice)**.**
Soient $f \in \mathscr{L}(E, F)$ et $\mathscr{B}_E$ et $\mathscr{B}_F$ des bases de E et F. Alors la matrice associée à f dans les bases précédentes est la matrice formée des n matrices des coordonnées des images de la base de E exprimées dans la base de F.

Le choix des bases influence la représentation de l'application linéaire f associée à A. L'idée consiste alors à partir de l'étude de f de trouver la "meilleure" base pour résoudre le système (et donc le "meilleur" changement de coordonnées). La notion sous-jacente est algébrique.

Exemple II.15.

Soit $\mathscr{B}$, $\mathscr{B}'$ deux bases de E et P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$. Alors P est la matrice de l'application de $(E, \mathscr{B})$ dans $(E, \mathscr{B})$ qui envoie les élément de $\mathscr{B}$ sur ceux de $\mathscr{B}'$.

Cependant, on vérifie aussi que P est la matrice de l'identité $id : E \to E$, où l'on a muni la source de la base $\mathscr{B}'$ et l'arrivée de $\mathscr{B}$.

Question II.16.

Pour une application linéaire donnée $f : E \to F$, il existe une famille de matrices possibles qui la représentent, en fonction des bases choisies. Le problème revient à déterminer des représentants "efficaces". Une question associée est de savoir combien de représentants permettent de décrire toutes les applications linéaires.

Exercice 12. Choisissez deux familles de vecteurs. Y-a-t-il 0, 1, plusieurs applications linéaires permettant de passer de l'une à l'autre ?

III Matrices équivalentes et théorème du rang

III.1 Le rang d'une application linéaire ou d'une matrice

Définition III.1 (Rappel : Rang d'une application linéaire)**.**

Le rang d'une application linéaire f est la dimension de son image : $rg(f) = \dim(Im(f))$.

Propriété III.2.

Soit A la matrice associée à f dans des bases quelconques. Alors $rg(f)$ est égal à la dimension de l'espace engendré par les colonnes de A, que l'on note aussi $rg(A)$.

Preuve : L'image de f est engendrée par les vecteurs $f(e_k)$, où e_k est une base de E. En particulier, il existe une famille de $\mathrm{rg}(f)$ vecteurs $(f(e_{i_1}), \cdots, f(e_{i_{\mathrm{rg}(f)}}))$ libre. Donc les colonnes des coordonnées de $f(e_{i_1})$, ..., $f(e_{i_{\mathrm{rg}(f)}})$ dans la base de F sont aussi libres. Or ce sont exactement les colonnes i_1, ..., $i_{\mathrm{rg}(f)}$ de la matrice A. Donc la matrice A est de rang au moins $\mathrm{rg}(f)$, c'est-à-dire que $\mathrm{rg}(A) \geq \mathrm{rg}(f)$.

L'inégalité $\mathrm{rg}(f) \geq \mathrm{rg}(A)$ s'obtient en faisant le raisonnement précédent en sens inverse : on trouve $\mathrm{rg}(A)$ colonnes de la matrice A qui sont libres, on vérifie que les $\mathrm{rg}(A)$ vecteurs correspondant dans F sont aussi libres, ce qui permet de conclure. □

III.2 Matrices équivalentes

Définition III.3.

Soient A et B deux matrices de $\mathscr{M}_{p,n}(\mathbb{R})$. Ces matrices sont **équivalentes** *s'il existe deux matrices $P \in GL_n(\mathbb{R})$ et $Q \in GL_p(\mathbb{R})$ telles que $A = QBP^{-1}$.*

Propriété III.4.

Deux matrices sont équivalentes si et seulement si elles peuvent être associées à la même application linéaire. [9]

9. Quitte à choisir des bases différentes.

Preuve : Commençons par le sens indirect. Soient A, $B \in \mathscr{M}_{p,n}(\mathbb{R})$ associées à la même application linéaire $g : E \to F$. Soient $\mathscr{B}_E$, $\mathscr{B}_F$ des bases de E et F respectivement telles que A soit la représentation matricielle de $g : (E, \mathscr{B}_E) \to (F, \mathscr{B}_F)$. Soient P la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$ et Q la matrice de passage de $\mathscr{B}_F$ à $\mathscr{B}'_F$.

Il suffit de montrer que, pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$, on a $AX = QBP^{-1}X$. Soit $X \in \mathscr{M}_{n,1}(\mathbb{R})$. Soit $u \in E$ dont les coordonnées dans $\mathscr{B}_E$ sont X.

Alors les coordonnées de u dans $\mathscr{B}'_E$ sont $P^{-1}X$. Les coordonnées de $g(u)$ dans $\mathscr{B}'_F$ sont $BP^{-1}X$ par définition de B. Enfin, les coordonnées de $g(u)$ dans $\mathscr{B}_F$ sont $QBP^{-1}X = AX$, ce qui finit la démonstration. □

Propriété III.5.

La relation "être équivalentes" est une relation d'équivalence sur l'ensemble des matrices $\mathscr{M}_{n,p}(\mathbb{R})$.

Preuve : Il faut démontrer la réflexivité, la symétrie et la transitivité de cette relation. Notons-la $\sim$.

1. Réflexivité : soit $A \in \mathscr{M}_{p,n}(\mathbb{R})$. Choisissons $P = I_n$ et $Q = I_p$. On obtient $A \sim A$.
2. Symétrie : soient A, $B \in \mathscr{M}_{p,n}(\mathbb{R})$ telles que $A \sim B$. Soient P et Q telles que $A = QBP^{-1}$. Alors $B = Q^{-1}A(P^{-1})^{-1}$, et Q^{-1} et P^{-1} sont dans $GL_p(\mathbb{R})$ et $GL_n(\mathbb{R})$ respectivement (les groupes linéaires sont clos par inversion). Donc $B \sim A$.
3. Transitivité : soient A, B, $C \in \mathscr{M}_{p,n}(\mathbb{R})$ telles que $A \sim B$ et $B \sim C$. Soient P, Q, P', Q' telles que $A = QBP^{-1}$ et $B = Q'C(P')^{-1}$. Alors $A = Q(Q'C(P')^{-1})P^{-1} = (QQ')C(PP')^{-1}$, et QQ' et PP' sont dans $GL_p(\mathbb{R})$ et $GL_n(\mathbb{R})$ respectivement (les groupes linéaires sont clos par multiplication). Donc $A \sim C$.

□

Exercice 13. Montrez que, si $A \sim B$ sont deux matrices et λ un réel, alors $\lambda A \sim \lambda B$. La réciproque est-elle vraie ?

Attention : $\mathscr{M}_{n,p}(\mathbb{R})$ n'est pas un groupe pour la multiplication. Même si $p = n$, les classes d'équivalence de $\mathscr{M}_n(\mathbb{R})$ ne sont a priori closes ni par addition, ni par multiplication. Autrement dit, si $A_1 \sim B_1$ et $A_2 \sim B_2$, il n'est pas vrai a priori que $A_1 + B_1$ est équivalent à $A_2 + B_2$.

Exercice 14. Trouvez un contre-exemple à la remarque précédente !

Proposition III.6.

Le rang est un **invariant** *pour les matrices équivalentes : si A et B sont équivalentes alors $rg(A) = rg(B)$.*

Preuve : Soient P et Q inversibles telles que $A = QBP^{-1}$.

La matrice P^{-1} étant inversible, $\mathrm{Im}(BP^{-1}) = \mathrm{Im}(B)$. Donc $\mathrm{Im}(A) = \mathrm{Im}(QB) = Q\mathrm{Im}(B)$. Enfin, Q étant inversible, $Q\mathrm{Im}(B)$ et $\mathrm{Im}(B)$ ont même dimension. Donc $\mathrm{rg}(A) = \mathrm{rg}(B)$. □

Définition III.7.

Soit $A \in \mathscr{M}_{p,n}(\mathbb{R})$. L'ensemble des matrices équivalentes à A s'appelle la **classe d'équivalence de** *A (pour la relation d'équivalence précédente). C'est l'ensemble*

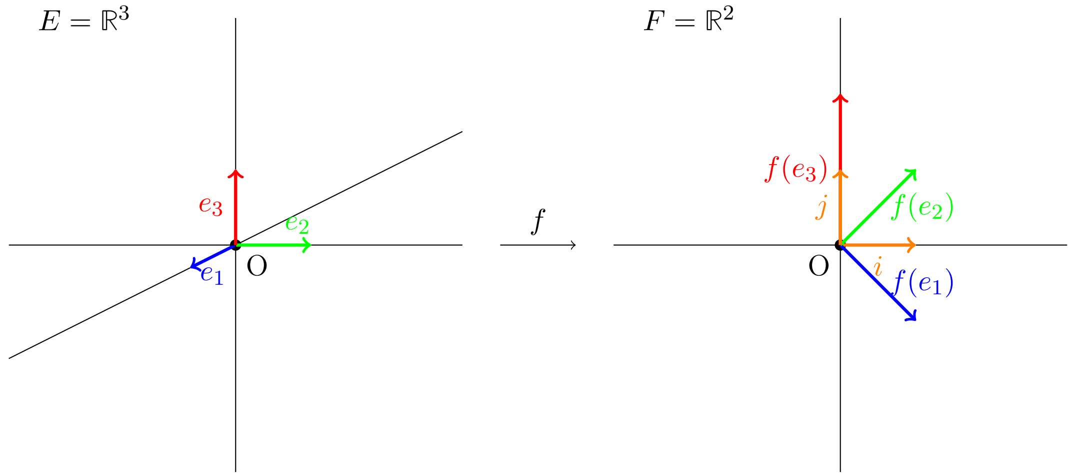
$$\{Q^{-1}AP; P \in GL_n(\mathbb{R}); Q \in GL_p(\mathbb{R})\}.$$

Par ce qui précède, si Ω est une classe d'équivalence, alors $\Omega \subset \{\mathrm{rg} = k\}$ pour un certain entier k. Il n'est pas évident a priori que $\Omega = \{\mathrm{rg} = k\}$ pour ce même entier.

III.3 Réduction d'une matrice par équivalence, exemple

Question : On voudrait maintenant déterminer, pour une application linéaire donnée, les matrices les "plus simples" pouvant être associées à f. On travaille sur un exemple.

Nous allons travailler sur l'exemple suivant. Soit f une application linéaire de $E = \mathbb{R}^3$ dans $F = \mathbb{R}^2$. On note (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$ et (i, j) celle de $\mathbb{R}^2$. On suppose que f agit de la façon suivante :



La matrice de f dans ces deux bases est $A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}$. On cherche de nouvelles bases qui vont permettre de simplifier la matrice de f. Il y a plusieurs étapes :

1. On commence par transformer la base de F en choisissant parmi les vecteurs images de la base ceux qui forment une base de $\mathrm{Im}(f)$. Ici, (f_1, f_2, f_3) sont 3 vecteurs en dimension 2, donc ils sont liés. On peut choisir comme base f_1 et f_3 car ils ne sont pas colinéaires, donc engendrent un espace de dimension 2 qui est donc $\mathbb{R}^2$ tout entier. [10]
2. Dans la nouvelle base de F, la matrice s'écrit $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ (ce qui était difficile, c'était de trouver les coordonnées de f_2 dans la base (f_1, f_3)).
3. On cherche maintenant à simplifier encore la matrice. Pour cela on va transformer la base canonique de E. On procède en 2 étapes :
   - ▷ On remet dans l'ordre les vecteurs dont l'image est l'un des vecteurs de la nouvelle base de F, en gardant le même ordre que sur F : e_1 (pour f_1) et ensuite e_3 (pour f_3).
   - ▷ On laisse ensuite ceux dont l'image vaut 0 : pourquoi y en a-t-il ? Parce que la famille (f_1, f_2, f_3) est liée : $f_2 = f_1 + f_3$, donc $f_1 - f_2 + f_3 = 0$. On en déduit que $f(e_1 - e_2 + e_3) = 0$. On pose alors $u = e_1 - e_2 + e_3$: (e_1, e_3, u) est une base car la coordonnée de u selon e_2 n'est pas nulle et que (e_1, e_3, e_2) est une base.

10. Attention, peut-être que dans le cas général, on ne pourra pas avoir toute une base de F, mais seulement une partie : cela dépend de la dimension de l'image de f, c'est-à-dire de $r = \mathrm{rg}(f)$.

4. Dans les bases (e_1, e_3, u) et (f_1, f_3), la matrice de f est donc $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Nous avons trouvé la forme la plus "simple" sous laquelle peut s'écrire une matrice associée à f.[11]
5. On écrit enfin la relation entre les matrices : on note P la matrice de passage vers la nouvelle base de E et Q la matrice de passage vers la nouvelle base de F.

Exercice 15. Dans l'exemple ci-dessus, expliciter les matrices P et Q, et vérifier l'égalité trouvée.

Exercice 16. Les nouvelles bases construites ci-dessus ne sont pas les seules dans lesquelles la matrice de f est B. En trouver d'autres.

III.4 Théorème de réduction par équivalence et théorème du rang

En généralisant l'exemple précédent, on peut énoncer le théorème suivant :

Théorème 2.
Soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$ une matrice ; on note r son rang. Alors il existe $P \in GL_n(\mathbb{R})$ et $Q \in GL_p(\mathbb{R})$ telles que

$$A = Q \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P^{-1}. \tag{1.1}$$

Autrement dit, toute matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ est équivalente à une matrice de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ avec $r \leq n$.

Preuve : On reprend l'exemple proposé. On choisit $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ et on se donne $\mathscr{B}_E = (e_1, \cdots, e_n)$ et $\mathscr{B}_F$ les bases canoniques associées. On note f l'application linéaire associée à A dans $\mathscr{B}_E$ et $\mathscr{B}_F$. On note $r = \operatorname{rg}(f)$.

- ▷ Première étape : choix de la nouvelle base de F. Si on appelle $f_1, \cdots, f_n$ les images de $e_1, \cdots, e_n$ par f, on peut choisir r vecteurs parmi la famille $(f_1, \cdots, f_n)$ formant une famille libre (par définition du rang)
  - ▷ On construit avec ces r vecteurs le début d'une base de F , notée $(f'_1, \cdots, f'_r)$. Comme dans l'exemple précédents, ce sont les f_i qui sont libres ; il y a du choix !
  - ▷ Comme il est tout à fait possible que $r < p$, il y a besoin de compléter les vecteurs précédents pour former une base : c'est toujours possible, il suffit de prendre une base d'un supplémentaire de $\operatorname{Im}(f)$.

  On obtient donc une nouvelle base $\mathscr{B}_F = (f'_1, \cdots, f'_p)$ de F, dans laquelle si $k \leq r$ alors f'_k est l'image d'un des vecteurs d'un des e_i, et sinon un vecteur qui complète l'image de f pour construire une base de F.
- ▷ Deuxième étape : choisir la nouvelle base de E. Elle se réalise également en deux temps.
  - ▷ Pour $i \leq r$ on choisit e'_i parmi les vecteurs $(e_1, \cdots, e_n)$ de sorte qu'on ait $f(e'_i) = f'_i$ (on met dans l'ordre les e_i). Il reste à choisir les suivants. On va les choisir de sorte que leur image soit nulle.

11. On peut remarquer que la base trouvée n'est pas nécessairement unique ! Trouver d'autres bases dans lesquelles ces écritures sont aussi simples.

▷ Pour finir, le plus difficile : la construction d'une base de $\mathrm{Ker}(f)$. Notons $e''_{r+1}, \cdots, e''_n$ les vecteurs de $(e_1, \cdots, e_n)$ que l'on n'a pas encore utlisé. Comme $(f'_1, \cdots, f'_r)$ engendre $\mathrm{Im}(f)$, pour tout $i > r$, le vecteur $f(e''_i)$ est une combinaison linéaire de $(f_1, \cdots, f_r)$. Par exemple, pour $i = r+1$, on peut écrire

$$f_{r+1} = \sum_{i=1}^{r} \lambda_i f_i.$$

Alors $e'_{r+1} = e''_{r+1} - \sum_{i=1}^{r} \lambda_i e_i$ vérifie $f(e'_{r+1}) = 0$ et la famille $(e'_1, \cdots, e'_r, e'_{r+1})$ est de même rang que $(e'_1, \cdots, e'_r, e''_{r+1})$ qui est libre. On peut recommencer le raisonnement pour les suivants, ce qui permet de construire $e'_{r+1}, \cdots, e'_n$ libres et d'image nulle par f : il s'agit en fait de la base du noyau $\mathrm{Ker}(f)$ de f.

□

Conséquence III.8 (Théorème du rang)**.** *Pour toute application linéaire $f : E \to F$,*

$$\dim(E) = rg(f) + \dim(Ker(f)).$$

PREUVE : Soit r le rang de f. Soient $\mathscr{B}_E = (e_1, \cdots, e_n)$ et $\mathscr{B}_F = (f_1, \cdots, f_p)$ des bases de E et F respectivement telles que l'écriture de f dans ces bases soit $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Alors $\mathrm{Im}(f) = \mathrm{Vect}(f_1, \cdots, f_r)$ et $\mathrm{Ker}(f) = \mathrm{Vect}(e_{r+1}, \cdots, e_n)$. Par conséquent,

$$\dim(E) = n = r + (n-r) = \mathrm{rg}(f) + \dim(\mathrm{Ker}(f)),$$

d'où le théorème du rang. □

Exercice 17. Reconnaître des fonctions linéaires associées à ces opérations sur les bases (symétries/projections ; symétries/rotations ; cycles/transpositions...).

Exercice 18. Savoir-faire : mettre des matrices sous la forme de l'Équation (1.1). Reconnaître des matrices équivalentes.

III.5 Conclusion

Les classes d'équivalences pour l'équivalence sont uniquement fonction du rang de l'application linéaire associée (ou de la matrice considérée) : le rang est un ***invariant complet*** pour cette relation d'équivalence.

Exercice 19.

1. Montrer que $\mathscr{M}_{n,p}(\mathbb{R})$ a exactement $\min\{n, p\} + 1$ classes d'équivalences pour cette relation.
2. En déduire quelles sont les classes d'équivalences des matrices de $\mathscr{M}_{n,1}(\mathbb{R})$. Interprétez ce résultat.
3. Dans $M_n(\mathbb{R})$, décrire la classe d'équivalence associée à la matrice identité I_n. Interpréter.

Remarque III.9. *Limites : Le rang est insuffisant pour rendre compte de l'information apportée concernant la "ressemblance" géométrique des applications linéaires de E dans E. Par exemple, dans ce cas, on peut itérer les applications. Pour conserver le lien avec les matrices, cela contraint à prendre la même base au départ et à l'arrivée, contrainte que l'on ne s'est pas imposée dans ce qui précède. C'est l'objet de l'étude des matrices semblables.*

III.6 Compléments d'algèbre : groupes

Définition III.10.

Un groupe est un ensemble G muni d'une loi "sympathique". Une loi de composition interne est une application de $G \times G$ dans G ; c'est une loi de groupe si on peut lui trouver un élément neutre 1, qu'elle est associative, et que tous les éléments de G admettent un inverse.

Remarque III.11.

En général, il n'y a pas besoin de construire la loi : celle-ci existe déjà. Il est donc beaucoup plus simple dans ce cas de vérifier qu'un ensemble G est un groupe. Il suffit de vérifier que :

- ▷ *la loi est interne dans G : pour tous (g, g') dans $G \times G$, on a aussi $g\dot{g}' \in G$ (G est stable pour la loi).*
- ▷ *G contient l'unité 1.*
- ▷ *G est stable pour l'inverse : pour tout $g \in G$, g^{-1} est bien défini et $g^{-1} \in G$.*

Il y a encore un cas plus simple, le cas des sous-groupes :

Définition III.12. *Soit G un groupe multiplicatif (c'est dire qu'on note sa loi comme une multiplication). One dit que H est un* **sous-groupe** *de G si :*

- ▷ *$1 \in H$ (H contient l'unité).*
- ▷ *$hh' \in H$ pour tous h, h' dans H (H est stable pour la multiplication).*
- ▷ *$h^{-1} \in H$ pour tout h dans H (H est stable pour l'inverse).*

Exemple III.13.

Voici quelques exemples de groupes et de sous-groupes.

- ▷ *$\mathbb{R}$ est un groupe additif, $\mathbb{Z}$ et $\mathbb{Q}$ sont des sous-groupes (additifs) de $\mathbb{R}$.*
- ▷ *$\mathbb{R}^*$ est un groupe multiplicatif, $\mathbb{Q}^*$ est un sous-groupe multiplicatif de $\mathbb{R}^*$. Pourquoi pas $\mathbb{Z}^*$?*
- ▷ *$\mathbb{R}^n$ est un groupe additif, mais $\mathbb{R}^n \setminus \{0\}$ n'est pas un groupe multiplicatif en général (savez-vous définir une multiplication ?)*
- ▷ *$\mathscr{M}_{n,p}(\mathbb{R})$ est un groupe additif, mais il n'y a pas de multiplication, sauf si $n = p$.*
- ▷ *$\mathscr{M}_n(\mathbb{R})$ est un groupe additif, et possède une multiplication qui est stable, associative et admet une unité (I_n). En revanche, ce n'est pas un groupe. Comme cette multiplication est aussi distributive avec l'addition, on dit que c'est une algèbre.*
- ▷ *$GL_n(\mathbb{R})$ est un sous-ensemble de $\mathscr{M}_n(\mathbb{R})$ et un groupe multiplicatif (mais pas additif !).*

Vous pourrez compléter cette liste en cours de semestre !

IV Matrices semblables et conjugaison

Dans cette partie on étudie le cas des endomorphismes : $E = F$.

Soit $f : E \to E$. L'espace de départ et d'arrivée étant identiques, on peut itérer f (c'est-à-dire le composer avec lui-même). On notera f^2, f^3... ces itérés.

Si l'on écrit la matrice A de f avec des bases différentes au départ (disons, $\mathscr{B}$) et à l'arrivée (disons, $\mathscr{B}'$), il n'y a pas de lien simple entre les puissances de f et celles de A. En effet, étant donné un vecteur $u \in E$ de coordonnées X dans $\mathscr{B}$, les coordonnées de $f(u)$ dans $\mathscr{B}'$ sont AX. Cependant, on ne connaît pas les coordonnées de $f(u)$ dans $\mathscr{B}$, ce qui empêche d'appliquer A à nouveau !

Dans ce cas, il sera donc utile d'avoir la même base au départ et à l'arrivée. On pourra chercher une base qui soit la plus simple possible pour représenter l'endomorphisme f, mais cette contrainte supplémentaire nous empêchera en général de trouver un représentant aussi simple que celui donné par le Théorème 2.

IV.1 Matrices semblables

On adapte la définition de matrices équivalentes.

Définition IV.1 (Rappel : Matrices semblables)**.**
Soient A et B deux matrices de $\mathscr{M}_n(\mathbb{R})$. Ces matrices sont semblables (sur $\mathbb{R}$) s'il existe $P \in GL_n(\mathbb{R})$ tel que[12] $B = P^{-1}AP$.

On dit que A et B sont semblables sur $\mathbb{C}$ si on peut trouver $P \in GL_n(\mathbb{C})$ tel que $B = P^{-1}AP$.

Propriété IV.2.
Deux matrices sont semblables si et seulement si elles peuvent être associées au même endomorphisme (quitte à changer de base, mais en gardant la même base au départ et à l'arrivée).

Preuve : On reprend la démonstration de la propriété similaire pour les matrices équivalentes. Commençons par le sens indirect. Soient A, $B \in \mathscr{M}_n(\mathbb{R})$ associées au même endomorphisme $g : E \to E$. Soit $\mathscr{B}$ (respectivement, $\mathscr{B}'$) une base de E telle que A (respectivement, B) soit la représentation matricielle de $g : (E, \mathscr{B}) \to (E, \mathscr{B})$ (respectivement, de $g : (E, \mathscr{B}') \to (E, \mathscr{B}')$). Soient P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$. On se retrouve dans la situation de la démonstration pour des matrices équivalentes avec $P = Q$, d'où $A = PBP^{-1}$.

Réciproquement, soient A, B deux matrices semblables, et soit P telle que $A = PBP^{-1}$. Alors on peut réaliser A et B comme matrices d'un même endomorphisme linéaire de $\mathbb{R}^n$ dans deux bases, et P sera la matrice de passage entre ces deux bases. □

Propriété IV.3.
La relation "être semblables" est une relation d'équivalence sur l'ensemble des matrices $\mathscr{M}_n(\mathbb{R})$.

Preuve : Il faut démontrer la réflexivité, la symétrie et la transitivité de cette relation. Notons-la $\sim'$.

1. Réflexivité : soit $A \in \mathscr{M}_n(\mathbb{R})$. Choisissons $P = I_n$. On obtient $A \sim' A$.
2. Symétrie : soient A, $B \in \mathscr{M}_n(\mathbb{R})$ telles que $A \sim' B$. Soit P telle que $A = PBP^{-1}$. Alors $B = P^{-1}A(P^{-1})^{-1}$, et $P^{-1} \in GL_n(\mathbb{R})$ (les groupes linéaires sont clos par inversion). Donc $B \sim' A$.
3. Transitivité : soient A, B, $C \in \mathscr{M}_n(\mathbb{R})$ telles que $A \sim' B$ et $B \sim' C$. Soient P, P' telles que $A = PBP^{-1}$ et $B = P'C(P')^{-1}$. Alors $A = P(P'C(P')^{-1})P^{-1} = (PP')C(PP')^{-1}$, et $PP' \in GL_n(\mathbb{R})$ respectivement (les groupes linéaires sont clos par multiplication). Donc $A \sim' C$.

□

12. C'est-à-dire que $A = PBP^{-1}$. Comparez avec la définition de matrices équivalentes.

Propriété IV.4.

Deux matrices semblables sont équivalentes. En particulier, elles sont de même rang.

Définition IV.5.

Soit $A \in \mathscr{M}_n(\mathbb{R})$. L'ensemble des matrices semblables à A s'appelle la **classe de similitude** *de A. C'est l'ensemble*

$$\{P^{-1}AP; \quad P \in GL_n(\mathbb{R})\}.$$

Par définition de la relation "être semblable", si $f : E \to E$ est une applicaiton linéaire et A une représentation matricielle de f, alors la classe de similitude de A est l'ensemble des représentations matricielles de f.

Propriété IV.6.

$\mathscr{M}_n(\mathbb{R})$ est une algèbre, c'est à dire que c'est un groupe additif et qu'il est stable pour la multiplication.

Exercice 20. Soient A_1, A_2 deux matrices semblables. Soient B_1, B_2 deux matrices semblables. Démontrer ou donner un contre-exemple des assertions suivantes :

- $A_1 + B_1$ et $A_2 + B_2$ sont semblables ;
- A_1B_1 et A_2B_2 sont semblables ;
- λA_1 et λA_2 sont semblables, où λ est un réel quelconque.

IV.2 Endomorphismes conjugués

Définition IV.7 (Rappel : Endomorphismes)**.**

- *Un* **endomorphisme** *de E est une application linéaire de E dans E. L'ensemble des endomorphismes de E est noté $L(E)$.*
- *L'ensemble des endomorphismes bijectifs de E est appelé le* **groupe linéaire** *de E. On le note $GL(E)$.*

Exercice 21.

1. Montrer que $GL(E)$ est un groupe multiplicatif.
2. * Quel est le lien entre $GL_n(\mathbb{R})$ et $GL(E)$? Montrer que ces deux ensembles sont isomorphes (comme groupes). (Rappeler la définition d'un morphisme de groupe, puis d'un isomorphisme).

Propriété IV.8.

$\mathscr{M}_n(\mathbb{R})$ est l'ensemble des matrices associées aux endomorphismes d'un $\mathbb{R}$-espace vectoriel.

Définition IV.9.

On dit que deux endomorphismes f et g de $L(E)$ sont **conjugués** *s'il existe $h \in GL(E)$ telle que $f = h^{-1} \circ g \circ h$.*

Exemple IV.10.

Soient $\mathcal{P}$ le plan euclidien et O un point de $\mathcal{P}$. Soit f la rotation autour de O d'angle $\pi/3$ et h la symétrie axiale relativement à une droite Δ passant par 0. Quel est la transformation conjuguée $g = h \circ f \circ h^{-1}$?

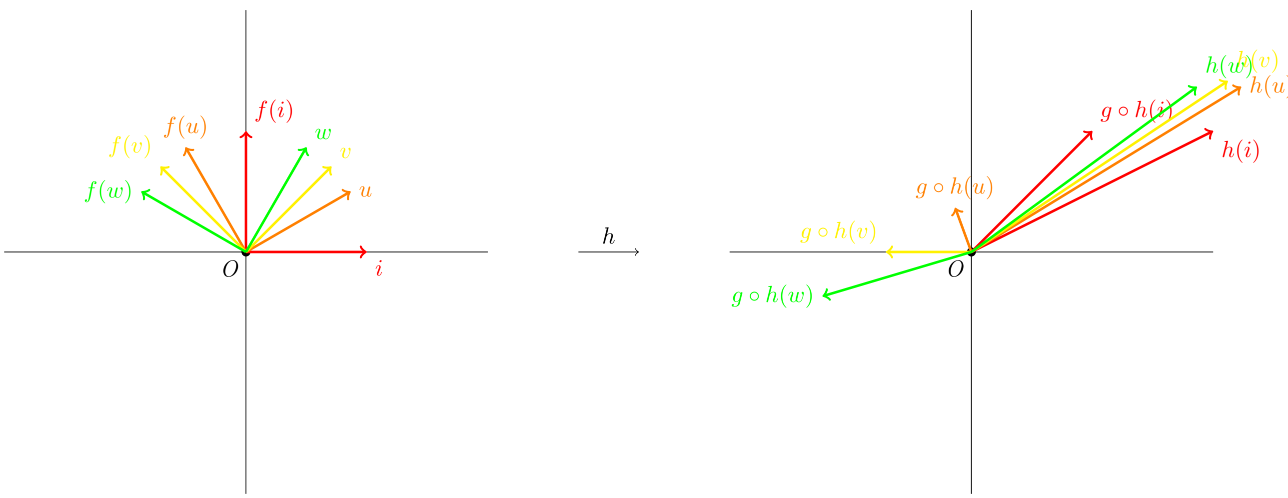
Exemple IV.11.

Prenons $E = \mathbb{R}^2$, muni de sa base canonique $\mathscr{B} = (i, j)$. Soit f la rotation d'angle $\pi/4$ centrée en 0, et h l'endomorphisme dont la matrice dans la base canonique est

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad H^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

La matrice représentant $g = h \circ f \circ h^{-1}$ dans la base canonique est alors :

$$G = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ 2 & -3 \end{pmatrix}.$$



Remarquons que le fait que f et g sont conjugués peut se réécrire :

$$h \circ f = g \circ h.$$

Par exemple, si u est un vecteur et $v = f(u)$ son image par f, alors $h(v)$ est l'image de $h(u)$ par $g = h \circ f \circ h^{-1}$. L'action de g est l'image de l'action de f par h.

Exercice 22. Soit $\mathcal{P}$ le plan euclidien, O un point de ce plan et Δ une droite passant par O. Soient r la rotation d'angle $\Pi/6$ centrée en O et s la symétrie axiale d'axe Δ. Décrire géométriquement les transformations :

1. $s \circ r \circ s^{-1}$ (conjuguée d'une rotation par une symétrie) ;
2. $r \circ s \circ r^{-1}$ (conjuguée d'une symétrie par une rotation).

Exercice 23. Soient f, $g \in L(E)$ et $h \in GL(E)$ tels que $g = h \circ f \circ h^{-1}$. Montrer que :

1. pour tout $m \geq 0$, les endomorphismes f^m et g^m sont conjugués ;
2. pour tout $\lambda \in \mathbb{R}$, les endomorphismes $f - \lambda id$ et $g - \lambda id$ sont conjugués ;
3. $h(\mathrm{Im}(f)) = \mathrm{Im}(g)$;
4. $h(\mathrm{Ker}(f)) = \mathrm{Ker}(g)$.

Propriété IV.12.

f et g sont conjugués si et seulement si on peut leur associer une même matrice, quitte à prendre deux bases différentes pour f et pour g.

f et g sont conjugués si et seulement si pour toute base, les matrices de f et g dans cette base sont semblables.

Preuve : Soit $\mathscr{B}$ une base de E. Soit A la matrice de f dans $\mathscr{B}$.

Supposons que f et g soient conjuguées. Soit $h \in GL(E)$ telle que $.f = h^{-1} \circ g \circ h$. Soit $\mathscr{B}'$ l'image de $\mathscr{B}$ par h, et P la matrice de h dans la base $\mathscr{B}$. Alors P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$. Par conséquent, la matrice associée à $g = h \circ f \circ h^{-1}$ dans $\mathscr{B}$ est PAP^{-1}. Il s'ensuit que :

- ▷ La matrice de g dans la base $\mathscr{B}'$ est $P^{-1}(PAP^{-1})P = A$, qui est la même que la matrice de f dans $\mathscr{B}$..
- ▷ $B = P \circ A \circ P^{-1}$, où B est la matrice de g dans la base $\mathscr{B}$. Donc les matrices de f et g dans $\mathscr{B}$ sont semblables.

Il nous reste maintenant deux réciproques à démontrer. Supposons pour commencer que l'on peut associer une même matrice A à f et g. Soit $\mathscr{B}'$ une base de E telle que la matrice de g dans $\mathscr{B}'$ soit A. Soit h l'endomorphisme qui envoie $\mathscr{B}$ sur $\mathscr{B}'$. Alors $f = h^{-1} \circ g \circ h$, donc f et g sont conjugués. Cela finit la démonstration de la première assertion.

Supposons maintenant que la matrice B de g dans $\mathscr{B}$ soit semblable à celle de f. Soit $P \in GL_n(\mathbb{R})$ telle que $A = P^{-1}BP$. Soit h l'endomorphisme de E donc la matrice dans la base $\mathscr{B}$ soit P. Alors $f = h^{-1} \circ g \circ h$, donc f et g sont conjugués. □

Il est donc équivalent de chercher des matrices semblables ou bien des endomorphismes conjugués.

IV.3 Quelques invariants de similitudes

Définition IV.13.

Un **invariant de similitude** *est une fonction définie sur $M_n(\mathbb{R})$ qui est constante sur les classes de similitudes (c'est-à-dire que deux matrices semblables ont le même invariant).*

Remarque IV.14.

On peut dire de manière équivalente que cette fonction ne dépend que de l'endomorphisme associé, et pas de la base dans laquelle on choisit de le représenter.

Propriété IV.15.

Ce qui suit sont des invariants de similitude :

- ▷ *Le rang.*
- ▷ *La trace.*
- ▷ *Le déterminant.*
- ▷ *Les valeurs propres (quelles sont les valeurs de cet invariant ?).*
- ▷ *Les dimensions des sous-espaces propres.*
- ▷ *Le polynôme caractéristique.*

Preuve : À faire en exercice. Pour chaque invariant, trouver des matrices qui ne sont pas semblables mais qui ont les mêmes invariants. □

Exercice 24. Quels sont les liens entre ces différents invariants ?

Remarque IV.16.

Le polynôme caractéristique est le plus fin. Le déterminant, les valeurs propres et le polynôme caractéristiques sont assez longs à calculer.

IV.4 Exemple de classes de conjugaison

Définition IV.17 (Rappel : Espaces supplémentaires)**.**

Soient F et G deux sous-espaces vectoriels. On dit qu'ils sont supplémentaires si $E = F \oplus G$, c'est à dire si pour tout $u \in E$ il existe un unique couple $(u_F, u_G) \in F \times G$ tel que $u = u_F + u_G$.

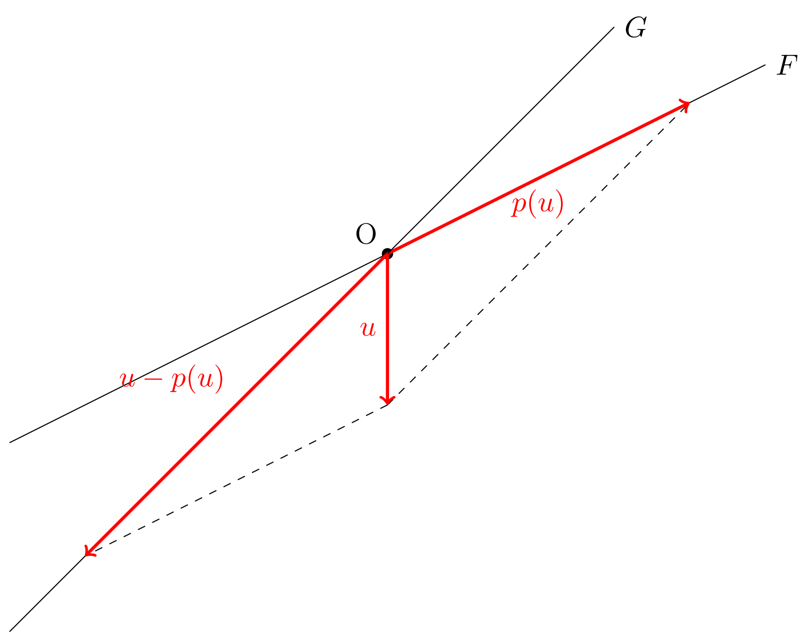
Les projections

Définition IV.18 (Rappel : Projection)**.**

Soit F et G deux sous-espaces vectoriels supplémentaires. On appelle projection sur F parallèlement à G (ou selon G) l'application qui à tout vecteur u dans E associe le vecteur $p(u)$ tel que

- ▷ *$p(u)$ est dans F ;*
- ▷ *$u - p(u)$ est dans G.*

En dimension 2, on retrouve la construction des coordonnées d'un vecteur :



Preuve : Tout d'abord, montrons que p est bien définie. Soit $u \in E$. Comme F et G sont supplémentaires, il existe une unique paire $(u_F, u_G) \in F \times G$ telle que $u = u_F + u_G$. Donc $p(u) = u_F$ convient, et est le seul choix possible.

La décomposition d'un vecteur de E dans des espaces supplémentaires est linéaire : pour tous u, $v \in E$ et tous λ, $\mu \in \mathbb{R}$,

$$(\lambda u + \mu v)_F = \lambda u_F + \mu v_F \quad (\lambda u + \mu v)_G = \lambda u_G + \mu v_G.$$

Comme $p(u) = u_F$, l'application p est elle aussi linéaire. $\square$

Propriété IV.19.

Soient F, G deux espaces supplémentaires dans E et p la projection sur F parallèlement à G. L'ensemble des vecteurs fixes de p est égal à F, et le noyau de p est égal à G.

Preuve : Soit $u \in F$. Alors $u = u + 0$, où $u \in F$ et $0 \in G$. Donc $p(u) = u$. Réciproquement, soit $u \in E$ et supposons que $p(u) = u$. Comme $p(u) \in F$, on sait que $u \in F$. Cela montre la première partie de la propriété.

Soit $u \in G$. Alors $u = 0 + u$, où $0 \in F$ et $u \in G$. Donc $p(u) = 0$. Réciproquement, soit $u \in E$ et supposons que $p(u) = 0$. Comme $u - p(u) \in G$, on sait que $u \in G$, ce qui termine la démonstration. $\square$

Propriété IV.20.

Soit $p \in L(E)$. Alors $p \circ p = p$ si et seulement si p est une projection. Dans ce cas, $E = Ker(p) \oplus Ker(p - id)$, et p est la projection sur $Ker(p - id)$ parallèlement à $Ker(p)$.

Preuve : Soit p la projection sur F parallèlement à G. Pour tout $u \in E$, on a $p(u) \in F$. Donc $p(u)$ est fixé par p, donc $p(p(u)) = p(u)$.

On peut aussi le montrer avec les matrices. Soient $\mathscr{B}_F$ et $\mathscr{B}_G$ des bases de F et G respectivement. Alors $\mathscr{B} = \mathscr{B}_F \cup \mathscr{B}_G$ est une base de E. La matrice associée à p dans $\mathscr{B}$ est

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

On vérifie alors que $A^2 = A$.

Réciproquement, soit p un endomorphisme tel que $p \circ p = p$. Posons $F = \mathrm{Im}(p)$ et $g := \mathrm{Ker}(p)$. Alors :

1. D'une part, si $u \in F$, alors il existe $v \in E$ tel que $u = p(v)$. Donc $p(u) = p(p(v)) = p(v) = u$. Donc $F \subset \{u \in E : \ p(u) = u\}$.
2. D'autre part, soit $u \in E$ tel que $p(u) = u$. Alors $u \in F$.

Donc $F = \mathrm{Im}(p) = \{u \in E : \ p(u) = u\} = \mathrm{Ker}(p - id)$.

Par le théorème du rang, $\dim(E) = \dim(F) + \dim(G)$. De plus, si $u \in F \cap G$, alors $p(u) = u = 0$, donc $F \cap G = \{0\}$. Donc F et G sont supplémentaires dans E. Enfin, en écrivant $u = u_F + u_G$ avec $(u_F, u_G) \in F \times G$, on remarque que $p(u) = u_F$, ce qui montre que p est bien la projection sur $F = \mathrm{Ker}(p - id)$ parallèlement à $G = \mathrm{Ker}(p)$. $\square$

Les deux propositions suivantes permettent de décrire les classes de similitudes de projections.

Proposition IV.21.

Soit p la projection sur F parallèlement à G. Soit g un endomorphisme conjugué à p :

$$g = h \circ p \circ h^{-1}.$$

Alors g est la projection sur $h(F)$ parallèlement à $h(G)$.

PREUVE : Soit $u \in E$. L'endomorphisme h étant inversible, les espaces $h(F)$ et $h(G)$ sont supplémentaires dans E.

Soit $u \in h(F)$. Alors $h^{-1}(u) \in F$, donc $p(h^{-1}(u)) = h^{-1}(u)$, donc $g(u) = h(h^{-1}(u)) = u$.

Soit $u \in h(G)$. Alors $h^{-1}(u) \in G$, donc $p(h^{-1}(u)) = 0$, donc $g(u) = h(0) = 0$.

Donc u est bien la projection sur $h(F)$ parallèlement à $h(G)$. □

Remarque IV.22.

On peut aussi utiliser la propriété précédente. Si $g = h \circ p \circ h^{-1}$, alors

$$g \circ g = h \circ p \circ h^{-1} \circ h \circ p \circ h^{-1} = h \circ p \circ p \circ h^{-1} = h \circ p \circ h^{-1} = g,$$

donc g est la projection sur $Im(g)$ parallèlement à $Ker(g)$. Enfin, $Im(g) = h(Im(p)) = h(F)$ et $Ker(g) = h(Ker(p)) = h(G)$.

En particulier, si p est une projection, alors sa classe de similitude est incluse dans l'ensemble des projections de même rang (i.e. sur des sous-espaces de même dimension).

Proposition IV.23.

Soient p, q deux projections de E. Les endomorphismes p et q sont conjugués si et seulement si $rg(p) = rg(q)$.

PREUVE : Le sens direct est une conséquence du fait que le rang est un invariant de similitude.

Supposons que $\mathrm{rg}(p) =' (q)$. Soient $\mathscr{B}_F$ une base de $\mathrm{Im}(p)$ et $\mathscr{B}_G$ une base de $\mathrm{Ker}(p)$. Les espaces $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ étant supplémentaires, $\mathscr{B} = \mathscr{B}_F \cup \mathscr{B}_G$ est une base de E.

Soient $\mathscr{B}'_F$ une base de $\mathrm{Im}(q)$ et $\mathscr{B}'_G$ une base de $\mathrm{Ker}(q)$. Comme p et q sont de même rang, $\mathscr{B}_F$ et $\mathscr{B}'_F$ sont de même cardinalité, de même que $\mathscr{B}_G$ et $\mathscr{B}'_G$. Il existe donc une unique application linéaire h qui envoie les éléments de $\mathscr{B}_F$ sur ceux de $\mathscr{B}'_F$, et les $\mathscr{B}_G$ sur ceux de $\mathscr{B}'_G$.

Soit $u \in F$. Alors $h(u) \in \mathrm{Im}(q)$ et $u \in \mathrm{Im}(p)$, donc $q(h(u)) = h(u) = h(p(u))$. De même, pour tout $u \in G$, on a $q \circ h(u) = 0 = h(0) = h(p(u))$. Par linéarité de h, on a finalement $q(h(u)) = h(p(u))$ pour tout u, donc $q = h \circ p \circ h^{-1}$. □

Conséquence IV.24.

Soit p une projection. La classe de similitude de p est l'ensemble des projections de même rang que p.

Les classes de similitudes sont donc beaucoup plus petites que les classes d'équivalence : non seulement la relation de similitude préserve le rang, mais aussi le fait d'être une projection !

Les symétries

La présentation de symétrie suit celle des projections. Les démonstrations sont essentiellement identiques, et ne sont pas reproduites.

Définition IV.25. *Soit F et G deux sous-espaces vectoriels supplémentaires. On appelle* **symétrie** *d'axe F parallèlement à G l'application s qui à tout vecteur $u \in E$ associe le vecteur $s(u) = u_F - u_G$. En particulier, s est bien un endomorphisme de E.*

On peut donner une définition équivalente, qui s'appelle aussi une propriété caractéristique :

Définition IV.26. *Soit F et G deux sous-espaces vectoriels supplémentaires. On appelle symétrie d'axe F parallèlement à G l'application qui à tout vecteur u dans E associe le vecteur $s(u)$ tel que*

- ▷ *$u + s(u)$ est un vecteur de F ;*
- ▷ *$u - s(u)$ est dans G.*

Propriété IV.27. *1. Les symétries sont bijectives.*

2. L'ensemble des vecteurs fixes d'une symétrie d'axe F parallèlement à G est égal à F.

Propriété IV.28. *Soit $s \in L(E)$. Alors $s \circ s = id$ si et seulement si s est une symétrie. Dans ce cas, $E = Ker(s - id) \oplus Ker(s + id)$, et s est la symétrie d'axe $Ker(s - id)$ parallèlement à $Ker(s + id)$.*

Proposition IV.29.

Soit s la symétrie d'axe F parallèlement à G. Soit g un endomorphisme conjugué à p :

$$g = h \circ s \circ h^{-1}.$$

Alors g est la symétrie d'axe $h(F)$ parallèlement à $h(G)$.

Exemple IV.30.

Reprenons l'un des exercices précédents. Soit $\mathcal{P}$ le plan euclidien, O un point de ce plan et Δ une droite passant par O. Soient r la rotation d'angle $\Pi/6$ centrée en O et s la symétrie axiale d'axe Δ.

Alors s est la symétrie d'axe Δ parallèlement à Δ', où Δ' est la droite orthogonale à Δ passant par O. Par ce qui précède, $r \circ s \circ r^{-1}$ est la symétrie d'axe $r(\Delta)$ parallèlement à $r(\Delta')$. Commes les images par des rotations de droites orthogonales sont orthogonales, la symétrie d'axe $r(\Delta)$ parallèlement à $r(\Delta')$ n'est autre que la symétrie axiale d'axe $r(\Delta)$.

Proposition IV.31.

Soient s, t deux symétries de E. Les endomorphismes s et t sont conjugués si et seulement si les dimensions de leurs axes sont égales :

$$\dim(Ker(s - id)) = \dim(Ker(t - id)).$$

Conséquence IV.32.

Soit s une symétrie. La classe de similitude de s est l'ensemble des symétries dont l'axe a même dimension que l'axe de s.

IV.5 Sous-espaces stables

Revenons au cas général.

Propriété IV.33.

Toute application linéaire préserve les sous-espaces vectoriels.

En particulier, l'image d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel de dimension au plus celle du sous-espace vectoriel de départ.

PREUVE : Soient E, F deux espaces vectoriels et $f : E \to F$. Soit L un sous-espace vectoriel de E. Alors :

- ▷ Soit $u \in f(L)$. Soit $v \in L$ tel que $f(v) = u$. Soit $\lambda \in \mathbb{R}$. Alors $\lambda v \in L$, donc $\lambda u = \lambda f(v) = f(\lambda v) \in f(L)$.
- ▷ Soient u_1, $u_2 \in f(L)$. Soient v_1, $v_2 \in L$ tels que $f(v_1) = u_1$ et $f(v_2) = u_2$. Alors $v_1 + v_2 \in L$, donc $u_1 + u_2 = f(v_1) + f(v_2) = f(v_1 + v_2) \in f(L)$.

Donc $f(L)$ est un sous-espace vectoriel de F.

Par le théorème du rang, $\dim(f(L)) \leq \dim(\mathrm{Im}(E)) \leq \dim(E)$. □

Exercice 25. Donner la nature de l'image d'une droite, d'un plan. Que peut-on dire si f est bijective ? Pourquoi ces propriétés sont-elles invariantes par conjugaison ?

Exercice 26. **Réciproquement, si $\dim(E)$, $\dim(F) \geq 2$, on pourrait démontrer que toute application de E dans F qui préserve les sous-espace vectoriel est une application linéaire. (C'est difficile)

Définition IV.34 (Ensembles fixes, ensembles stables)**.**

- ▷ *Un ensemble A de E est dit* **fixe** *par f si pour tout $u \in A$, on a $f(u) = u$.*
- ▷ *Un ensemble A de E est dit* **stable** *par f s'il contient son image par f, c'est-à-dire si $f(A) \subset A$.*

Propriété IV.35.

Soit $f \in L(E)$, et u un vecteur non nul.

- ▷ *u est un vecteur propre associé à f si et seulement si la droite vectorielle $\mathbb{R}u$ est stable par f.*
- ▷ *u est un vecteur fixe pour f si et seulement si u est un vecteur de l'espace propre $Ker(f - id)$ associé à 1.*

PREUVE : Commençons par le premier énoncé. Supposons que u est un vecteur propre de f, et soit λ la valeur propre associée. Soit $v \in \mathbb{R}u$. Alors il existe $\mu \in \mathbb{R}$ tel que $v = \mu u$. Alors

$$f(v) = f(\mu u) = \mu f(u) = \mu\lambda u \in \mathbb{R}u.$$

Donc $f(\mathbb{R}u) \subset \mathbb{R}u$: la droite vectorielle $\mathbb{R}u$ est stable par f.

Réciproquement, supposons que $\mathbb{R}u$ est stable par f. Alors $f(u) \in \mathbb{R}u$, donc il existe $\lambda \in \mathbb{R}$ tel que $f(u) = \lambda u$.

Passons au second énoncé. u est un vecteur propre de l'espace propre $\mathrm{Ker}(f - id)$ associé à 1 si et seulement si $f(u) - u = 0$, c'est-à-dire si et seulement si $f(u) = u$, donc si et seulment si u est un vecteur fixe pour f. □

Exercice 27. Donner des exemples d'applications linéaires admettant un ou plusieurs vecteurs fixes ; n'admettant pas de vecteurs fixes mais des ensembles stables.

Notation : Soient $f \in L(E)$ et $\lambda \in \mathbb{R}$ une valeur propre de f. On note $E_\lambda = \mathrm{Ker}(f - \lambda id)$ le sous-espace propre associé à λ.

Proposition IV.36.

La dimension des sous-espaces propres est un invariant de conjugaison.

Preuve : Soient f un endomorphisme et $g = h \circ f \circ h^{-1}$ un endomorphisme conjugué à f. Soit $\lambda \in \mathbb{R}$ une valeur propre de f.

Pour tout $u \in E_\lambda$,

$$g(h(u)) = h(f(u)) = h(\lambda u) = \lambda h(u).$$

Donc $h(u) \in \mathrm{Ker}(g - \lambda id)$. Par conséquent, $h(\mathrm{Ker}(f - \lambda id)) \subset \mathrm{Ker}(g - \lambda id)$. Réciproquement, en écrivant $f = h^{-1} \circ g \circ h$, on montre que $h^{-1}(\mathrm{Ker}(g-\lambda id)) \subset \mathrm{Ker}(f-\lambda id)$, et donc $(\mathrm{Ker}(g-\lambda id)) \subset h(\mathrm{Ker}(f - \lambda id))$. Finalement, on a donc démontré quelconque

$$h(\mathrm{Ker}(f - \lambda id)) = \mathrm{Ker}(g - \lambda id).$$

Or, h étant inversible, elle préserve la dimension des sous-espaces vectoriels. Donc $\dim(\mathrm{Ker}(f - \lambda id)) = \dim(\mathrm{Ker}(g - \lambda id))$. □

IV.6 Cas des endomorphismes diagonalisables

Les sous-espaces propres ne sont pas des invariants, mais leur dimension oui. Par conséquent, lorsqu'on peut décomposer l'espace selon les sous-espaces propres, les classes de conjugaison seront plus simples à décrire : c'est le cas des endomorphismes diagonalisables.

Définition IV.37 (Rappel : sous-espaces vectoriels en somme directe)**.**

Soient E_1, E_2, $\cdots$, E_k k sous-espaces vectoriels de E. On dit que ces sous-espaces sont **en somme directe** *si la concaténation des bases de $E_1, \cdots, E_k$ forme une base de $E_1 + E_2 + \cdots + E_k$. On note alors $E_1 \oplus E_2 \oplus \cdots \oplus E_k$ ce sous-espace vectoriel de E.*

Remarque IV.38.

On peut dire de manière équivalente que pour tout $u \in E_1 \oplus E_2 \oplus \cdots \oplus E_k$, il existe une unique décomposition $u = e_1 + e_2 + \cdots e_k$ telle que $e_i \in E_i$ pour tout $i \in \{1, \cdots, k\}$.

Exercice 28. Prenons $E = \mathbb{R}^2$. Trouver trois sous-espaces vectoriel E_1, E_2, E_3 tels que $E_1 \cap E_2 = E_2 \cap E_3 = E_1 \cap E_3 = \{0\}$, mais E_1, E_2, E_3 ne sont pas en somme directe.

Exercice 29. Savoir construire géométriquement ces décompositions en dimension 2 et 3.

Propriété IV.39.

Les sous-espaces propres sont toujours en somme directe.

Preuve : Soit f un endomorphisme de E. Soient λ_1, $\cdots$, λ_k k valeurs propres distinctes de f. Nous voulons démontrer que, pour toute décomposition $0 = u_1 + \cdots + u_k$ de 0 dans $E_{\lambda_1} + \cdots + E_{\lambda_k}$, tous les u_i sont nuls. Nous allons procéder par récurrence sur k.

Initialisation : si $k = 1$, le résultat est immédiat (si $u_1 = 0$, alors $u_1 = 0$).

Hérédité : supposons la propriété vraie au rang $k \geq 1$. Soit $0 = u_1 + \cdots + u_{k+1}$ une décomposition de 0 dans $E_{\lambda_1} + \cdots + E_{\lambda_{k+1}}$. Alors

$$0 = f(0) = \sum_{i=1}^{k+1} f(u_i) = \sum_{i=1}^{k+1} \lambda_i u_i.$$

En soustrayant l'égalité $\lambda_{k+1} 0 = \lambda_{k+1} u_1 + \cdots + \lambda_{k+1} u_{k+1}$, on obtient

$$0 = \sum_{i=1}^{k+1} (\lambda_i - \lambda_{k+1}) u_i = \sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) u_i,$$

car le dernier coefficient $(\lambda_{k+1} - \lambda_{k+1})$ est nul. Ensuite, par la propriété d'hérédité, les vecteurs $(\lambda_i - \lambda_{k+1}) u_i$ sont nuls pour $1 \leq i \leq k$. Or $\lambda_i \neq \lambda_{k+1}$ pour $i \leq k$, donc les vecteurs u_i sont nuls pour tout $1 \leq i \leq k$. Finalement, il reste $u_{k+1} = 0$, ce qui achève cette démonstration. □

Définition IV.40 (Rappel : diagonalisabilité)**.**

Soit $f \in L(E)$. L'endomorphisme f est **diagonalisable** *sur $\mathbb{R}$ (respectivement sur $\mathbb{C}$) si on peut trouver k réels (resp. complexes) $\lambda_1, \cdots, \lambda_k$ tels que*

$$E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}.$$

Propriété IV.41.

$A \in \mathscr{M}_n(\mathbb{R})$ est diagonalisable sur $\mathbb{R}$ s'il existe une matrice $D \in \mathscr{M}_n(\mathbb{R})$ diagonale qui est semblable à A.

Preuve : Soient $\mathscr{B}_1, \cdots, \mathscr{B}_k$ des bases de $E_{\lambda_1}, \cdots, E_{\lambda_k}$ respectivement. Comme ces espaces sont en somme directe et leur somme est E, la famille $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_k$ est une base de E. Soit P la matrice de passage de la base canonique à $\mathscr{B}$. Par définition des sous-espaces propres, la matrice $P^{-1}AP$ est diagonale :

$$P^{-1}AP = \begin{pmatrix} \lambda_1 Id_{K_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k Id_{K_k} \end{pmatrix},$$

où $K_i = \dim(E_{\lambda_i})$. □

Remarque IV.42.

Comment se traduit géométriquement cette propriété ?

Propriété IV.43.

Les classes de similitudes pour les endomorphismes diagonalisables sont caractérisées par l'ensemble de leurs valeurs propres et la dimension de l'espace propre associé.

Preuve : Soient A, B deux matrices diagonalisables ayant les mêmes valeurs propres et les mêmes dimensions d'espaces propres associés. Alors A et B sont semblables respectivement à des matrices C, D diagonales. Il suffit de montrer que C et D sont semblables ; or, on peut conjuguer C à D en permutant les vecteurs de la base dans laquelle est représentée C.

Réciproquement, soient A, B deux matrices diagonalisables semblables. Alors $\dim(\mathrm{Ker}(A - \lambda Id)) = \dim(\mathrm{Ker}(B - \lambda Id))$ pour tout $\lambda \in \mathbb{R}$, donc A et B ont les mêmes valeurs propres et les mêmes dimensions d'espaces propres associés. □

Malheureusement, cette caractérisation n'est pas valable pour les matrices non diagonalisables.

Exercice 30. Donner des exemples de matrices ayant les mêmes valeurs propres et mêmes dimensions d'espaces propres, mais non semblables.

Exercice 31. Soit $k \in \mathbb{N}^*$. Montrer que si f est diagonalisable alors f^k est diagonalisable. Que pensez-vous de la réciproque ?

V Endomorphismes nilpotents, blocs de Jordan et tableaux de Young

Définition V.1.
Soient $f \in L(E)$ et $p \in \mathbb{N}$. On dit que f est **nilpotent** *d'indice p si $f^{p-1} \neq 0$ et $f^p = 0$.*

Exercice 32.

1. Montrer que $p \leq n^2$.
2. * Montrer qu'en fait $p \leq n$.

Remarque V.2.
Dire que $f^p = 0$ est équivalent à $Ker(f^p) = E$.

Exercice 33. Petites propriétés à chercher. Soit $f \in L(E)$ quelconque.

1. Démontrer que $\mathrm{Ker}(f) \subset \mathrm{Ker}(f^2)$.
2. Démontrer que si $p \in \mathbb{N}$ vérifie $\mathrm{Ker}(f^p) = \mathrm{Ker}(f^{p+1})$, alors $\mathrm{Ker}(f^k) = \mathrm{Ker}(f^p)$ pour tout $k \geq p$.
3. Démontrer que si $v \in \mathrm{Ker}(f^p)$ mais $v \notin \mathrm{Ker}(f^{p-1})$, alors pour tout $k \in \{1, \cdots, p-1\}$, on a $f^k(v) \in \mathrm{Ker}(f^{p-k})$ mais $f^k(v) \notin \mathrm{Ker}(f^{p-k-1})$.
4. Soit L est un sous-espace vectoriel tel que $\mathrm{Ker}(f) \oplus L \subset F$. Montrer que si $(v_1, v_2, \cdots, v_k)$ est une famille libre de L, alors $(f(v_1), f(v_2), \cdots, f(v_k))$ est encore une famille libre.

Preuve : Premier point : soit $u \in \mathrm{Ker}(f)$. Alors $f^2(u) = f(f(u)) = f(0) = 0$. Donc $u \in \mathrm{Ker}(f^2)$. On a bien montré que $\mathrm{Ker}(f) \subset \mathrm{Ker}(f^2)$.

Second point : il suffit de montrer que $\mathrm{Ker}(f^{k+1}) = \mathrm{Ker}(f^k)$ pour tout $k \geq p$. Soit $k \geq p$. Alors $\mathrm{Ker}(f^k) \subset \mathrm{Ker}(f^{k+1})$ par le même argument que précédemment ; nous allons montrer l'inclusion réciproque. Soit $u \in \mathrm{Ker}(f^{k+1})$. Alors $f^{p+1}(f^{k-p}(u)) = 0$, donc $f^{k-p}(u) \in \mathrm{Ker}(f^{p+1}) = \mathrm{Ker}(f^p)$, donc $0 = f^p(f^{k-p}(u)) = f^k(u)$. Ainsi, $u \in \mathrm{Ker}(f^k)$, ce que l'on voulait démontrer.

Troisième point : soit $v \in \mathrm{Ker}(f^p)$ tel que $v \notin \mathrm{Ker}(f^{p-1})$. Soit $k \in \{1, \cdots, p-1\}$. D'une part, $0 = f^p(v) = f^{p-k}(f^k(v))$, donc $f^k(v) \in \mathrm{Ker}(f^{p-k})$. D'autre part, $0 \neq f^{p-1}(v) = f^{p-k-1}(f^k(v))$, donc $f^k(v) \notin \mathrm{Ker}(f^{p-k-1})$.

Quatrième point : Soit L un tel sous-espace vectoriel et $(v_1, v_2, \cdots, v_k)$ est une famille libre de L. Soit $(\lambda_1, \cdot, \lambda_k)$ tel que

$$\sum_{i=1}^{k} \lambda_i f(v_i) = 0.$$

Il suffit de montrer que tous les λ_i sont nuls. Mais on remarque que $\sum_{i=1}^{k} \lambda_i v_i \in \mathrm{Ker}(f)$ par la condition sur $(\lambda_1, \cdot, \lambda_k)$, et $\sum_{i=1}^{k} \lambda_i v_i \in L$ car tous les v_i sont dans L. Donc $\sum_{i=1}^{k} \lambda_i v_i \in \mathrm{Ker}(f) \cap L = \{0\}$, et $\sum_{i=1}^{k} \lambda_i v_i = 0$. Par liberté de cette famille de vecteurs, tous les λ_i sont nuls. □

V.1 Le théorème des noyaux itérés

Théorème 3 (Noyaux itérés)**.**
Soit $f \in L(E)$. Alors :

1. *La suite des noyaux est croissante : pour tout $j \in \mathbb{N}$, on a $Ker(f^j) \subset Ker(f^{j+1})$.*
2. *La suite se stabilise : il existe $p \leq n$ tel que $Ker(f^p) = Ker(f^{p+1})$ et :*
   (a) *Pour tout $j < p$, $Ker(f^j) \subsetneq Ker(f^{j+1})$,*
   (b) *Pour tout $j \geq p$, $Ker(f^j) = Ker(f^{j+1})$.*
3. *La suite est de moins en moins strictement croissante : pour tout $j \geq 0$,*

$$\dim(Ker(f^{j+2})) - \dim(Ker(f^{j+1})) \leq \dim(Ker(f^{j+1})) - \dim(Ker(f^j)).$$

Définition V.3.
Avec les notations du théorème précédent, $Ker(f^p)$ s'appelle le **sous-espace caractéristique** *associé à (la valeur propre) 0.*

Preuve : Idées générales de la preuve du théorème :

1. Soit $j \in \mathbb{N}$ et $v \in \mathrm{Ker}(f^j)$. Alors $f^{j+1}(v) = f(f^j(v)) = f(0) = 0$, donc $v \in \mathrm{Ker}(f^{j+1})$.
2. On montre d'abord l'existence de p, sans s'occuper des conditions supplémentaires : la suite des noyaux $(\mathrm{Ker}(f^j))_{j \geq 0}$ est croissante, donc la suite des dimensions des noyaux est elle aussi croissante. La suite des dimensions est donc une suite croissante d'entiers majorée par $\dim(E) = n$. Elle est donc constante à partir d'un certain rang : il existe un entier p' tel que $\dim(\mathrm{Ker}(f^{p'})) = \dim(\mathrm{Ker}(f^{p'+1}))$.

   Soit p le plus petit entier tel que $\dim(\mathrm{Ker}(f^p)) = \dim(\mathrm{Ker}(f^{p+1}))$. Alors la suite $(\dim(\mathrm{Ker}(f^j)))_{0 \leq j \leq p}$ est une suite d'entiers strictement croissante. En particulier [13], $\dim(\mathrm{Ker}(f^j)) \geq \dim(\mathrm{Ker}(f^0)) + j = j$ pour tout entier $j \leq p$. Il suit que $n \geq \dim(\mathrm{Ker}(f^p)) \geq p$. On obtient au passage la première propriété.

   Pour montrer la seconde propriété, on choisit $j \geq p$, et on se donne $v \in \mathrm{Ker}(f^{j+1})$. Il suffit de démontrer que $v \in \mathrm{Ker}(f^j)$. Or $0 = f^{j+1}(v) = f^{p+1}(f^{j-p}(v))$, donc $f^{j-p}(v) \in \mathrm{Ker}(f^{p+1}) = \mathrm{Ker}(f^p)$, donc $0 = f^p(f^{j-p}(v)) = f^j(p)$.

13. Il faudrait en écrire la démonstration par récurrence.

3. Cette partie est la plus délicate. Pour tout $j < p$, on sait que $\mathrm{Ker}(f^j) \subsetneq \mathrm{Ker}(f^{j+1})$. Soit donc L_j un supplémentaire de $\mathrm{Ker}(f^j)$ dans $\mathrm{Ker}(f^{j+1})$. En particulier,

$$\dim(L_j) = \dim(\mathrm{Ker}(f^{j+1})) - \dim(\mathrm{Ker}(f^j)).$$

Soit $j \geq 1$, et soit $v \in L_j$ non nul. Alors $v \in \mathrm{Ker}(f^{j+1}) \setminus \mathrm{Ker}(f^j)$, donc $f(v) \in \mathrm{Ker}(f^j) \setminus \mathrm{Ker}(f^{j-1})$. Par conséquent, $f(L_j) \subset \mathrm{Ker}(f^j)$ et $f(L_j) \cap \mathrm{Ker}(f^{j-1}) = \{0\}$, donc $f(L_j)$ et $\mathrm{Ker}(f^{j-1})$ sont en somme directe dans $\mathrm{Ker}(f^j)$. Ainsi,

$$\dim(f(L_j)) = \dim(f(L_j) \oplus \mathrm{Ker}(f^{j-1})) - \dim(\mathrm{Ker}(f^{j-1})) \leq \dim(\mathrm{Ker}(f^j)) - \dim(\mathrm{Ker}(f^{j-1})).$$

Enfin, par l'exercice précédent, l'image d'une base de L_j est une famille libre, donc

$$\dim(f(L_j)) = \dim(L_j) = \dim(\mathrm{Ker}(f^{j+1})) - \dim(\mathrm{Ker}(f^j)),$$

ce qui conclut la preuve.

□

Remarque V.4.

L'orbite d'un vecteur du noyau le plus grand décrit successivement tous les noyaux de niveaux inférieurs : pour tous $0 \leq k \leq j < p$, si $v \in Ker(f^{j+1}) \setminus Ker(f^j)$, alors $f^k(v) \in Ker(f^{j+1-k}) \setminus Ker(f^{j-k})$. En particulier, la famille $(v, f(v), \cdot, f^j(v))$ est libre.

La démonstration précédente permet de comprendre la structure des noyaux emboités, et de savoir construire des bases efficaces pour décrire f l'action de f dans $Ker(f^p)$. On commence par choisir L_{p-1} et une base $\mathscr{B}_{p-1}$ de celui-ci. Alors chaque vecteur de $\mathscr{B}_{p-1}$ fournit une famille libre de la forme $(v, f(v), \cdot, f^{p-1}(v))$, et chacune de ces familles sont indépendantes.

Les vecteurs $f(v)$ pour $v \in \mathscr{B}_{p-1}$ sont libres et dans $Ker(f^{p-1}) \setminus Ker(f^{p-2})$. On complète le sous-espace engendré par ceux-ci en L_{p-2}, et une on se donne une famille libre $\mathscr{B}_{p-2}$ complétant l'image de $\mathscr{B}_{p-1}$ en une base de L_{p-2}. Alors chaque vecteur de $\mathscr{B}_{p-2}$ fournit une nouvelle famille libre de la forme $(v, f(v), \cdot, f^{p-2}(v))$.

On fabrique ainsi des "chaines" de vecteurs formant finalement une base de $Ker(f^p)$. Si f est nilpotente, on a construit une base de $E = Ker(f^p)$ dans laquelle la matrice de f s'écrit avec des 1 ou des 0 sous la diagonale, et des 0 partout ailleurs : c'est ce qu'on appelle des matrices de Jordan.

V.2 Forme de Jordan et tableaux de Young des endomorphismes nilpotents

Lorsque f est nilpotent, le théorème des noyaux emboités se traduit par le fait que le sous-espace caractéristique est égal à E tout entier. L'indice de nilpotence est alors égal à m avec les notations du théorème. La preuve du théorème des noyaux emboités permet de construire des bases de E dans lesquelles la matrice s'écrit très simplement : en effet, il s'agit de suivre les orbites des vecteurs des noyaux, en partant des vecteurs dont l'orbite est la plus grande.

Exemple V.5 (Cas d'un endomorphisme d'indice de nilpotence maximal)**.**

Soit $f \in L(E)$ d'indice $n = \dim(E)$. On peut trouver une base de E de la forme $(v, f(v), \cdots, f^{n-1}(v))$. La matrice associée à f dans cette base est :

$$J_n = \begin{pmatrix} 0 & 0 & \cdots & & 0 \\ 1 & 0 & & & \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

Réciproquement, tout $f \in L(E)$ admettant une matrice de cette forme est nilpotent d'indice n.
Conclusion : *Les endomorphismes nilpotents d'indice n forment une seule classe de conjugaison. En termes de matrices, on peut dire que les matrices nilpotentes d'indice n sont toutes dans la même classe de similitude, qui est celle de J_n.*

Exercice 34. Montrer qu'un endomorphisme représenté par J_n est nilpotent d'indice n.

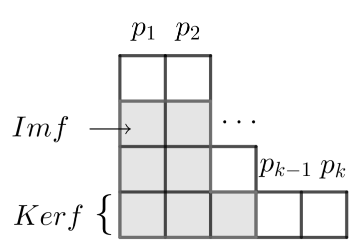
Définition V.6.
Soit $n \in \mathbb{N}^*$. La matrice J_n définie précédemment s'appelle aussi **bloc de Jordan** *de taille (ou d'indice) n.*

Propriété V.7 (Forme de Jordan d'un endomorphisme nilpotent)**.**
Soit $f \in L(E)$ nilpotent d'indice p, alors il existe une décomposition entière de n sous la forme $n = \sum_{i=1}^{k} p_i$ avec $(p_i)_{i\in\{1;k\}}$ décroissante et $p_1 = p$, de sorte que f admette une matrice de la forme :

$$\begin{pmatrix} J_{p_1} & 0 & \cdots & 0 \\ 0 & J_{p_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{p_k} \end{pmatrix}$$

Cette matrice s'appelle **la matrice de Jordan** *associée à f.*

Propriété V.8 (Tableau de Young associé à un endomorphisme nilpotent)**.**
Soit $f \in L(E)$ nilpotent d'indice p. On peut établir une bijection entre les matrices de Jordan des endomorphismes nilpotents et les décompositions entières de n sous la forme $n = \sum_{i=1}^{k} p_i$, avec $k \in \{1, \cdots, n\}$, et $(p_i)_{i\in\{1;k\}}$ un k-uplet décroissant d'entiers non nuls. Cette décomposition peut se représenter à l'aide d'un tableau de Young :



Théorème 4 (Classification des endomorphismes nilpotents).

Soit $f \in L(E)$ nilpotent d'indice p. La classe de conjugaison de f est caractérisée par son tableau de Young associé. Autrement dit, les tableaux de Young constituent un invariant complet pour les classes de conjugaison des endomorphismes nilpotents : deux endormorphismes nilpotents sont conjugués si et seulement s'ils ont le même tableau de Young.

PREUVE : Soit f un endomorphisme nilpotent. Son tableau de Young, et donc la matrice de Jordan associée, est caractérisée par la suite $(\dim(\mathrm{Ker}(f^k)))_{0\leq k\leq n-1}$. Il n'existe donc qu'un plus une matrice de Jordan dans la classe de similitude de f. De plus, par la construction précédente, il en existe au moins une. On a donc bien défini une fonction de l'ensemble des endormorphismes nilpotents à valeur dans les tableaux de Young.

De plus, si deux endomorphismes nilpotents sont conjugués, alors ils ont les mêmes représentations matricielles, donc peuvent être représentés par le même bloc de Jordan. Leur tableaux de Young associés sont les mêmes.

Réciproquement, si deux endomorphismes nilpotent ont le même tableau de Young, alors leurs classes de similitudes s'intersectent (car elles contiennent toutes deux le même bloc de Jordan), donc sont les mêmes. Par conséquent, ces deux endomorphismes sont conjugués. □

Remarque V.9.

Le nombre de classes d'équivalences d'endomorphismes nilpotents d'un espace vectoriel de dimension n est donc égal au nombre de tableaux de Young de taille n, donc au nombre de partitions $p(n)$ de n, c'est-à-dire au nombre de décompositions $n = \sum_{i=1}^{k} p_i$ où les p_i sont décroissants et non nuls.

Ce nombre de partition a été extensivement étudié. S'il n'y a pas de formule élémentaire explicite connue, il existe des formules de récurrences efficaces permettant de le calculer ; par exemple, $p(10) = 42$, donc il y a 42 classes de similitude d'endomorphismes nilpotent de $\mathbb{R}^{10}$. De plus, il existe des estimations de l'ordre de grandeur de $p(n)$. Par exemple, Hardy et Ramanujan ont montré en 1918 que

$$p(n) \sim_{n\to+\infty} \frac{1}{4\sqrt{3}n} e^{\frac{\pi\sqrt{2}}{\sqrt{3}}\sqrt{n}}.$$

Si le nombre de classes de conjugaison croît moins vite qu'une exponentielle, il croît tout de même très vite. Les premières valeurs sont répertoriées par l'On-line Encyclopedy of Integer Sequences (OEIS), suite A000041.

VI Décomposition de Dunford ou forme de Jordan-Chevalley.

Nous traitons maintenant le cas général d'un endomorphisme f de E, ou E est un $\mathbb{C}$-espace vectoriel. L'idée repose sur la décomposition de E en somme directe de sous-espace vectoriels stables associés aux valeurs propres. Nous avons vu qu'en général, il n'y a pas de base de vecteurs propres ; il faut donc compléter les sous-espaces propres.

VI.1 Trigonalisation

S'il n'est pas possible de diagonaliser, en revanche, on peut quand même simplifier l'écriture de la matrice représentant f : cette simplification se traduit par l'existence d'une base dans laquelle la matrice associée à f est triangulaire.

Attention : cette trigonalisation n'est possible que [14] dans $\mathbb{C}$.

Exercice 35. Trouver une matrice 2×2 non trigonalisable dans $\mathbb{R}$.

Propriété VI.1. *Pour tous $n \in \mathbb{N}^*$ et $A \in \mathscr{M}_n(\mathbb{C})$, il existe une matrice triangulaire supérieure T qui est semblable à A dans $\mathbb{C}$.*

Autrement dit, pour tout $f \in L(E)$ où E est un $\mathbb{C}$-espace vectoriel, on peut trouver une base de E dans laquelle la matrice associée à f est triangulaire.

Preuve : On procède par récurrence sur la dimension n de l'espace vectoriel E.

Initialisation : pour $n = 1$, toute matrice est triangulaire, donc n'importe quelle base convient.

Hérédité : Soit $n \geq 1$. Supposons la propriété vraie au rang n, et montrons-là au rang $n+1$. Le polynôme caractéristique de f est de degré $n+1 \geq 1$; il admet donc au moins un racine $\lambda \in \mathbb{C}$. La dimension de l'espace propre $E_\lambda = \mathrm{Ker}(f - \lambda id)$ est strictement positive. Soit v_λ vecteur propre pour la valeur propre λ, et F un sous-espace supplémentaire de $\mathbb{C}v_\lambda$. Remarquons que F est de dimension n.

Donnons-nous une base $\mathscr{B}_F = (e_1, \cdots, e_n)$ de F, et posons $\mathscr{B} := (v_\lambda, e_1, \cdots, e_n)$. Dans la base $\mathscr{B}$, la matrice de f est de la forme

$$A' = \begin{pmatrix} \lambda & \ell \\ 0 & A \end{pmatrix},$$

où ℓ est un vecteur ligne et A une matrice $n \times n$. Par la propriété de récurrence, on peut trouver une matrice de passage P de taille $n \times n$ telle que $P^{-1}AP$ soit triangulaire supérieure. Posons alors

$$P' := \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}, \quad (P')^{-1} := \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix}.$$

Un calcul matriciel par blocs donne alors :

$$(P'^{-1}AP) = \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} \cdot \begin{pmatrix} \lambda & \ell \\ 0 & A \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} \cdot \begin{pmatrix} \lambda & \ell P \\ 0 & AP \end{pmatrix} = \begin{pmatrix} \lambda & \ell P \\ 0 & P^{-1}AP \end{pmatrix},$$

qui est bien triangulaire supérieure. □

VI.2 Sous-espaces caractéristiques

Définition VI.2.

Soit $f \in L(E)$ et λ une valeur propre associée. On appelle **sous-espace caractéristique** *associé à λ le sous-espace vectoriel $Ker(f - \lambda id)^p$, où p est le plus petit entier tel que $Ker(f - \lambda id)^p = Ker(f - \lambda id)^{p+1}$.*

14. Ou, plus généralement, dans un corps *algébriquement clos*.

Notation : On note $(\lambda_i)_{i\in\{1;k\}}$ l'ensemble des valeurs propres associées à f et pour $i \in \{1;k\}$, on note $K_i = \mathrm{Ker}(f-\lambda id)^{p_i}$ le sous-espace caractéristique associé à λ_i et n_i sa dimension.

L'objectif de ce paragraphe est de démontrer que tout endomorphisme admet une matrice qui s'écrit comme la somme d'une matrice diagonale et d'une matrice de Jordan (c'est la décomposition de Dunford). Pour cela, on va démontrer que :

1. pour tout $i \in \{1;k\}$, K_i est stable pour f ;
2. $E = \bigoplus_{i\in\{1;k\}} K_i$.

L'un des arguments essentiels repose sur le lemme des noyaux, qui permet d'appliquer les résultats d'arithmétique sur les polynômes aux décompositions sur les sous-espace vectoriel stables de f.

Notation : Pour $i \in \{1;k\}$, on note $P_i \in \mathbb{R}[X]$ défini par : $P_i(X) = (X-\lambda_i)^{p_i}$. Alors on peut remarquer que $(f-\lambda_i id)^{p_i}$ s'écrit aussi $P_i(f)$.

Remarque VI.3.

Les polynômes $P_1, \cdots, P_k$ sont deux à deux premiers entre eux.

Preuve : Voir le cours de L2. On peut dire par exemple qu'ils n'ont pas de facteurs premiers communs. □

VI.3 Étape 1 : stabilité des sous-espaces caractéristiques

Propriété VI.4.

Pour tout $i \in \{1;k\}$, le sous-espace vectoriel K_i est stable par f.

Preuve : On peut démontrer plus généralement que pour tout polynôme $P \in \mathbb{C}[X]$, $\mathrm{Ker}(P(f))$ est stable par f. Il faut démontrer que si $u \in \mathrm{Ker}(P(f))$, alors $f(u) \in \mathrm{Ker}(P(f))$. Soit $u \in \mathrm{Ker}(P(f))$. Posons $P(X) := \sum_{k=0}^{n} a_k X^k$. Alors [15] :

$$\begin{aligned}P(f)(f(u)) &= \left(\sum_{k=0}^{n} a_k f^k\right)(f(u)) = \sum_{k=0}^{n} a_k f^{k+1}(u) = f\left(\left(\sum_{k=0}^{n} a_k f^k\right)(u)\right)\\ &= f(P(f)(u)) = f(0) = 0,\end{aligned}$$

donc $f(u) \in \mathrm{Ker}(P(f))$. □

On en déduit qu'on peut alors restreindre l'étude de f sur K_i à l'étude d'un endomorphisme sur K_i (et non plus sur E). Si on note f_i cet endomorphisme, alors $f_i - \lambda_i id$ est nilpotent. On peut donc lui associer une matrice de Jordan (qui est de taille n_i).

Il reste à démontrer qu'on peut reconstruire f entièrement à partir de la connaissance de $f_1, f_2, \cdots, f_k$. C'est l'objet de l'étape 2.

15. Ce qui suit repose sur l'observation que $f \circ P(f) = P(f) \circ f$.

VI.4 Étape 2 : somme directe des sous-espaces caractéristiques

On utilise ici le lemme des noyaux :

Lemme VI.5.
Soient $f \in L(E)$ et $(Q_i)_{i\in\{1;k\}} \in (\mathbb{C}[X])^k$ des polynômes deux à deux premiers entre eux. Alors :

$$Ker\left(\Pi_{i=1}^{k} Q_i(f)\right) = Ker(Q_1(f)) \oplus Ker(Q_2(f)) \oplus \cdots \oplus Ker(Q_k(f)).$$

PREUVE : La preuve généralise la proposition, vue plus tôt, que les sous-espaces propres sont en somme directe. La méthode de résolution est similaire (récurrence sur k). □

Exercice 36. Reprendre la démonstration du fait que les sous-espaces propres sont en somme directe, et l'adapter pour obtenir la démonstration du lemme ci-dessus.

Le lemme appliqué à $(P_j)_{j\in\{1;k\}}$ montre alors que les sous-espaces caractéristiques sont en somme directe (on rappelle que $K_i = \mathrm{Ker}(P_i(f)) = \mathrm{Ker}((f - \lambda_i id)^{p_i}))$:

$$K_1 \oplus K_2 \oplus \cdots \oplus K_k \subset E$$

Cela permet donc de choisir une base de Jordan pour chacun des f_i et de "recoller" ces bases pour former une nouvelle famille libre.

VI.5 Théorème de Cayley Hamilton et décomposition de Dunford

Pour terminer, il reste à démontrer que cette famille est en fait une base de E, c'est à dire que la somme recouvre E tout entier. Pour cela, on relie les dimensions des sous-espaces caractéristiques aux multiplicités des valeurs propres dans le polynôme caractéristique.

Notation : Le polynôme caractéristique associé à f est défini par $\chi_f(X) := det(Xid - f)$. Il est scindé sur $\mathbb{C}$, et ses racines sont les valeurs propres $\lambda_1,\cdots,\lambda_k$ de f. En notant m_i la multiplicité de λ_i dans χ_f,

$$\chi_f(X) = \Pi_{i=1}^{k}(X - \lambda_i)^{m_i}.$$

Propriété VI.6.
Soient E un $\mathbb{C}$-espace vectoriel et $f \in L(E)$. Avec les notations précédentes, $\dim(K_i) = m_i$ *pour tout $i \in \{1;k\}$.*

PREUVE : Soit $\mathscr{B}$ une base dans laquelle la matrice représentant f est triangulaire supérieure. Alors la matrice A de $(\lambda_i f - id)$ dans cette base a exactement m_i zéros sur la diagonale ; quitte à permuter les éléments de la base, on peut supposer que les m_i derniers coefficients diagonaux de A sont nuls.

Alors A^{m_i} est triangulaire supérieure, et le dernier bloc $m_i \times m_i$ est nul. Il s'ensuite que $\dim(\mathrm{Ker}(f - \lambda_i id)^{m_i}) \geq m_i$. Mais, par définition de l'indice de nilpotence, $\dim(K_i) \geq \dim(\mathrm{Ker}(f - \lambda_i id)^{m_i})$, donc $\dim(K_i) \geq m_i$. Il suffit maintenant de montrer que $\dim(K_i) \leq m_i$

De plus, toujours dans la base $\mathscr{B}$, la matrice représentant f a $n - m_i$ coefficients diagonaux non nuls. Comme cette matrice est triangulaire supérieure, ses itérées ont toujours $n - m_i$ coefficients diagonaux non nuls, et sont donc de rang an moins $n - m_i$. Par le théorème du rang, $\dim(\mathrm{Ker}((f - \lambda_i id)^k)) \leq n - (n - m_i) = m_i$ pour tout m, et en particulier $\dim(K_i) = \dim(\mathrm{Ker}((f - \lambda_i id)^{p_i})) \leq m_i$.
□

Conséquence VI.7.

$$E = Ker(f - \lambda_1 id)^{p_1} \oplus Ker(f - \lambda_2 id)^{p_2} \oplus \cdots \oplus Ker(f - \lambda_k id)^{p_k}.$$

Preuve : Les sous-espaces $\mathrm{Ker}(f - \lambda_i id)^{p_i}$ sont en somme directe, donc

$$\begin{aligned} &\dim\left(\mathrm{Ker}(f - \lambda_1 id)^{p_1} \oplus \mathrm{Ker}(f - \lambda_2 id)^{p_2} \oplus \cdots \oplus \mathrm{Ker}(f - \lambda_k id)^{p_k}\right) \\ &= \sum_{i=1}^{k} \dim\left(\mathrm{Ker}(f - \lambda_i id)^{p_i}\right) = \sum_{i=1}^{k} m_i = \deg(\chi_f) = \dim(E). \end{aligned}$$

□

En utilisant le lemme des noyaux, on en déduit que $E = \mathrm{Ker}\left(\Pi_{i=1}^k (f - \lambda_i id)^{p_i}\right)$, ce qui permet de montrer le théorème de Cayley Hamilton énoncé ci-dessous.

Théorème 5 (Cayley Hamilton)**.**
Soient E un $\mathbb{C}$-espace vectoriel de dimension finie et $f \in L(E)$. Le polynôme caractéristique χ_f de f est un polynôme annulateur de f, c'est à dire qu'il vérifie $\chi_f(f) = 0$.

Preuve : En utilisant ce qui précède, il suffit de vérifier que $\Pi_{i=1}^k (X - \lambda_i)^{p_i}$ divise χ_f. Or on a vu que l'indice de nilpotence p_i était égal au plus à la dimension de l'espace associé :

$$p_i \leq \dim(\mathrm{Ker}(f - \lambda_i id)^{p_i}) = m_i.$$

Par conséquent,

$$\chi_f(X) = \Pi_{i=1}^k (X - \lambda_i)^{m_i - p_i} \cdot \Pi_{i=1}^k (X - \lambda_i)^{p_i},$$

□

Exercice 37. * Il y a d'autres démonstrations de ce théorème, en particulier l'une d'elle utilise de la topologie... et elle est très courte ! Indice : montrez le théorème de Cayley-Hamilton pour des endomorphismes diagonalisables, puis cherchez un argument topologique pour l'étendre à tous les endomorphismes.

Remarque VI.8.
On peut remarquer que $\mu_f(X) = \Pi_{i=1}^k (X - \lambda_i)^{p_i}$ est un polynôme annulateur de f. C'est le plus petit possible, et il divise le polynôme caractéristique ; il s'appelle le polynôme minimal.

Exercice 38. Vérifier que μ_f est en effet le plus petit polynôme annulateur de f.

Théorème 6 (Décomposition de Dunford)**.**
Soient E un $\mathbb{C}$-espace vectoriel de dimension finie et $f \in L(E)$. Il existe une base dans laquelle la matrice associée à f s'écrit sous la forme $D + N$, où D est une matrice diagonale de $\mathscr{M}_n(\mathbb{C})$ et N une matrice de Jordan vérifiant $ND = DN$.

Exemple de matrice réduite dans une base de Jordan :

$$
\begin{array}{cc}
 & \begin{array}{ccc} \xleftrightarrow{\qquad m_1 \qquad} & \cdots & \xleftrightarrow{\qquad m_k \qquad} \end{array} \\
\begin{array}{c} K_1 \\ \\ \vdots \\ \\ K_k \end{array} & \begin{pmatrix}
\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 1 & \lambda_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & \lambda_1 \end{bmatrix} & & 0 \\
 & \ddots & \\
0 & & \begin{bmatrix} \lambda_k & 0 & \cdots & 0 \\ 1 & \lambda_k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k \end{bmatrix}
\end{pmatrix}
\end{array}
$$

VI.6 Applications

Voici deux approfondissements possibles :

- ▷ La classification des matrices semblables : déterminer un invariant de similitude complet pour les matrices carrées $n \times n$ complexes.
- ▷ Les exponentielles de matrices : l'exponentielle d'un endomorphisme f est définie par la somme (convergente)

$$exp(f) := \sum_{n=0}^{+\infty} \frac{f^n}{n!}.$$

Déterminez explicitement, dans une base bien choisie, $exp(f)$ quand f est diagonalisable, puis quand f est nilpotente, puis quand f n'a qu'une seule valeur propre. Décrire $exp(f)$ dans le cas général.