Université de Lille

M31 Algèbre Linéaire 2020-2021

Poly du cours M201, année 2008-2009

S. Delaunay

Sommaire

Cours d'Algèbre Linéaire et Affine

Dans tout ce cours, le corps K désignera soit le corps $\mathbb{R}$ des réels, soit le corps $\mathbb{C}$ des complexes.

0. Rappels d'algèbre linéaire

Nous allons commencer par quelques rappels d'algèbre linéaire indispensables, sans démonstrations (celles-ci étant à chercher en exercice).

0.1. Espaces vectoriels

Un espace vectoriel est un ensemble stable par combinaison linéaire.

Définition. *Un ensemble E muni d'une opération interne notée $+$ et d'une opération externe est un espace vectoriel sur un corps K, s'il vérifie les propriétés suivantes :*
1) L'ensemble $(E, +)$ est un groupe commutatif :
- l'opération $+$ est associative et commutative,
- il existe un élément neutre noté 0_E qui vérifie pour tout $u \in E$, $u + 0_E = u$,
- tout élément de E admet un symétrique (ou opposé), pour tout $u \in E$, il existe $v \in E$ tel que $u + v = 0_E$ on notera cet opposé $-u$.
2) Pour tout $u, v \in E$ et pour tout $\lambda, \mu \in K$, on a
- $1_K.u = u$,
- $\lambda.(u + v) = \lambda.u + \lambda.v$,
- $(\lambda + \mu).u = \lambda.u + \mu.u$,
- $\lambda.(\mu.v) = (\lambda\mu).v$.

Les éléments de E sont appelés *vecteurs*, les éléments de K sont appelés *scalaires*.

Exemples : Le corps $\mathbb{C}$ des nombres complexes est un espace vectoriel sur $\mathbb{C}$ et sur $\mathbb{R}$, l'ensemble $\mathbb{R}^n$, $n \in \mathbb{N}$, est un espace vectoriel sur $\mathbb{R}$.

Définition. *Un sous-ensemble non vide E' de E est un sous-espace vectoriel de E, s'il est stable par combinaisons linéaires.*

Exemple *: $\mathbb{R}^2$ est un sous-espace vectoriel de $\mathbb{R}^3$.*

Proposition. *Soient $E_1, \ldots, E_p$ des sous-espaces vectoriels de E, l'ensemble*

$$F = E_1 + \ldots + E_p = \{u_1 + \ldots + u_p,\ u_1 \in E_1, \ldots, x_p \in E_p\}$$

somme des sous-espaces vectoriels $E_1, \ldots, E_p$, est un sous-espace vectoriel de E. C'est le plus petit sous-espavce vectoriel de E contenant la réunion $E_1 \cup \ldots \cup E_p$.

Définition. *Deux sous-espaces vectoriels, E_1 et E_2, de E sont dits supplémentaires si $E_1 \cap E_2 = \{0\}$ et $E = E_1 + E_2$, on note alors $E = E_1 \oplus E_2$ et tout vecteur de E se décompose de manière unique en somme d'un vecteur de E_1 et d'un vecteur de E_2.*

Familles libres, génératrices. Bases

Définitions. *- Soient $(u_1, u_2, \dots, u_n)$ des vecteurs de E, on dit que la famille $(u_1, u_2, \dots, u_n)$ engendre l'espace vectoriel E si tout vecteur de E s'écrit comme combinaison linéaire des vecteurs $(u_1, u_2, \dots, u_n)$, c'est-à-dire si, pour tout $v \in E$, il existe des scalaires $(\lambda_1, \lambda_2, \dots, \lambda_n)$ tel que $v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n$.*
- Soient $(u_1, u_2, \dots, u_n)$ des vecteurs de E, on dit que la famille $(u_1, u_2, \dots, u_n)$ est libre, si pour tout $(\lambda_1, \dots, \lambda_n) \in K^n$, l'égalité $\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n = 0$ implique
$\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.
- Une famille de vecteurs, génératrice et libre, est appelée une base de E. Si E admet une base de n vecteurs, toutes les bases de E ont le même nombre, n, de vecteurs et n est la dimension de E.

Dans ce cours, tous les espaces vectoriels seront de dimension finie.

Si $(u_1, u_2, \dots, u_n)$ est une base de E, alors pour tout $v \in E$, il existe un unique n-uplet $(\lambda_1, \dots, \lambda_n) \in K^n$ tel que

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n.$$

0.2 Applications linéaires

Définition. *Soient E et F deux espaces vectoriels, une application $f : E \to F$ est dite K-linéaire ou linéaire si pour tout $(u, v) \in E^2$ et pour tout $(\lambda, \mu) \in K^2$,*

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

- Les applications linéaires sont encore appelées *homomorphismes* d'espaces vectoriels, on note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F.

- Les applications linéaires de E dans lui-même sont appelées *endomorphismes* de E, on note $\mathcal{L}(E)$ leur ensemble.

- Les applications linéaires *bijectives* de E dans F sont appelées *isomorphismes.*

- Les applications linéaires *bijectives* de E dans lui-même sont les *automorphismes* de E.

- Les applications linéaires de E dans le corps de base, K, sont appelées *formes linéaires* de E.

On définit sur $\mathcal{L}(E, F)$ une addition, pour f et $g \in \mathcal{L}(E, F)$ et pour $u \in E$, par

$$(f + g)(u) = f(u) + g(u),$$

et une multiplication externe, pour $\lambda \in K$, par

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

Nous allons maintenant rappeler quelques propositions évidentes que nous ne démontrerons pas.

Proposition. *L'ensemble $\mathcal{L}(E,F)$ muni de ces deux opérations a une structure d'espace vectoriel.*

Par ailleurs, la composée de deux applications linéaires est encore une application linéaire.

Proposition. *Soient E, F, G trois espaces vectoriels sur un corps K, soient $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$, alors $f \circ g \in \mathcal{L}(E,G)$.*

Proposition. *Soit E un espace vectoriel, l'ensemble $\mathcal{L}(E)$ muni de la somme et de la composition des applications est un anneau.*

Proposition. *Si f est une application linéaire bijective de E dans F, alors, son application inverse f^{-1} est une application linéaire bijective de F dans E, c'est-à-dire un isomorphisme entre E et F. Si une telle applicationj existe, on dit que les espaces vectoriels E et F sont isomorphes. Ils sont alors de même dimension.*

Noyau. Image

Proposition (définitions). *Soient E et F deux espaces vectoriel et $f \in \mathcal{L}(E,F)$. Notons E' un sous-espace vectoriel de E et F' un sous-espace vectoriel de F.*
- L'ensemble $f(E')$ est un sous-espace vectoriel de F. En particulier l'ensemble suivant appelé image de f :

$$\operatorname{Im} f = f(E) = \{f(u),\ u \in E\} = \{y \in F,\ \exists u \in E,\ v = f(u)\}.$$

L'application f est surjective si et seulement si $\operatorname{Im} f = F$.
- L'ensemble $f^{-1}(F') = \{u \in E,\ f(u) \in F'\}$ est un sous-espace vectoriel de E. En particulier l'ensemble suivant appelé noyau de f :

$$\ker f = \{u \in E,\ f(u) = 0_F\}.$$

L'application f est injective si et seulement si $\ker f = \{0_E\}$.

Théorème du rang. *Soient E et F deux espaces vectoriels (de dimensions finies) et $f \in \mathcal{L}(E,F)$, on a*

$$\dim E = \dim \ker f + \dim \operatorname{Im} f.$$

Démonstration : Supposons E de dimension n, et notons m la dimension du sous-espace vectoriel $\ker f$ de E. Soit $(e_1, e_2, \ldots, e_m)$ une base de $\ker f$, on la complète par des vecteurs $e_{m+1}, \ldots, e_n$ en une base de E. Les vecteurs $f(e_1), \ldots, f(e_n)$ engendrent $\operatorname{Im} f$, mais comme les vecteurs $e_1, \ldots, e_m$ sont dans le noyau, le sous-espace $\operatorname{Im} f$ est engendré par les vecteurs $f(e_{m+1}), \ldots, f(e_n)$. Montrons que ces vecteurs forment une base de $\operatorname{Im} f$, pour cela il suffit de démontrer qu'ils sont linéairement indépendants.

Supposons qu'il existe des scalaires $\lambda_{m+1}, \ldots, \lambda_n$ tels que $\lambda_{m+1} f(e_{m+1}) + \ldots + \lambda_n f(e_n) = 0$, on a alors

$$f(\lambda_{m+1} e_{m+1} + \ldots + \lambda_n e_n) = 0,$$

c'est-à-dire $\lambda_{m+1} e_{m+1} + \ldots + \lambda_n e_n \in \ker f$, par conséquent, il existe des scalaires $\mu_1, \ldots, \mu_m$ tels que

$$\lambda_{m+1} e_{m+1} + \ldots + \lambda_n e_n = \mu_1 e_1 + \ldots + \mu_m e_m.$$

Mais, comme les vecteurs $e_1, \ldots, e_m, e_{m+1}, \ldots, e_n$ forment une base de E, ceci implique la nullité de tous les coefficients. Ainsi les vecteurs $f(e_{m+1}), \ldots, f(e_n)$ forment bien une base de $\operatorname{Im} f$. Le sous-espace vectoriel de $\operatorname{Im} f$ de F est donc de dimension $n - m$ et, on a bien,

$$n = \dim E = m + (n - m) = \dim \ker f + \dim \operatorname{Im} f.$$

0.3 Matrices

Nous allons maintenant, avant d'aborder les déterminants, terminer ce chapitre de rappels par quelques mots sur les matrices.

Définition. *On appelle matrice à n lignes et m colonnes, à coefficients dans K, le tableau suivant :*

$$M = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}$$

On note $M_{n,m}(K)$, l'ensemble de ces matrices et $M_n(K)$ cet ensemble si $m = n$.

On définit sur $M_{n,m}(K)$ une addition, composante par composante et une multiplication externe par un élément de K sur chaque composante. Muni de ces deux opérations, l'ensemble $M_{n,m}(K)$ a une structure d'espace vectoriel.

Définitions. *Soit $M \in M_n(K)$ une matrice carrée. La matrice $M = (a_{ij})_{1 \leq i,j \leq n}$ est dite*
- triangulaire supérieure si $a_{ij} = 0$ pour tout (i, j) tel que $i > j$,
- triangulaire inférieure si $a_{ij} = 0$ pour tout (i, j) tel que $i < j$,
- diagonale si $a_{ij} = 0$ pour tout (i, j) tel que $i \neq j$,
- symétrique si $a_{ij} = a_{ji}$ pour tout (i, j).

Définition. *Soit $M = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m} \in M_{n,m}(K)$, on appelle transposée de M, la matrice ${}^tM = (b_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ où $b_{ij} = a_{ji}$ pour tout (i, j).*

Produit de matrices

On définit également un *produit* de matrices pour $A = (a_{ik})_{1 \leq i \leq n, 1 \leq k \leq m} \in M_{n,m}(K)$ et $B = (b_{kj})_{1 \leq k \leq m, 1 \leq j \leq p} \in M_{m,p}(K)$ par

$$A.B = (c_{ij})_{1 \leq i \leq n, 1 \leq j \leq p} \in M_{n,p}(K), \text{ où } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

A la place (ij), le coefficient c_{ij} est le produit scalaire de la i-ème ligne de la matrice A et de la j-ème colonne de la matrice B. Remarquons que ce produit n'est possible que si le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B et qu'il n'est évidemment pas commutatif, même dans le cas des matrices carrées.

Définition. *Une matrice $A \in M_n(K)$ est dite inversible s'il existe une matrice $B \in M_n(K)$ telle que $A.B = B.A = I_n$ où I_n désigne la matrice diagonale telle que $a_{ii} = 1$ pour tout i. La matrice I_n est appelée matrice identité, elle est élément neutre pour le produit des matrices.*

Proposition. *L'ensemble $M_n(K)$ muni de l'addition et du produit des matrices est un anneau (non commutatif). L'ensemble des matrices inversibles de $M_n(K)$ est un groupe pour le produit des matrices, c'est le groupe linéaire noté $GL_n(K)$.*

Nous allons maintenant relier matrices et applications linéaires.

Soient E et F deux espaces vectoriels, on note m la dimension de E et n la dimension de F. Soient $\mathcal{B} = (e_1, \dots, e_m)$ une base de E et $\mathcal{B}' = (f_1, \dots, f_n)$ une base de F. Soit $\varphi \in \mathcal{L}(E, F)$.

Définition. *On appelle matrice de φ dans les bases $\mathcal{B}$ et $\mathcal{B}'$, la matrice dont les colonnes sont les coordonnées des vecteurs $\varphi(e_j)_{1\leq j\leq m}$ exprimées dans la base $\mathcal{B}'$.*

Si pour tout j, $1 \leq j \leq m$, on a

$$\varphi(e_j) = \sum_{i=1}^{n} a_{ij} f_i,$$

la matrice de φ, relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$ est la matrice $A = (a_{ij})_{1\leq i\leq n, 1\leq j\leq m}$.

Proposition. *L'application de $\mathcal{L}(E, F)$ dans $M_{n,m}(K)$ qui à une application linéaire φ associe sa matrice dans des bases données de E et F est un isomorphisme d'espaces vectoriels.*

Ainsi, la matrice d'une application linéaire dépend des bases dans lesquelles elle est exprimée. Néanmoins, certaines propriétés de l'application linéaire, et par là de sa matrice, sont conservées quelques soient les bases dans lesquelles on exprime cette matrice. C'est là un point essentiel puisque, dans la suite de ce cours, il s'agira de déterminer des bases dans lesquelles la matrice d'une application linéaire donnée soit la plus simple possible, triangulaire ou diagonale par exemple. C'est pourquoi, il faut bien comprendre les changements de bases et leurs matrices.

Changements de basess

Soit E un espace vectoriel de dimension n et soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ deux bases de E. Pour j, $1 \leq j \leq n$, on note $(a_{ij})_{1\leq i\leq n}$ les coordonnées du vecteur $e'j$ dans la base $\mathcal{B}$, on a

$$e'_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Définition et propriétés. *La matrice $P = (a_{ij})_{1\leq i\leq n}$ est appelée matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Elle est inversible et la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$ est la matrice P^{-1}.*

Si u est un vecteur de E, on peut exprimer ses coordonnées dans la base $\mathcal{B}$ ou dans la base $\mathcal{B}'$, on note

$$u = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} x'_i e'_i.$$

Si on note

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix},$$

on a

$$X = P.X'.$$

Proposition. *Soient E un espace vectoriel, $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ deux bases de E . Si $\varphi \in \mathcal{L}(E)$, notons M la matrice de φ dans la base $\mathcal{B}$ et M' la matrice de φ dans la base $\mathcal{B}'$, alors on a*

$$M' = P^{-1}.M.P.$$

On dit alors que les matrices M et M' sont semblables.

Avant d'aborder les déterminants, nous allons tenter d'illustrer ce qui précède par quelques exemples en dimension 2 et 3.

0.4 Exemples. Homothéties, projections, symétries

Homothéties vectorielles

Soit E un espace vectoriel et $\lambda \in K$, on note h_λ l'application :

$$\begin{aligned} h_\lambda : E &\longrightarrow E \\ u &\longmapsto \lambda.u \end{aligned}$$

Cette application est linéaire et bijective, on a $h_\lambda = \lambda \operatorname{id}_E$, $\ker h_\lambda = \{0_E\}$ et $\operatorname{Im} h_\lambda = E$. Quelque soit le base choisie, la matrice de h_λ s'écrit

$$M_\lambda = \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & \ldots & \lambda \end{pmatrix}.$$

Projections vectorielles

Définition et propriétés. Soient E_1 et E_2 deux sous-espaces supplémentaires d'un espace vectoriel E. Tout vecteur u de E s'écrit de manière unique $u = u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$, on appelle *projection vectorielle* (ou projecteur) de E sur E_1 parallèlement à E_2 l'application

$$\begin{aligned} p : E &\longrightarrow E \\ u &\longmapsto u_1. \end{aligned}$$

Cette application est linéaire et vérifie les propriétés suivantes :

1) $p \circ p = p$.

2) $\operatorname{Im} p = E_1$ et $\ker p = E_2$.

3) E_1 est l'ensemble des vecteurs invariants par p.

Si E est de dimension 2, les espaces E_1 et E_2 sont des droites vectorielles, notons e_1 une base de E_1 et e_2 une base de E_2, les vecteurs e_1 et e_2 forment une base de E dans cette base, la matrice de la projection p sur E_1 parallèlement à E_2 s'écrit

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Si E est de dimension 3, supposons E_1 de dimension 2 engendré par e_1 et e_2, et E_2 de dimension 1 engendré par e_3. Dans la base (e_1, e_2, e_3) de E la matrice de la projection p sur E_1 parallèlement à E_2 s'écrit

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Symétries vectorielles

Définition et propriétés. Soient E_1 et E_2 deux sous-espaces supplémentaires d'un espace vectoriel E. Tout vecteur u de E s'écrit de manière unique $u = u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$, on appelle *symétrie vectorielle* de E par rapport à E_1 parallèlement à E_2 l'application

$$\begin{aligned} s : E &\longrightarrow E \\ u &\longmapsto u_1 - u_2. \end{aligned}$$

Cette application est linéaire et vérifie les propriétés suivantes :

1) $s \circ s = \mathrm{id}_E$.

2) $\operatorname{Im} s = E$ et $\ker s = \{0\}$.

3) E_1 est l'ensemble des vecteurs invariants par s.

4) E_2 est l'ensemble des vecteurs tels que $s(u) = -u$.

Si E est de dimension 2, les espaces E_1 et E_2 sont des droites vectorielles, notons e_1 une base de E_1 et e_2 une base de E_2, les vecteurs e_1 et e_2 forment une base de E dans cette base, la matrice de la symétrie s par rapport à E_1 parallèlement à E_2 s'écrit

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

Si E est de dimension 3, supposons E_1 de dimension 2 engendré par e_1 et e_2, et E_2 de dimension 1 engendré par e_3. Dans la base (e_1, e_2, e_3) de E la matrice de la symétrie s par rapport à E_1 parallèlement à E_2 s'écrit

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

I. Les Déterminants

I.1. Formes multilinéaires alternées

I.1.1.Définitions et premières propriétés.

Définition 1. *Soient $E_1, \ldots E_p, F$ des espaces vectoriels sur K, une application*

$$\begin{aligned}\varphi : E_1 \times \ldots \times E_p &\longrightarrow F \\ (x_1, \ldots, x_p) &\longmapsto \varphi(x_1, \ldots, x_p)\end{aligned}$$

est dite p-linéaire, si en tout point, les p applications partielles sont linéaires.

Ce qui revient à dire que φ est linéaire par rapport à chacune de ses variables. Si l'on se fixe $i \in \{1, \ldots, p\}$ et des vecteurs $x_j \in E_j$ pour $j \in \{1, \ldots, p\} \setminus \{i\}$ alors, l'application

$$\begin{aligned}\varphi_i : E_i &\longrightarrow F \\ x &\longmapsto \varphi(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_p)\end{aligned}$$

est linéaire.

Dans le cas particulier où $E_1 = \ldots = E_p = E$ et $F = K$, on parle alors de formes p-linéaires et on note $\mathcal{L}_p(E)$ leur ensemble. Si $p = 1$, ce sont les formes linéaires, si $p = 2$ les formes bilinéaires.

Exemple : On note $a = (a_1, a_2)$ et $b = (b_1, b_2)$ des vecteurs de $\mathbb{R}^2$, et on définit

$$\varphi(a, b) = a_1 b_2 - a_2 b_1$$

L'application φ est bilinéaire alternée, en effet si x et y sont des vecteurs de $\mathbb{R}^2$, α et β des réels, on vérifie que

$$\varphi(a, \alpha x + \beta y) = \alpha\varphi(a, x) + \beta\varphi(a, y)$$

et

$$\varphi(\alpha x + \beta y, b) = \alpha\varphi(x, b) + \beta\varphi(y, b).$$

Définition 2 . *Une forme p-linéaire, φ sur E est dite alternée, si*

$$\varphi(x_1, \ldots, x_p) = 0$$

dès que deux vecteurs parmi les x_i, $1 \leq i \leq p$, sont égaux.

Une conséquence de cette définition est la propriété suivante :

Propriété 1. *Soit φ une forme p-linéaire alternée sur E, on ne change pas la valeur de $\varphi(x_1, \ldots, x_p)$, en ajoutant à un des vecteurs une combinaison linéaire des autres.*

Démonstration.

Soient $x_1, \ldots, x_p$ des vecteurs de E et $\lambda_1, \ldots, \lambda_p$ des éléments de K. On a

$$\varphi(x_1, \ldots, \underbrace{x_i + \sum_{k=1, k\neq i}^{n} \lambda_k x_k}_{i^{\text{eme}}\text{place}}, \ldots, x_p) = \varphi(x_1, \ldots, x_p) + \sum_{k=1, k\neq i}^{n} \lambda_k \varphi(x_1, \ldots, \underbrace{x_k}_{i^{\text{eme}}\text{place}}, \ldots, x_p).$$

La somme qui apparaît dans le membre de droite est nulle, en effet la forme multilinéaire φ étant alternée, pour tout $k \neq i$, $\varphi(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_p) = 0$ car x_k apparaît deux fois. Remarquons que cette propriété implique que $\varphi(x_1, \ldots, x_p) = 0$ dès que les vecteurs $x_1, \ldots, x_p$ sont linéairement dépendants.

Afin de pouvoir définir le déterminant, nous allons faire quelques rappels sur le groupe des permutations.

I.1.2 Permutations

Définition 3. *Pour tout $n \in \mathbb{N}^*$, on appelle groupe des permutations de $\{1, \ldots, n\}$, noté S_n le groupe des bijections de $\{1, \ldots, n\}$ dans $\{1, \ldots, n\}$. On a $\operatorname{card} S_n = n!$.*

Définition 4. *On appelle transposition sur i et j, la permutation notée $\tau_{i,j}$ qui échange i et j.*

Théorème . *Toute permutation se décompose en produit de transpositions. Ce produit n'est pas unique mais, pour $\sigma \in S_n$ fixé, la parité du nombre de ces transpositions est fixé, c'est la signature de la permutation σ.*

La démonstration de ce théorème se fait par récurrence sur n, on ne la fera pas ici. Cette décomposition n'est pas unique, mais, pour σ fixée, la parité du nombre de transpositions entrant dans la décomposition de σ est fixe. C'est ce qui définit la signature de la permutation σ.

Définition 5. *Soit $\sigma \in S_n$, on appelle signature de σ, notée $\varepsilon(\sigma)$, le nombre appartenant à $\{-1, 1\}$ défini par*

$$\varepsilon(\sigma) = \frac{\prod_{1 \leq i < j \leq n}(\sigma(i) - \sigma(j))}{\prod_{1 \leq i < j \leq n}(i - j)}$$

Il est clair que la signature d'une transposition est égale à -1. Lorsque $\varepsilon(\sigma) = 1$, on dit que la permutation est paire, lorsque $\varepsilon(\sigma) = -1$ on dit qu'elle est impaire. Le sous-groupe des permutations paires est appelé groupe alterné et noté A_n.

Propriété. *Soient σ et σ' deux éléments de S_n, on a*

$$\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma').$$

Ce qui signifie que la signature est un morphisme de groupes de S_n sur le groupe multiplicatif $\{-1, 1\}$.

Proposition 1. *Soit φ une forme p-linéaire sur E. Alors, φ est alternée si et seulement si pour toute permutation σ de S_p et pour tout $(x_1, \ldots, x_p) \in E^p$ on a*

$$\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) = \varepsilon(\sigma)\varphi(x_1, \ldots, x_p).$$

Compte tenu de la décomposition de σ en produit de transpositions et de la propriété de morphisme de la signature, nous allons démontrer la proposition suivante, dont on peut déduire la Proposition 1 :

Proposition 2. *Soit φ une forme p-linéaire sur E. Alors, φ est alternée si et seulement si pour toute transposition τ, on a*

$$\varphi(x_{\tau(1)}, \dots, x_{\tau(p)}) = -\varphi(x_1, \dots, x_p).$$

Démonstration de la proposition 2.

Soient $i, k \in \{1, \dots, p\}$, $i < k$. On suppose φ alternée, on note $\tau = \tau_{ik}$ la transposition sur $\{i, k\}$. On a

$$\begin{aligned} 0 &= \varphi(x_1, \dots, x_{i-1}, \underbrace{x_i + x_k}_{i^{\text{eme}}\text{place}}, x_{i+1}, \dots, x_{k-1}, \underbrace{x_k + x_i}_{k^{\text{eme}}\text{place}}, x_{k+1}, \dots, x_p) \\ &= \varphi(x_1, \dots, x_p) + \varphi(x_{\tau(1)}, \dots, x_{\tau(p)}) \end{aligned}$$

car $x_{\tau(j)} = x_j$ pour $j \neq i$ et $j \neq k$, et $x_{\tau(i)} = x_k$, $x_{\tau(k)} = x_i$. D'où le résultat.

Réciproquement, si pour toute transposition τ, on a $\varphi(x_{\tau(1)}, \dots, x_{\tau(p)}) = -\varphi(x_1, \dots, x_p)$, alors, si $x_i = x_k$ on a d'une part,

$$\varphi(x_{\tau(1)}, \dots, x_{\tau(p)}) = \varphi(x_1, \dots, x_p)$$

et d'autre part,

$$\varphi(x_{\tau(1)}, \dots, x_{\tau(p)}) = -\varphi(x_1, \dots, x_p)$$

d'où $\varphi(x_1, \dots, x_p) = -\varphi(x_1, \dots, x_p)$, c'est-à-dire $\varphi(x_1, \dots, x_p) = 0$, ce qui prouve que la forme est alternée.

I.2. Déterminants

I.2.1. Définition et propriétés

Théorème. *Soit E un K-espace vectoriel de dimension n. L'ensemble des formes n-linéaires alternées sur E est un K-espace vectoriel de dimension 1. De plus, si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E, il existe une unique forme n-linéaire alternée Δ vérifiant*

$$\Delta(e_1, \dots, e_n) = 1\,.$$

Démonstration.

On se fixe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E. Soient $x_1, \dots, x_n$, n vecteurs de E. Pour $1 \leq j \leq n$, on note $(x_{ij})_{1 \leq i \leq n}$ les coordonnées du vecteur x_j dans la base $\mathcal{B} = (e_1, \dots, e_n)$, on a $x_j = \sum_{i=1}^{n} x_{ij} e_i$ et,

$$\varphi(x_1, \dots, x_n) = \varphi\left(\sum_{i=1}^{n} x_{i1} e_i, \dots, \sum_{i=1}^{n} x_{in} e_i\right),$$

en utilisant le fait que φ est, d'une part n-linéaire, d'autre part alternée, on obtient

$$\begin{aligned}\varphi(x_1,\dots,x_n) &= \sum_{\sigma\in S_n} x_{\sigma(1)1}\cdots x_{\sigma(n)n}\varphi(e_{\sigma(1)},\dots,e_{\sigma(n)})\\ &= \sum_{\sigma\in S_n} x_{\sigma(1)1}\cdots x_{\sigma(n)n}\varepsilon(\sigma)\varphi(e_1,\dots,e_n)\\ &= \left(\sum_{\sigma\in S_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}\right)\varphi(e_1,\dots,e_n),\end{aligned}$$

d'après la proposition 1. Notons $\Delta(x_1,\dots,x_n)=\sum_{\sigma\in S_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}$. Comme $\varphi(e_1,\dots,e_n)$ est une constante qui ne dépend ni des vecteurs x_i, ni des permutations σ mais uniquement de la base $\mathcal{B}$, en démontrant que Δ est une forme n-linéaire alternée non nulle nous aurons démontré que toute forme n-linéaire est un multiple de Δ et ainsi prouvé que l'ensemble des formes n-linéaires alternées sur E est un K-espace vectoriel de dimension 1.

Démontrons donc que Δ est une forme n-linéaire alternée non nulle. Soit $(x_1,\dots,x_n)\in E^n$, on suppose qu'il existe i et $k\in\{1,\dots,n\}$ tels que $x_i=x_k$. On a alors

$$\begin{aligned}\Delta(x_1,\dots,x_n) &= \sum_{\sigma\in S_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}\\ &= \sum_{\sigma\in A_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}+\sum_{\sigma\notin A_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}.\end{aligned}$$

Notons τ la transposition qui échange i et k, l'application

$$\begin{aligned}A_n &\longrightarrow S_n\setminus A_n\\ \sigma &\longmapsto \sigma\circ\tau\end{aligned}$$

est une bijection. On peut donc écrire

$$\begin{aligned}\Delta(x_1,\dots,x_n) &= \sum_{\sigma\in A_n}\varepsilon(\sigma)x_{\sigma(1)1}\cdots x_{\sigma(n)n}+\sum_{\sigma\in A_n}\varepsilon(\sigma\circ\tau)x_{\sigma\circ\tau(1)1}\cdots x_{\sigma\circ\tau(n)n}\\ &= \sum_{\sigma\in A_n}x_{\sigma(1)1}\cdots x_{\sigma(n)n}-\sum_{\sigma\in A_n}x_{\sigma(1)1}\cdots x_{\sigma(n)n}=0,\end{aligned}$$

car $x_i=x_k$. Ce qui prouve que Δ est alternée. Par aileurs, Δ est clairement n-linéaire et non nulle et, pour $1\le j\le n$, on a

$$e_j=\sum_{i=1}^n\delta_{ij}e_j \quad\text{où}\quad \delta_{ij}=\begin{cases}1 & \text{si } i=j\\ 0 & \text{si } i\neq j\end{cases}$$

d'où

$$\Delta(e_1,\dots,e_n)=\sum_{\sigma\in S_n}\varepsilon(\sigma)\delta_{\sigma(1)1}\cdots\delta_{\sigma(n)n}=\delta_{11}\cdots\delta_{nn}=1.$$

L'espace des formes n-linéaires alternées étant de dimension 1, la forme Δ est unique.

Définition. *Avec les notations ci-dessus, le déterminant des vecteurs $x_1, \ldots, x_n$ dans la base $\mathcal{B} = (e_1, \ldots, e_n)$ est*

$$\det_{\mathcal{B}}(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n)\,.$$

Théorème. *Soient $x_1, \ldots, x_n$ des vecteurs de E, les propriétés suivantes sont équivalentes*
(i) les vecteurs $x_1, \ldots, x_n$ sont linéairement dépendants
(ii) Pour toute base B de E, $\det_B(x_1, \ldots, x_n) = 0$
(iii) Il existe une base B de E telle que $\det_B(x_1, \ldots, x_n) = 0$

Démonstration.

(i)⇒(ii) Quelque soit la base, le déterminant est une forme n-linéaire alternée, il prend donc la valeur 0 dès que les vecteurs sont linéairement dépendants.

(ii)⇒(iii) évident

(iii)⇒(i) On a vu que pour toute forme n-linéaire alternée φ, si $\mathcal{B} = (e_1, \ldots, e_n)$, pour tout $(x_1, \ldots, x_n) \in E^n$ on a

$$\varphi(x_1, \ldots, x_n) = \varphi(e_1, \ldots, e_n)\det_{\mathcal{B}}(x_1, \ldots, x_n).$$

Donc, s'il existe une base dans laquelle le déterminant des $x_1, \ldots, x_n$ est nul, alors, pour toute forme n-linéaire alternée φ, on a $\varphi(x_1, \ldots, x_n) = 0$. Or, si les x_i étaient linéairement indépendants, ils formeraient une base de E et il existerait une forme n-linéaire alternée prenant la valeur 1 sur cette base, ce qui est contradictoire. Les x_i sont donc linéairement dépendants.
Nous venons de définir le déterminant d'un système de n vecteurs, nous allons maintenant relier cette notion à celle de déterminant d'une matrice carrée ou d'un endomorphisme de E.

I.2.2 Déterminant d'une matrice

Définition. *Soit $A = (a_{i,j})_{1 \le i,j \le n} \in M_n(K)$. On appelle déterminant de A, noté* $\det A$, *le déterminant des vecteurs colonnes de la matrice A. On a*

$$\det A = \begin{vmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,n} \end{vmatrix} = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \ldots a_{\sigma(n),n}.$$

Le déterminant de A est le déterminant des vecteurs $c_j = \sum_{i=1}^{n} a_{i,j} e_i$ dans la base $(e_1, \ldots, e_n)$.

Avant d'en déduire les propriétés du déterminant, voyons ce que signifie cette définition si $n = 3$. Le groupe symétrique S_3 contient 6 éléments que nous noterons :

$$\mathrm{id}, \tau_1 = (23), \tau_2 = (13), \tau_3 = (12), \sigma_1 = (123), \sigma_2 = (132),$$

les permutations τ_1, τ_2, τ_3 sont des transpositions et σ_1, σ_2 sont des 3-cycles. Calculons le déterminant suivant :

$$\begin{aligned}
\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} =& \varepsilon(\mathrm{id}) a_{\mathrm{id}(1),1} a_{\mathrm{id}(2),2} a_{\mathrm{id}(3),3} + \varepsilon(\tau_1) a_{\tau_1(1),1} a_{\tau_1(2),2} a_{\tau_1(3),3} \\
& + \varepsilon(\tau_2) a_{\tau_2(1),1} a_{\tau_2(2),2} a_{\tau_2(3),3} + \varepsilon(\tau_3) a_{\tau_3(1),1} a_{\tau_3(2),2} a_{\tau_3(3),3} \\
& + \varepsilon(\sigma_1) a_{\sigma_1(1),1} a_{\sigma_1(2),2} a_{\sigma_1(3),3} + \varepsilon(\sigma_2) a_{\sigma_2(1),1} a_{\sigma_2(2),2} a_{\sigma_(3),3} \\
& = a_{1,1} a_{2,2} a_{3,3} + (-1) a_{1,1} a_{3,2} a_{2,3} + (-1) a_{3,1} a_{2,2} a_{1,3} \\
& + (-1) a_{2,1} a_{1,2} a_{3,3} + a_{2,1} a_{3,2} a_{1,3} + a_{3,1} a_{1,2} a_{2,3}.
\end{aligned}$$

On retrouve bien un développement du déterminant 3×3.

Propriétés. *Soient* $A, B \in M_n(K)$*, on a*
1) $\det A = \det {}^t\!A$
2) $\det A$ *dépend linéairement des colonnes (resp. des lignes) de* A*.*
3) Pour tout $\lambda \in K$, $\det(\lambda A) = \lambda^n \det A$.
4) $\det A \neq 0 \iff A$ *inversible.*
5) $\det A.B = \det A. \det B$
6) $\det A^{-1} = (\det A)^{-1}$

Démonstration.

1) $\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ et $\det {}^t\!A = \sum_{\rho \in S_n} \varepsilon(\rho) a_{1\rho(1)} \cdots a_{n\rho(n)}$ or, le produit $a_{1\rho(1)} \cdots a_{n\rho(n)}$ reste invariant si l'on fait subir à ses facteurs une permutation quelconque. On a donc pour $\sigma \in S_n$,

$$a_{1\rho(1)} \cdots a_{n\rho(n)} = a_{\sigma(1)(\rho\sigma)(1)} \cdots a_{\sigma(n)(\rho\sigma)(n)}$$

ainsi, en prenant $\sigma = \rho^{-1}$, on obtient

$$a_{1\rho(1)} \cdots a_{n\rho(n)} = a_{\rho^{-1}(1)1} \cdots a_{\rho^{-1}(n)n}$$

or, l'application

$$\begin{aligned} S_n &\longrightarrow S_n \\ \rho &\longmapsto \rho^{-1} \end{aligned}$$

est une bijection, on a donc

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\rho \in S_n} \varepsilon(\rho) a_{1\rho(1)} \cdots a_{n\rho(n)},$$

c'est-à-dire $\det A = \det {}^t\!A$.

2) Le déterminant est une forme linéaire alternée, il dépend donc linéairement des colonnes de A. En raison de la propriété 1, il dépend également linéairement des lignes de A.

3) Soit $\lambda \in K$, par n-linéarité du déterminant, on a

$$\det(\lambda A) = \det(\lambda c_1, \ldots, \lambda c_n) = \lambda^n \det(c_1, \ldots, c_n)$$

4) On a $\det A \neq 0 \iff$ les vecteurs $c_1, \ldots, c_n$ sont des vecteurs linéairements indépendants. Ce qui revient à dire que ces vecteurs forment une base de E, ce qui équivaut à dire que la matrice A est inversible.

5) Notons $A = (a_{ij})$ et $B = (b_{ij})$, rappelons que si $C = AB$, alors $C = (c_{ij})$ où

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

on a alors

$$\det C = \sum_{\sigma \in S_n} \varepsilon(\sigma) \left(\sum_{k=1}^{n} a_{\sigma(1)k} b_{k1} \right) \cdots \left(\sum_{k=1}^{n} a_{\sigma(n)k} b_{kn} \right)$$

et

$$\det A \det B = \sum_{\sigma, \rho \in S_n} \varepsilon(\sigma)\varepsilon(\rho) a_{\sigma(1)1} \cdots a_{\sigma(n)n} b_{\rho(1)1} \cdots b_{\rho(n)n}.$$

Un peu de gymnastique sur les permutations permet de se convaincre de l'égalité, mais nous verrons comment obtenir une autre démonstration au paragraphe suivant.

6) Si $\det A \neq 0$, on a

$$\det AA^{-1} = \det I_n = 1 = (\det A)(\det A^{-1})$$

d'où l'égalité $\det(A^{-1}) = (\det A)^{-1}$.

I.2.3 Déterminant d'un endomorphisme

Proposition 1. *Soit $u \in \mathcal{L}(E)$ et $B = (e_1, \ldots, e_n)$ une base de E, le déterminant $\det_B(u(e_1), \ldots, u(e_n))$ ne dépend pas de la base B, on le note $\det u$.*

Démonstration.

Soient $B = (e_1, \ldots, e_n)$ et $B' = (e'_1, \ldots, e'_n)$ deux bases de E et P la matrice de passage de B à B' qui exprime les coordonnées des vecteurs $(e'_j)_{1 \leq j \leq n}$ dans la base B. Si A est la matrice de u dans la base B et A' la matrice de u dans la base B', on a $A' = P^{-1}AP$ d'où

$$\det A' = \det P^{-1}AP = (\det P^{-1})(\det A)(\det P) = (\det P)^{-1}(\det A)(\det P) = \det A.$$

Proposition 2. *Pour tout $(x_1, \ldots, x_n) \in E^n$, et pour toute forme n-linéaire alternée φ, on a*

$$\varphi(u(x_1), \ldots, u(x_n)) = (\det u)\varphi(x_1, \ldots, x_n)$$

Remarques :

1) Cette proposition pourrait aussi être prise comme définition du déterminant de l'endomorphisme u.

2) Si on prend pour φ la forme n-linéaire alternée $\det_B$, alors on retrouve

$$\det u = \det_B(u(e_1), \ldots, u(e_n))$$

où $B = (e_1, \ldots, e_n)$ est une base de E.

3) L'application $\varphi \circ u$ est une forme n-linéaire alternée.

Démonstration.

Nous allons examiner trois cas.

1^{er}*cas* : On suppose le système $(x_1, \ldots, x_n)$ lié, alors le système $(u(x_1), \ldots, u(x_n))$ est également lié, on a donc pour toute forme n-linéaire alternée φ, $\varphi(u(x_1), \ldots, u(x_n)) = 0$ et $\varphi(x_1, \ldots, x_n) = 0$ d'où l'égalité.

2^{ieme}*cas* : On suppose cette fois $(x_1, \ldots, x_n)$ libre mais $(u(x_1), \ldots, u(x_n))$ lié. Le système de n vecteurs $(x_1, \ldots, x_n)$ étant libre, c'est une base de E et on peut écrire

$$\det u = \det_{(x_1, \ldots, x_n)}(u(x_1), \ldots, u(x_n))$$

mais, comme le système $(u(x_1), \ldots, u(x_n))$ est supposé lié, on a, d'une part, $\det u = 0$ et, d'autre part $\varphi(u(x_1), \ldots, u(x_n)) = 0$, d'où l'égalité.

3^{ieme}*cas* : On suppose les deux systèmes $(x_1, \ldots, x_n)$ et $(u(x_1), \ldots, u(x_n))$ libres, ce qui signifie que u est un isomorphisme. Si Δ est l'unique forme n-linéaire alternée prenant la valeur 1 sur la base $(x_1, \ldots, x_n)$ on a pour toute φ,

$$\varphi(u(x_1), \ldots, u(x_n)) = \Delta(u(x_1), \ldots, u(x_n))\varphi(x_1, \ldots, x_n),$$

or, $\Delta(u(x_1), \ldots, u(x_n)) = \det_{(x_1, \ldots, x_n)}(u(x_1), \ldots, u(x_n))$, d'où le résultat.

Nous remarquons que si $x_1, \ldots, x_n$ sont n vecteurs quelconques de E, le déterminant de $(x_1, \ldots, x_n)$ dans la base $(e_1, \ldots, e_n)$ est égal au déterminant de l'endomorphisme défini par $u(e_i) = x_i$ pour $1 \leq i \leq n$. Par ailleurs, en définissant le déterminant d'une matrice comme le déterminant de l'endomorphisme qu'elle représente dans une base donnée et en utilisant la proposition 2 pour définir le déterminant d'un endomorphisme, les propriétés du déterminant énoncées au paragraphe précédent sont faciles à démontrer (exercice).

Proposition 3. *Soient $u, v \in \mathcal{L}(E)$, on a*
(i) $\det(u \circ v) = (\det u)(\det v)$
(ii) $\det \mathrm{Id}_E = 1$

(iii) Si $u \in \mathcal{L}(E)$, on a $u \in GL(E) \iff \det u \neq 0$, dans ce cas $\det u^{-1} = (\det u)^{-1}$

Démonstration.

(i) En utilisant la proposition 2, on a d'une part,

$$\varphi(u \circ v(x_1), \ldots, u \circ v(x_n)) = \det(u \circ v)\varphi(x_1, \ldots, x_n),$$

et, d'autre part

$$\varphi(u \circ v(x_1), \ldots, u \circ v(x_n)) = (\det u)\varphi(v(x_1), \ldots, v(x_n)) = (\det u)(\det v)\varphi(x_1, \ldots, x_n).$$

ce qui prouve que $\det(u \circ v) = \det u \det v$. Remarquons que si A est la matrice de u dans une base B et A' celle de v on retrouve

$$\det(u \circ v) = \det AA' = \det A \det A' = \det u \det v.$$

(ii) Pour tout $u \in \mathcal{L}(E)$, on a $\mathrm{id}_E \circ u = u$ d'où

$$\det u \det \mathrm{id}_E = \det u$$

d'où $\det \mathrm{id}_E = 1$.

(iii) immédiat en utilisant (i) et (ii).

I.2.4. Méthodes de calcul

La formule définissant le déterminant ne rend pas très aisé le calcul de celui-ci. Nous allons donc étudier quelques méthodes pratiques de calcul.

Propriétés.
1) On ne change pas la valeur du déterminant d'une matrice en ajoutant à une colonne (resp. ligne) une combinaison linéaire des autres colonnes (resp.lignes).
2) Si on effectue une permutation sur les colonnes (ou les lignes), le déterminant est multiplié par la signature de la permutation.

Ce sont de simples rappels des propriétés des formes n-linéaires alternées, elles ont déjà été démontrées mais elles seront utiles pour les calculs de déterminants.

Théorème. *(Développement par blocs) Soit M une matrice carrée d'ordre n telle que pour $i > p$ et $j < p+1$, on ait $a_{i,j} = 0$, M est de la forme*

$$M = \begin{pmatrix} A & C \\ O & B \end{pmatrix}$$

où A est une matrice carrée d'ordre p, B une matrice carrée d'ordre $q = n - p$ et O désigne la matrice $q \times p$ dont tous les termes sont nuls. On a alors

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

Démonstration.

Notons $\mathcal{B} = (\underbrace{e_1, \ldots, e_p}_{\mathcal{B}'}, \underbrace{e_{p+1}, \ldots, e_n}_{\mathcal{B}''})$ une base de E, E' le sous-espace vectoriel de E engendré par le système $\mathcal{B}'$ et E'' le sous-espace vectoriel de E engendré par le système $\mathcal{B}''$, on a $\dim E' = p$ et $\dim E'' = q = n - p$.
Considérons les vecteurs colonnes de la matrice M :

$$a_1, \ldots, a_p, c_{p+1} + b_{p+1}, \ldots, c_n + b_n$$

avec pour $1 \leq i \leq p$, $a_i \in E'$ et pour $p+1 \leq j \leq n$, $c_j \in E'$, $b_j \in E''$.
Soit

$$\begin{aligned} f : (E')^p &\longrightarrow K \\ (x_1, \ldots, x_p) &\longmapsto \det_{\mathcal{B}}(x_1, \ldots, x_p, c_{p+1} + b_{p+1}, \ldots, c_n + b_n), \end{aligned}$$

l'application f est une forme p-linéaire alternée sur E', on a donc $\forall (x_1, \ldots, x_p) \in (E')^p$,

$$f(x_1, \ldots, x_p) = f(e_1, \ldots, e_p) \det_{\mathcal{B}'}(x_1, \ldots, x_p),$$

or,

$$f(a_1, \ldots, a_p) = \det M = f(e_1, \ldots, e_p) \det A$$

et

$$f(e_1, \ldots, e_p) = \det_{\mathcal{B}}(e_1, \ldots, e_p, c_{p+1} + b_{p+1}, \ldots, c_n + b_n).$$

Les $(c_j)_{p+1 \leq j \leq n}$ étant combinaison linéaire des $(e_i)_{1 \leq i \leq p}$ on a :

$$f(e_1, \ldots, e_p) = \det_{\mathcal{B}}(e_1, \ldots, e_p, b_{p+1}, \ldots, b_n).$$

Or, l'application

$$\begin{aligned} g : (E'')^q &\longrightarrow K \\ (x_{p+1}, \ldots, x_n) &\longmapsto \det_{\mathcal{B}}(e_1, \ldots, e_p, x_{p+1}, \ldots, x_n) \end{aligned}$$

est q-linéaire alternée donc

$$g(x_{p+1}, \ldots, x_n) = g(e_{p+1}, \ldots, e_n) \det_{\mathcal{B}''}(x_{p+1}, \ldots, x_n).$$

Mais,

$$g(b_{p+1}, \ldots, b_n) = g(e_{p+1}, \ldots, e_n) \det B$$

et

$$g(e_{p+1}, \ldots, e_n) = \det_{\mathcal{B}}(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n) = 1.$$

D'où

$$\det M = f(e_1, \ldots, e_p) \det A = g(b_{p+1}, \ldots, b_n) \det A = \det A \det B.$$

Nous allons maintenant étudier la méthode consistant à développer le déterminant suivant une ligne où une colonne.

Définitions. *Soit $A = (a_{i,j})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n. Pour tout (i,j) on appelle* mineur *de l'élément $a_{i,j}$, le déterminant $\Delta_{i,j}$ de la matrice carrée d'ordre $n-1$ obtenue en supprimant la i-ième ligne et la j-ième colonne de A. Le scalaire $A_{i,j} = (-1)^{i+j}\Delta_{i,j}$ s'appelle le* cofacteur *de $a_{i,j}$*

Théorème. *(développement suivant une ligne ou une colonne) Soit $A = (a_{i,j})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n, $A_{i,j}$ les cofacteurs des éléments de A, alors pour tout (i,j),*

$$\det A = \sum_{k=1}^{n} a_{k,j}A_{k,j} = \sum_{k=1}^{n} a_{i,k}A_{i,k}.$$

Démonstration.

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Pour $1 \leq j \leq n$ notons a_j les vecteurs colonnes de A. On a $a_j = \sum_{k=1}^{n} a_{kj}e_k$. Pour j, $1 \leq j \leq n$ fixé, la linéarité du déterminant nous donne

$$\begin{aligned}
\det A &= \det_{\mathcal{B}}(a_1, \ldots, a_j, \ldots, a_n) \\
&= \det_{\mathcal{B}}(a_1, \ldots, \sum_{k=1}^{n} a_{kj}e_k, \ldots, a_n) \\
&= \sum_{k=1}^{n} a_{kj} \underbrace{\det_{\mathcal{B}}(a_1, \ldots, a_{j-1}, e_k, a_j \ldots, a_n)}_{D_{kj}}
\end{aligned}.$$

En utilisant les propriétés des formes n-linéaires alternées, on obtient :

$$\begin{aligned}
D_{kj} &= (-1)^{j-1}\det_{\mathcal{B}}(e_k, a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n) \\
&= (-1)^{k-1}(-1)^{j-1}\det_{\mathcal{B}'}(e_k, a_1, \ldots, a_{j-1}, a_{j+1} \ldots, a_n)
\end{aligned}$$

où

$$\mathcal{B}' = (e_k, e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n).$$

On considère la matrice suivante qui exprime les vecteurs $e_k, a_1, \ldots, a_n$ dans la base $\mathcal{B}'$:

$$\begin{bmatrix} 1 & \\ 0 & \\ \vdots & M_{kj} \\ 0 & \end{bmatrix}$$

On a donc $D_{kj} = (-1)^{j+k}\det M_{kj} = A_{kj}$ et $\det A = \sum_{k=1}^{n} a_{kj}A_{kj}$.

Définition. *La matrice $(A_{i,j})_{1\leq i,j\leq n}$ des cofacteurs de A est appelée* comatrice *de A, on la note $\tilde{A}$.*

Proposition. *Soit $A \in M_n(K)$, alors*

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

d'où, si $\det A \neq 0$,

$$A^{-1} = \frac{1}{\det A}\, {}^t\tilde{A}$$

Démonstration.

Par définition de $\tilde{A}$ et ${}^t\tilde{A}$ on a $A{}^t\tilde{A} = (\gamma_{ij})_{1\leq i,j\leq n}$ où $\gamma_{ij} = \sum_{k=1}^n a_{ik}A_{jk}$.
Remarquons la chose suivante, si l'on note M la matrice obtenue à partir de A en substituant à la j-ème ligne une ligne $(\beta_1, \ldots, \beta_n)$, $\beta_j \in K$, on a $\det M = \sum_{k=1}^n \beta_k A_{jk}$, en développant par rapport à la j-ème ligne. En particulier, si dans A on remplace la j-ième ligne par la l-ième alors, si $l \neq j$, le déterminant de cette matrice est nul et si $l = j$ c'est le déterminant de A. On a donc

$$\sum_{k=1}^n a_{lk}A_{jk} = \delta_{lj} \det A.$$

Où δ_{lj} est le symbole de Kronecker

$$\delta_{lj} = \begin{cases} 1 & \text{si } l = j \\ 0 & \text{si } l \neq j \end{cases}$$

de même

$$\sum_{k=1}^n a_{kl}A_{kj} = \delta_{lj} \det A.$$

Ainsi les coefficients de $A{}^t\tilde{A}$ sont

$$\gamma_{ij} = \sum_{k=1}^n a_{ik}A_{jk} = \delta_{ij} \det A.$$

D'où ${}^t\tilde{A}.A = A.{}^t\tilde{A} = (\det A)\,\mathrm{I}_n$.

Exemple: Notons $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, on a $\tilde{A} = \begin{pmatrix} 0 & 1 & -1 \\ -3 & 2 & 1 \\ 3 & -1 & 1 \end{pmatrix}$ et ${}^t\tilde{A} = \begin{pmatrix} 0 & -3 & 3 \\ 1 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix}$.
Par ailleurs, $\det A = 3$ et on a bien

$$A{}^t\tilde{A} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -3 & 3 \\ 1 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 3I_n = (\det A)I_n.$$

I.2.5. Rang

Définitions. *Soient E et F des espaces vectoriels sur K.*
1) Si $u : E \longrightarrow F$ est une application linéaire, on appelle rang de u, la dimension du sous-espace vectoriel $\operatorname{Im} u$ de F. On le note $\operatorname{rg} u$.
2) Si $A \in M_{n,p}(K)$, on appelle rang de A, la dimension du sous-espace vectoriel engendré par les colonnes de A. On le note $\operatorname{rg} A$.
3) Si $x_1, \ldots, x_p$ sont des vecteurs de E, on appelle rang du système $(x_1, \ldots, x_p)$ la dimension du sous-espace vectoriel engendré par $(x_1, \ldots, x_p)$, c'est le nombre maximal de vecteurs linéairement indépendants du système.

Proposition 1. *Si A est la matrice d'une application linéaire u exprimée dans une base $\mathcal{B}$ de E et une base $\mathcal{B}'$ de F, on a $\operatorname{rg} u = \operatorname{rg} A$.*

Démonstration.

Si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, le sous-espace vectoriel $\operatorname{Im} u$ de F est engendré par les vecteurs $u(e_1), \ldots, u(e_n)$ qui sont les vecteurs colonnes de la matrice A de u.

Remarquons que si $A \in M_n(K)$, alors $\operatorname{rg} A = n \iff A$ est inversible.

Proposition 2. *Soit $A \in M_{n,p}(K)$,*
1) si $P \in GL_n(K)$, alors $\operatorname{rg} PA = \operatorname{rg} A$.
2) si $Q \in GL_p(K)$, alors $\operatorname{rg} AQ = \operatorname{rg} A$.

Démonstration.

1) Soient

$$\begin{aligned} f : K^p &\longrightarrow K^n \\ X = (x_1, \ldots, x_p) &\longmapsto AX \end{aligned}$$

et

$$\begin{aligned} g : K^p &\longrightarrow K^n \\ X = (x_1, \ldots, x_p) &\longmapsto PAX \end{aligned}$$

On a $\operatorname{rg} A = \operatorname{rg} f = \dim \operatorname{Im} f$ et $\operatorname{rg} PA = \operatorname{rg} g = \dim \operatorname{Im} g$. Mais, comme P est inversible, $\ker f = \ker g$. Or, $\dim \operatorname{Im} f + \dim \ker f = \dim \operatorname{Im} g + \dim \ker g = p$, d'où $\operatorname{rg} A = \operatorname{rg} PA$.

2) Notons $a_1, \ldots, a_p$ les vecteurs colonnes de la matrice A, soit $X = (x_1, \ldots, x_p)$ un vecteur de K^p. Le vecteur AX s'écrit $AX = x_1 a_1 + \cdots + x_p a_p$, il appartient donc au sous-espace vectoriel engendré par les $(a_i)_{1 \leq i \leq p}$, si l'on note $X_1, \ldots, X_p$ les vecteurs colonnes de la matrice Q, pour tout i, $1 \leq i \leq p$, le vecteur AX_i appartient au sous-espace vectoriel engendré par les vecteurs colonnes de la matrice A. Ainsi les vecteurs colonnes de la matrice AQ appartiennent-ils à ce même sous-espace vectoriel. En écrivant $A = (AQ)Q^{-1}$, pour les mêmes raisons que précédemment, on peut dire que les vecteurs colonnes de la matrice A sont dans le sous-espace vectoriel engendré par les vecteurs colonnes de la matrice AQ, ce qui prouve que $\operatorname{rg} A = \operatorname{rg} AQ$.

Théorème. *Soient $A \in M_{n,p}(K)$ et r un entier $1 \leq r \leq \min(n, p)$, notons J_r la matrice*

$$J_r = \begin{pmatrix} \mathrm{I}_r & O \\ O & O \end{pmatrix}$$

où I_r *est la matrice unité* $r \times r$ *et* O *désigne une matrice nulle. Alors*

$$\operatorname{rg} A = r \iff \exists P \in GL_n(K), Q \in GL_p(K) \ / PAQ = J_r$$

Démonstration.

($\Leftarrow$) Il est clair que $\operatorname{rg} J_r = r$, par conséquent, s'il existe $P \in GL_n(K), Q \in GL_p(K)$ telles que $PAQ = J_r$, comme d'après la proposition 2, $\operatorname{rg} PAQ = \operatorname{rg} A$, on a $\operatorname{rg} A = \operatorname{rg} J_r = r$.

($\Rightarrow$) Supposons $\operatorname{rg} A = r$, notons $u : K^p \longrightarrow K^n$ l'application linéaire de matrice A dans les bases canoniques. On a $\operatorname{rg} A = \operatorname{rg} u = \dim \operatorname{Im} u$. Or, $\dim \operatorname{Im} u + \dim \ker u = p$. Il existe une base de K^p, $(x_1, \ldots, x_p)$, dont les $p - r$ derniers vecteurs forment une base de $\ker u$. Pour tout i, $1 \leq i \leq r$, notons $y_i = u(x_i)$, les vecteurs $y_1, \ldots, y_r$ engendrent $\operatorname{Im} u$. Par le théorème de la base incomplète, il existe des vecteurs $y_{r+1}, \ldots, y_n$ tels que les $(y_i)_{1 \leq i \leq n}$ forment une base de K^n. La matrice de u exprimée dans les bases $(x_1, \ldots, x_p)$, $(y_1, \ldots, y_n)$ est la matrice J_r.
La matrice A étant la matrice de u exprimée dans les bases canoniques de K^n et K^p, si P est la matrice qui exprime les vecteurs de la base canonique de K^n dans la base $(y_1, \ldots, y_n)$ et Q la matrice qui exprime les vecteurs de la base $(x_1, \ldots, x_p)$ dans la base canonique de K^p, alors la matrice PAQ est la matrice de u exprimée dans les bases $(x_1, \ldots, x_p)$ et $(y_1, \ldots, y_n)$, c'est-à-dire que ses colonnes sont les coordonnées des vecteurs $u(x_1), \ldots, u(x_p)$ dans la base $(y_1, \ldots, y_n)$, on a donc $PAQ = J_r$.

Corollaire. *Soit* $A \in M_{n,p}(K)$, *on a* $\operatorname{rg} {}^tA = \operatorname{rg} A$.

Démonstration.

Notons $r = \operatorname{rg} A$, alors, d'après le théorème précédent, il existe une matrice $P \in GL_n(K)$ et une matrice $Q \in GL_p(K)$ telles que $PAQ = J_r$. On a alors ${}^t(PAQ) = {}^tJ_r$, c'est-à-dire ${}^tQ{}^tA{}^tP = {}^tJ_r$. Ainsi, il existe une matrice ${}^tQ \in GL_p(K)$ et une matrice ${}^tP \in GL_n(K)$ telles que ${}^tQ{}^tA{}^tP = {}^tJ_r$, or la matrice ${}^tJ_r$ est dans $M_{p,n}(K)$ du même type que J_r dans $M_{n,p}(K)$, ce qui prouve, en raison de l'équivalence démontrée dans le théorème, que $\operatorname{rg} {}^tA = r$.

Théorème. *Deux matrices de* $M_{n,p}(K)$ *sont équivalentes si et seulement si elles ont même rang.*

Démonstration.

Rappelons que deux matrices A et B de $M_{n,p}(K)$ sont dites équivalentes, noté $A \sim B$, s'il existe $P \in GL_n(K)$ et $Q \in GL_p(K)$ telles que $B = PAQ$.
($\Rightarrow$) Soient A et $B \in M_{n,p}(K)$, on suppose $A \sim B$, il existe alors $P \in GL_n(K)$ et $Q \in GL_p(K)$ telles que $B = PAQ$, on a donc $\operatorname{rg} B = \operatorname{rg} PAQ = \operatorname{rg} AQ = \operatorname{rg} A$.
($\Leftarrow$) Supposons que A et B aient même rang, alors $A \sim J_r$ et $B \sim J_r$ d'où $A \sim B$.

Nous allons maintenant étudier le lien entre le rang et le déterminant. Remarquons que dans la matrice J_r, qui est de rang r par définition, le plus "grand" déterminant extrait non nul est de rang r, nous allons préciser cette notion.

Définition. *Soit* $A = (a_{i,j})_{\substack{1\le i\le n\\ 1\le j\le p}} \in M_{n,p}(K)$. *Soient* $I \subset \{1,\ldots,n\}$ *et* $J \subset \{1,\ldots,p\}$. *La matrice* $B = (a_{i,j})_{\substack{i\in I\\ j\in J}}$ *s'appelle matrice* extraite *de* A. *On appelle déterminant* extrait *de* A *le déterminant d'une matrice carrée extraite de* A.

Théorème. *Soient* $A \in M_{n,p}(K)$ *et* r *un entier* $1 \le r \le \min(n,p)$, *le rang de* A *est supérieur ou égal à* r *si et seulement si il existe une matrice carrée d'ordre* r *inversible extraite de* A.

Démonstration.

($\Rightarrow$) Supposons $\operatorname{rg} A \ge r$, il existe alors r colonnes de A linéairement indépendantes. On peut supposer que ce sont les r premières, notées $c_1,\ldots,c_r$, car le rang n'est pas affecté par une permutation des colonnes. On peut compléter le système $(c_1,\ldots,c_r)$ par des vecteurs de la base canonique de K^n, notons les $e'_{r+1},\cdots,e'_n \in \{e_1,\ldots e_n\}$, de façon à obtenir une autre base de K^n. On a alors $\det_{(e_1,\ldots,e_n)}(c_1,\ldots c_r,e'_{r+1},\ldots,e'_n) \neq 0$. En développant ce déterminant, successivement, par rapport à la dernière colonne, on obtient donc un déterminant $r\times r$ non nul qui est extrait de la matrice des colonnes $c_1,\ldots,c_r$ donc de A.
($\Leftarrow$) Supposons qu'il existe une matrice $B \in M_r(K)$ extraite de A et inversible. Notons $c_1,\ldots,c_r$ les vecteurs colonnes de A correspondants, ces vecteurs sont alors linéairement indépendants et la matrice A est de rang $\ge r$.

Corollaire. *Le rang d'une matrice est l'ordre maximum d'un déterminant extrait non nul de cette matrice.*

Démonstration.

Soit M une matrice, la matrice carrée d'ordre 1 déterminée par un élément non nul de M est inversible, par conséquent l'ensemble des ordres des sous-matrices carrées inversibles est non vide, il est majoré par r, il admet donc un plus grand élément qui ne peut-être que r.

Exemple : Soit $A = \begin{pmatrix} 5 & -1 & -3\\ 2 & 1 & -4\\ 3 & -2 & 1\\ 1 & 2 & -5\end{pmatrix}$, c'est la matrice d'une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^4$. Le rang d'une matrice n'est pas affecté par les opérations linéaires sur les colonnes, on a donc

$$\operatorname{rg} A = \operatorname{rg}\begin{pmatrix} 1 & 5 & -3\\ -1 & 2 & -4\\ 2 & 3 & 1\\ -2 & 1 & -5\end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 0 & 0\\ -1 & 7 & -7\\ 2 & -7 & 7\\ -2 & 11 & -11\end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 0 & 0\\ -1 & 7 & 0\\ 2 & -7 & 0\\ -2 & 11 & 0\end{pmatrix} = 2.$$

En effet, tous les déterminants 3×3 sont nuls mais pas tous les déterminants 2×2.

II. Réduction des endomorphismes

Soit u un endomorphisme de E, il s'agit de trouver une base de E dans laquelle la matrice de u s'exprime de la manière la plus simple possible. Pour cela nous allons nous intéresser aux droites (sous-espaces vectoriels de dimension 1) de E invariantes par u. Par exemple 1) si l'endomorphisme u est défini par

$$u\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

c'est-à-dire si u est une homothétie, toute droite est invariante par u.

2) si l'endomorphisme u est défini

$$u\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix}$$

seuls les axes Ox et Oy sont invariants par u.
Si une droite $D \subset E$ est invariante par un endomorphisme u, alors pour tout vecteur non nul $x \in D$, il existe $\lambda \in K$ tel que $u(x) = \lambda x$. C'est ce qui nous amène à la notion de vecteurs propres et valeurs propres.

II.1. Valeurs propres, vecteurs propres

Définition 1. *Soit $\lambda \in K$, λ est dite valeur propre de l'endomorphisme u si il existe un vecteur non nul $x \in E$ tel que*

$$u(x) = \lambda x,$$

le vecteur x est alors appelé vecteur propre de u pour la valeur propre λ.

On remarque que le vecteur x est vecteur propre de l'endomorphisme u pour la valeur propre λ si et seulement si $(u - \lambda \,\mathrm{id}_E)(x) = 0$, c'est-à-dire $x \in \ker(u - \lambda \,\mathrm{id}_E)$.

Définition 2. *Si λ est une valeur propre de u, on note E_λ l'ensemble des vecteurs $x \in E$ tels que $u(x) = \lambda x$, c'est le sous-espace vectoriel de E constitué des vecteurs propres de u associés à la valeur propre λ et du vecteur nul. On a*

$$E_\lambda = \ker(u - \lambda \,\mathrm{id}_E),$$

c'est le sous-espace propre associé à la valeur propre λ.

Il est clair que λ est valeur propre de u si et seulement si $E_\lambda \neq \{0\}$, c'est-à-dire si et seulement si l'endomorphisme $u - \lambda \,\mathrm{id}_E$ n'est pas injectif, ce qui équivaut à $u - \lambda \,\mathrm{id}_E$ non bijectif car E est supposé de dimension finie.

Proposition. *Soit P un polynôme et x un vecteur propre de u pour la valeur propre λ, alors, x est vecteur propre de $P(u)$ pour la valeur propre $P(\lambda)$.*

Rappelons que si $P(X) = \sum_{k=0}^{n} a_k X^k$ est un polynôme et si u est un endomorphisme de E, on note $P(u)$ l'endomorphisme

$$P(u) = a_0 \operatorname{id}_E + a_1 u + a_2 u^2 + \cdots + a_n u^n$$

qui à un vecteur x associe le vecteur

$$P(u)(x) = a_0 x + a_1 u(x) + a_2 u^2(x) + \cdots + a_n u^n(x),$$

où pour $1 \leq k \leq n$, on a $u^k = \underbrace{u \circ u \cdots \circ u}_{k\times}$.

Démonstration de la proposition.

Nous allons commencer par démontrer par récurrence sur k que pour tout $k \in \mathbb{N}$ et pour tout x vecteur propre de u pour la valeur propre λ on a $u^k(x) = \lambda^k x$.
La propriété est clairement vraie pour $k = 0$ et $k = 1$, supposons la vraie pour un k arbitrairement fixé et calculons alors $u^{k+1}(x)$:

$$u^{k+1}(x) = u(u^k(x)) = u(\lambda^k x) = \lambda^k u(x) = \lambda^{k+1} x,$$

d'où le résultat. Si maintenant $P(X) = \sum_{k=0}^{n} a_k X^k$, on a

$$P(u)(x) = \sum_{k=0}^{n} a_k u^k(x) = \sum_{k=0}^{n} a_k \lambda^k x = \left(\sum_{k=0}^{n} a_k \lambda_k \right) x = P(\lambda) x.$$

Ce qui prouve que x est un vecteur propre de l'endomorphisme $P(u)$ pour la valeur propre $P(\lambda)$.

Théorème. *Soient $\lambda_1, \ldots, \lambda_r$ des valeurs propres distinctes de u, alors si pour $1 \leq i \leq r$ x_i est un vecteur propre pour λ_i, les x_i sont linéairement indépendants.*

Démonstration.

On suppose $r \geq 2$. Notons, pour $1 \leq i \leq r$,

$$P_i(X) = (X - \lambda_1) \cdots (X - \lambda_{i-1})(X - \lambda_{i+1}) \cdots (X - \lambda_r) = \prod_{1 \leq j \leq r, j \neq i} (X - \lambda_j).$$

Pour tout i et pour tout j, $i \neq j$, on a $P_i(\lambda_j) = 0$ et, puique les valeurs propres sont supposées distinctes, $P_i(\lambda_i) \neq 0$. D'après la proposition précédente, le vecteur x_j est vecteur propre de $P_i(u)$ pour la valeur propre $P_i(\lambda_j) = 0$, on a donc $P_i(u)(x_j) = 0$. De même, x_i est vecteur propre de $P_i(u)$ pour la valeur propre $P_i(\lambda_i) \neq 0$, d'où, pour tout i, $1 \leq i \leq r$, on a $P_i(u)(x_i) = P_i(\lambda_i) x_i$.

Supposons qu'il existe des scalaires $a_1, \dots, a_r \in K$, tels que

$$a_1 x_1 + \cdots + a_r x_r = 0,$$

on a alors

$$P_i(u)(a_1 x_1 + \cdots + a_r x_r) = 0,$$

c'est-à-dire

$$a_1 P_i(u)(x_1) + \cdots + a_r P_i(u)(x_r) = a_i P_i(\lambda_i)(x_i) = 0.$$

Ce qui implique $a_i = 0$, car $P_i(\lambda_i) \neq 0$ et $x_i \neq 0$. Ceci étant vrai pour tout i on a obtenu $a_1 = a_2 = \ldots = a_r = 0$, ce qui prouve que les vecteurs $x_1, \dots, x_r$ sont linéairement indépendants.

Corollaire. *Si $\lambda_1, \dots, \lambda_r$ sont des valeurs propres distinctes, alors les sous-espaces propres correspondants, $E_{\lambda_1}, \dots, E_{\lambda_r}$ sont en somme directe.*

Démonstration.

Pour chaque i, $1 \leq i \leq r$, soit $x_i \in E_{\lambda_i}$. On suppose $x_1 + \cdots + x_r = 0$, nous allons montrer qu'alors $x_1 = x_2 = \ldots = x_r = 0$.
Notons r' le nombre minimal tel qu'il existe r' tels vecteurs non nuls, on peut supposer sans perdre de généralité que ce sont les r' premiers, c'est-à-dire que l'on a $x_1 + \cdots + x_{r'} = 0$, on a alors

$$u(x_1 + \cdots + x_{r'}) - \lambda_1(x_1 + \cdots + x_{r'}) = 0$$

c'est-à-dire

$$(\lambda_2 - \lambda_1)x_2 + \cdots + (\lambda_{r'} - \lambda_1)x_{r'} = 0.$$

Or, les vecteurs $x_2, \dots, x_{r'}$ sont linéairement indépendants et les valeurs propres sont supposées distinctes, de plus, r' est supposé minimal et l'on vient d'obtenir $r' - 1$ vecteurs non nuls dont la somme est nulle, cette contradiction prouve que les vecteurs $x_2, \dots, x_{r'}$ sont nuls et donc que tous les $(x_i)_{1 \leq i \leq r}$ sont nuls, la somme est donc directe.

II.2. Polynôme caractéristique

Proposition 1. *Soit u un endomorphisme de E et A sa matrice dans une base $\mathcal{B}$, on appelle polynôme caractéristique de u, ou de A le polynôme, qui ne dépend pas de la base,*

$$P_u(X) = P_A(X) = \det(A - X.I_n) = \det(u - X\,\mathrm{id}_E).$$

De plus, on a

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

Avant de démontrer ce théorème, voyons le cas $n = 2$, soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a

$$P_A(X) = \begin{vmatrix} a - X & b \\ c & d - X \end{vmatrix} = (a - X)(d - X) - bc = (-1)^2 X^2 + (-1)\underbrace{(a + d)}_{\operatorname{tr} A} X + \underbrace{ad - bc}_{\det A}.$$

Démonstration.

Vérifions tout d'abord que ce polynôme ne dépend pas de la base $\mathcal{B}$. Si B est la matrice de u dans une autre base, il existe une matrice inversible $P \in GL_n(E)$ telle que $B = P^{-1}AP$. D'où

$$B - XI_n = P^{-1}AP - X(P^{-1}I_nP) = P^{-1}(A - XI_n)P,$$

d'où

$$\det(B - XI_n) = \det(A - XI_n).$$

Par ailleurs, si $A = (a_{i,j})_{1 \leq i,j \leq n}$, on a

$$\begin{aligned} P_A(X) &= \begin{vmatrix} a_{11} - X & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - X & & \\ \vdots & & & \\ a_{n1} & \cdots & \cdots & a_{nn} - X \end{vmatrix} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \text{ où } b_{ij} = a_{ij} \text{ si } i \neq j, b_{ii} = a_{ii} - X \\ &= (a_{11} - X) \cdots (a_{nn} - X) + \sum_{\sigma \in S_n, \sigma \neq \mathrm{id}} \varepsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n}. \end{aligned}$$

Or, si $\sigma \neq \mathrm{id}$, il y a au plus $n - 2$ entiers k tels que $\sigma(k) = k$, et, donc, le polynôme

$$\sum_{\sigma \in S_n, \sigma \neq \mathrm{id}} \varepsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n}$$

est de degré au plus $n - 2$. Le polynôme $P_A(X)$ étant de degré n, ses termes de degré n et $n - 1$ proviennent du produit

$$(a_{11} - X) \cdots (a_{nn} - X) = (-1)^n X^n + (-1)^{n-1}(\operatorname{tr} A)X^{n-1} + \ldots.$$

Le terme constant, quant à lui est donné par $P_A(0) = \det A$.

Proposition 2. *Soit $\lambda \in K$, alors*

$$\lambda \text{ valeur propre de } u \iff P_u(\lambda) = 0.$$

Démonstration.

On a vu que λ est valeur propre de u si et seulement si l'endomorphisme $u - \lambda\,\mathrm{id}_E$ est non bijectif, ce qui équivaut à $\det(u - \lambda\,\mathrm{id}_E) = 0$, c'est-à-dire $P_u(\lambda) = 0$.

Remarquons que si E est de dimension n, le polynôme caractéristique d'un endomorphisme de E est de degré n, il a donc au plus n racines, l'endomorphisme u a donc au plus n valeurs propres. Si le corps $K = \mathbb{C}$, c'est-à-dire s'il est algébriquement clos, alors tout endomorphisme admet au moins une valeur propre.

Théorème. *Soient u un endomorphisme de E et P_u son polynôme caractéristique. Soient λ une valeur propre de u, $m(\lambda)$ sa multiplicité dans P_u et E_λ le sous-espace propre associé, alors on a*

$$1 \leq \dim E_\lambda \leq m(\lambda)$$

Démonstration.

Rappelons que si λ est une racine de P_u de multiplicité m, on a

$$P_u(X) = (\lambda - X)^m Q(X)$$

avec $Q(\lambda) \neq 0$.

Soit λ une valeur propre de u et E_λ son sous-espace propre, on note $p = \dim E_\lambda$ et $(e_1, \ldots, e_p)$ une base de E_λ. On complète cette base en une base de E $(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$. Dans cette base, la matrice de u est de la forme

$$A = \begin{bmatrix} \lambda I_p & C \\ 0 & B \end{bmatrix},$$

d'où $\det(A - XI_n) = \det((\lambda - X)I_p)\det(B - XI_{n-p}) = (\lambda - X)^p \det(B - XI_{n-p})$.

Ce qui prouve que $(\lambda - X)^p$ divise P_u, donc, par définition de la multiplicité d'une racine, on a $p \leq m(\lambda)$. Par aileurs, par définition d'une valeur propre et d'un sous-espace propre, on a $\dim E_\lambda \geq 1$.

II.3. Endomorphismes diagonalisables.

Définition 1. *On dit qu'une matrice* $A = (a_{i,j})_{1 \leq i,j \leq n}$ *est diagonale si* $a_{i,j} = 0$ *dès que* $i \neq j$.

Définition 2. *On dit que l'endomorphisme* u *de* E *est diagonalisable, s'il existe une base de* E *formée de vecteurs propres.*

Définition 3. *Soit* A *une matrice d'ordre* n*, on dit que* A *est diagonalisable sur* K *si il existe une matrice* $P \in M_n(K)$ *inversible telle que* $P^{-1}AP$ *soit diagonale.*

Théorème. *Soit* $u \in \mathcal{L}(E)$*, les propositions suivantes sont équivalentes*
(i) u *est diagonalisable*
(ii) P_u *a toutes ses racines dans le corps* K *et pour toute racine* λ *de* P_u*, l'ordre de multiplicité de* λ *est égal à la dimension du sous-espace propre correspondant* E_λ.
(iii) Il existe des valeurs propres $\lambda_1, \dots, \lambda_r$ *telles que*

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}.$$

Démonstration.

$(i) \Rightarrow (ii)$ Supposons u diagonalisable et notons $\lambda_1, \dots, \lambda_r$ ses valeurs propres et $m_1, \dots, m_r$ leur multiplicité. Soit A une matrice diagonale de u, on suppose que chaque $(\lambda_i)_{1 \leq i \leq r}$ apparaît h_i fois dans la diagonale de A, on a alors,

$$P_A(X) = P_u(X) = \prod_{i=1}^{r} (\lambda_i - X)^{h_i}.$$

Ce qui prouve que P_u se décompose en facteurs linéaires dans $K[X]$ et que $h_i = m_i$. Par ailleurs, pour tout $i, 1 \leq i \leq r$, il existe h_i vecteurs de la base de E vérifiant $u(x) = \lambda_i x$, il existe donc h_i vecteurs linéairement indépendants dans E_{λ_i}, d'où $\dim E_{\lambda_i} \geq h_i$, or, $h_i = m_i$ et l'on a démontré que $\dim E_{\lambda_i} \leq m_i = h_i$, d'où l'égalité.

$(ii) \Rightarrow (iii)$ On suppose que P_u a toutes ses racines dans K et que pour toute racine $(\lambda_i)_{1 \leq i \leq r}$ de multiplicité m_i on a $\dim E_{\lambda_i} = m_i$. On a alors

$$P_u(X) = \prod_{i=1}^{r} (\lambda_i - X)^{m_i}.$$

Notons $F = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, on sait que les sous-espaces propres sont en somme directe. On a $\dim F = \sum_{i=1}^{r} m_i = \deg P_u = n$ d'où $F = E$.

$(iii) \Rightarrow (i)$ Pour tout i, $1 \leq i \leq r$, on note $\mathcal{B}_i$ une base de E_{λ_i}, la base $\mathcal{B} = \cup_{i=1}^{r} \mathcal{B}_i$ est une base de E puisque E est somme directe des E_{λ_i}. Ainsi, il existe une base de E formée de vecteurs propres de u, ce qui prouve que u est diagonalisable.

Exemples :

1) Diagonalisons la matrice

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

Pour cela on détermine ses valeurs propres :

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & -2 \\ 1 & 1-\lambda \end{vmatrix} = (4-\lambda)(1-\lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

Ainsi, la matrice A admet deux valeurs propres distinctes, qui sont $\lambda_1 = 2$ et $\lambda_2 = 3$. Elle est diagonalisable. Déterminons une base de vecteurs propres :

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \iff x = y,$$

d'où le vecteur propre $u_1 = (1, 1)$ associé à la valeur propre $\lambda_1 = 2$.

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix} \iff x = 2y,$$

d'où le vecteur propre $u_2 = (2, 1)$ associé à la valeur propre $\lambda_2 = 3$. Dans la base (u_1, u_2), la matrice s'écrit

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

On a $A = PA'P^{-1}$ où

$$P = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}.$$

2) Soit

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

Démontrons que A est diagonalisable et trouvons une matrice P telle que $P^{-1}AP$ soit diagonale.

Commençons par calculer le polynôme caractéristique de A :

$$P_A(X) = \begin{vmatrix} 1-X & 0 & 0 \\ 0 & 1-X & 0 \\ 1 & -1 & 2-X \end{vmatrix} = (1-X)^2(2-X)$$

Les racines du polynôme caractéristique sont les réels 1 avec la multiplicité 2, et 2 avec la multiplicité 1.

Déterminons les sous-espaces propres associés : Soit E_1 le sous-espace propre associé à la valeur propre double 1.

$E_1 = \{V(x, y, z) \in \mathbb{R}^3 / \ A.V = V\}$,

$$V \in E_1 \iff \begin{cases} x = x \\ y = y \\ x - y + z = 0 \end{cases} \iff x - y + z = 0$$

E_1 est donc un plan vectoriel, dont les vecteurs $e_1 = (1, 1, 0)$ et $e_2 = (0, 1, 1)$ forment une base.

Soit E_2 le sous-espace propre associé à la valeur propre simple 2.

$E_2 = \{V(x, y, z) \in \mathbb{R}^3 / \ A.V = 2V\}$,

$$V \in E_2 \iff \begin{cases} x = 2x \\ y = 2y \\ x - y + 2z = 2z \end{cases} \iff x = 0, y = 0$$

E_2 est donc une droite vectorielle, dont le vecteur $e_3 = (0, 0, 1)$ est une base.

Les dimensions des sous-espaces propres sont égales aux la multiplicités des valeurs propres correspondantes, la matrice A est donc diagonalisable. Dans la base (e_1, e_2, e_3) l'endomorphisme représenté par A (dans la base canonique) a pour matrice

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

la matrice de passage

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

vérifie $P^{-1}AP = D$.

II.4. Trigonalisation

Il est clair que tous les endomorphismes ne sont pas diagonalisables, on peut néanmoins pour certains d'entre eux trouver une base de E dans laquelle la matrice est triangulaire supérieure.

Définition 1. *On dit qu'une matrice $A = (a_{i,j})_{1 \leq i,j \leq n}$ est triangulaire supérieure si $a_{i,j} = 0$ dès que $i > j$.*

Définition 2. *On dit qu'un endomorphisme u (ou une matrice A) est trigonalisable s'il existe une base dans laquelle la matrice de u soit triangulaire (ou s'il existe une matrice inversible P telle que $P^{-1}AP$ soit triangulaire).*

Théorème. *L'endomorphisme u est trigonalisable si et seulement si son polynôme caractéristique P_u a toutes ses racines dans K (i.e. est scindé sur K ou se décompose en facteurs linéaires sur $K[X]$).*

Remarquons que si $K = \mathbb{C}$, tout endomorphisme est trigonalisable, ce n'est évidemment pas le cas si $K = \mathbb{R}$.

Démonstration du théorème.

($\Rightarrow$) Si u est trigonalisable, il existe une base dans laquelle la matrice de u s'écrit

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

On a alors,

$$P_u(X) = P_A(X) = \prod_{i=1}^{n} (a_{ii} - X),$$

ce qui prouve que P_u a toutes ses racines dans K.

($\Leftarrow$) La démonstration se fait par récurrence sur n. Si $n = 1$, il n'y a rien à démontrer. Supposons le résultat vrai pour $n-1$, n étant arbitrairement fixé. Le polynôme P_u ayant au moins une racine dans K, notons λ l'une d'entre elles et x_1 un vecteur propre associé. Soit H l'hyperplan supplémentaire de la droite Kx_1. On considère alors une base de E, $(x_1, x_2, \ldots, x_n)$ avec pour $2 \leq i \leq n$, $x_i \in H$. La matrice de u dans cette base s'écrit

$$\begin{pmatrix} \lambda & \cdots & \\ 0 & & \\ \vdots & M & \\ 0 & & \end{pmatrix}$$

où M est une matrice carrée d'ordre $n-1$. On a

$$P_u(X) = (\lambda - X)\det(M - X\,\mathrm{id}_H) = (\lambda - X)P_M(X).$$

Notons v l'endomorphisme de H dont la matrice dans la base $(x_2, \ldots, x_n)$ est égale à M. Par hypothèse de récurrence, M est trigonalisable, en effet,

$$P_u(X) = (\lambda - X)P_v(X) = (\lambda - X)P_M(X),$$

et comme P_u est supposé scindé, P_v l'est également. Par conséquent, il existe une base $(y_2, \ldots, y_n)$ de H dans laquelle la matrice de v est triangulaire, ainsi dans la base $(x_1, y_2, \ldots, y_n)$ la matrice de u est triangulaire.

Exemple : Soit

$$A = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}$$

Démontrons que A est trigonalisable et trouvons une matrice P telle que $P^{-1}AP$ soit triangulaire supérieure.
Commençons par calculer le polynôme caractéristique de A :

$$P_A(X) = \begin{vmatrix} 1-X & 4 & 6 \\ 0 & 6-X & -3 \\ -1 & 4 & -X \end{vmatrix} = (3-X)(2-X)^2$$

Les racines du polynôme caractéristique sont les réels 3 avec la multiplicité 1, et 2 avec la multiplicité 2. Les racines sont dans le corps $\mathbb{R}$, la matrice est trigonalisable et peut-être diagonalisable.

Déterminons les sous-espaces propres associés : Soit E_3 le sous-espace propre associé à la valeur propre simple 3.

$E_3 = \{V(x, y, z) \in \mathbb{R}^3 / \ A.V = 3V\}$,

$$V \in E_3 \iff \begin{cases} x + 4y - 2z = 3x \\ \quad 6y - 3z = 3y \\ \quad -x + 4y = 3z \end{cases} \iff x = y = z$$

E_3 est donc la droite vectorielle engendrée par le vecteur $e_1 = (1, 1, 1)$.

Soit E_2 le sous-espace propre associé à la valeur propre double 2.

$E_2 = \{V(x, y, z) \in \mathbb{R}^3 / \ A.V = 2V\}$,

$$V \in E_2 \iff \begin{cases} x + 4y - 2z = 2x \\ \quad 6y - 3z = 2y \\ \quad -x + 4y = 2z \end{cases} \iff \begin{cases} x = z \\ 4y = 3z \end{cases}$$

E_2 est donc la droite vectorielle engendrée par $e_2 = (4, 3, 4)$.

La dimension de E_2 est égale à 1 alors que la multiplicité de la valeur propre 2 correspondante est égale à 2, par conséquent la matrice A n'est pas diagonalisable.

Soit e_3 le vecteur de la base canonique $e_3 = (0, 0, 1)$, les vecteurs e_1, e_2, e_3 forment une base de E. La matrice de passage est

$$P = \begin{pmatrix} 1 & 4 & 0 \\ 1 & 3 & 0 \\ 1 & 4 & 1 \end{pmatrix} \quad \text{et sa matrice inverse est} \quad P^{-1} = \begin{pmatrix} -3 & 4 & 0 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

On a $Ae_1 = 3e_1$ et $Ae_2 = 2e_2$, il reste à exprimer Ae_3 dans la base (e_1, e_2, e_3).

$$Ae_3 = Ak = -2i - 3j = -2(-3e_1 + e_2 - e_3) - 3(4e_1 - e_2) = -6e_1 + e_2 + 2e_3.$$

Ainsi l'endomorphisme qui a pour matrice A dans la base canonique (i, j, k) de E a pour matrice A' dans la base (e_1, e_2, e_3) où

$$A' = \begin{pmatrix} 3 & 0 & -6 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

On a $A' = P^{-1}AP$.

II.5. Sous-espaces stables

Définition. *Soient V un sous-espace vectoriel de E et u un endomorphisme de E. Le sous-espace vectoriel V est dit stable par u si pour tout vecteur $x \in V$, on a $u(x) \in V$.*

Remarquons que les sous-espaces propres de u sont évidemment stables par u.

Proposition 1. *Si V est un sous-espace vectoriel stable par u alors pour tout polynôme $P \in K[X]$, V est stable par $P(u)$.*

Démonstration.

On montre par récurrence sur n que pour tout $x \in V$, $u^{(n)}(x) \in V$, ainsi on a bien $P(u)(x) \in V$. En effet, $P(X) = \sum_{k=0}^{n} a_k X^k$ est un polynôme et si u est un endomorphisme de E, on note $P(u)$ l'endomorphisme

$$P(u) = a_0 \operatorname{id}_E + a_1 u + a_2 u^2 + \cdots + a_n u^n$$

qui à un vecteur x associe le vecteur

$$P(u)(x) = a_0 x + a_1 u(x) + a_2 u^2(x) + \cdots + a_n u^n(x),$$

ainsi, si pour tout $x \in V$, $u^{(n)}(x) \in V$, on a bien $P(u)(x) \in V$ car V est un espace vectoriel.

Proposition 2. *Soient u et v deux endomorphismes de E tels que $u \circ v = v \circ u$, alors* $\ker v$ *et* $\operatorname{Im} v$ *sont stables par u.*

Démonstration.

Soit $x \in \ker v$, on a $v(x) = 0$, d'où $v(u(x)) = u(v(x)) = 0$, donc $u(x) \in \ker v$.

Soit $y \in \operatorname{Im} v$, il existe $x \in E$ tel que $y = v(x)$, on a alors $u(y) = u(v(x)) = v(u(x))$ donc $u(y) \in \operatorname{Im} v$.

Proposition 3. *Soit V un sous-espace vectoriel non nul de E stable par u, alors le polynôme caractéristique de $u_{|V}$ divise P_u. De plus, si W est un sous-espace vectoriel non nul supplémentaire de V et stable par u on a*

$$P_u = P_{u_{|V}}.P_{u_{|W}}.$$

Démonstration.

On considère une base de V, $(e_1, \ldots, e_p)$, et on la complète en une base $(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ de E. La matrice de u dans cette base est de la forme

$$M = \begin{pmatrix} A & U \\ 0 & B \end{pmatrix}$$

où $A \in M_p(K)$ est la matrice de $u_{|V}$ dans la base $(e_1, \ldots, e_p)$. On a alors

$$\begin{aligned} P_u(X) &= \det(M - XI_n) = \det(u - X\,\mathrm{id}_E) \\ &= \det(A - XI_p)\det(B - XI_{n-p}). \end{aligned}$$

Or, $\det(A - XI_p) = \det(u_{|V} - X\,\mathrm{id}_V) = P_{u_{|V}}(X)$, ce qui prouve que $P_{u_{|V}}$ divise P_u.

Si maintenant on suppose que les vecteurs $(e_{p+1}, \ldots, e_n)$ forment une base de W et que W est supposé stable par u, alors, la matrice U est nulle et l'on a l'égalité

$$P_u = P_{u_{|V}}.P_{u_{|W}}.$$

Proposition 4. *On suppose u diagonalisable, on note $\lambda_1, \cdots, \lambda_r$ ses valeurs propres et $E_{\lambda_1}, \cdots, E_{\lambda_r}$ les sous-espaces propres correspondants. Soit V stable par u, alors on a*

$$V = V \cap E_{\lambda_1} \oplus \cdots \oplus V \cap E_{\lambda_r}$$

Démonstration.

Soit $x \in V$, comme $x \in E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, il existe vecteurs $x_1, \ldots, x_r$, avec pour $1 \leq i \leq r$, $x_i \in E_{\lambda_i}$ tels que $x = x_1 + \cdots + x_r$.

Le sous-espace V est stable par u, donc également par $P(u)$ pour tout $P \in K[X]$. Pour $1 \leq i \leq r$, on note

$$P_i(X) = \prod_{k=1, k \neq i}^{r} (\lambda_k - X).$$

On a $P_i(\lambda_j) = 0$ si $i \neq j$ et $P_i(\lambda_i) \neq 0$. On peut alors écrire

$$\begin{aligned} P_i(u)(x) &= P_i(u)(x_1 + \cdots + x_r) \\ &= P_i(\lambda_1)x_1 + \cdots + P_i(\lambda_r)x_r \\ &= P_i(\lambda_i)x_i. \end{aligned}$$

Or, $P_i(u)(x) \in V$ par stabilité, donc $x_i \in V$. Ainsi, pour tout i, $1 \leq i \leq r$, $x_i \in V \cap E_{\lambda_i}$, d'où le résultat.

II.6. Le Théorème de Hamilton-Cayley

Théorème. *1) Soient u un endomorphisme de E et P_u son polynôme caractéristique, on a $P_u(u) = 0$.*
2) Soient A une matrice carrée d'ordre n et P_A son polynôme caractéristique, on a $P_A(A) = 0$.

Avant de commencer la démonstration, rappelons que si A est la matrice de l'endomorphisme u dans la base $(e_1, \dots, e_n)$ de E, le polynôme $P_A(X) = \det(A - XI_n)$ ne dépend pas de la base, c'est aussi le polynôme $P_u(X) = \det(u - X\,\mathrm{id}_E)$. De même que l'on a défini précédemment un polynôme d'endomorphisme, on peut définir un polynôme de matrices, si $P(X) = \sum_{k=0}^{d} a_k X^k$, on a $P(A) = \sum_{k=0}^{d} a_k A^k$ où $A^0 = I_n$.

Exemple : En dimension 2, posons

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_n(\mathbb{R}).$$

On a

$$P_A(X) = X^2 - (a+d)X + ad - bc.$$

D'où

$$\begin{aligned} P_A(A) &= A^2 - (a+d)A + (ad-bc)I_2 \\ &= \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Démonstration du théorème.

On commence dans un premier temps par supposer que le polynôme $P_A = P_u$ a toutes ses racines dans K. Nous démontrons dans ce cas particulier le résultat du théorème par récurrence sur n.

Pour $n = 1$, $P_A(X) = -X + a$, $A = [a]$, on a bien $P_A(A) = 0$.

Pour n arbitrairement fixé, on suppose que pour toute matrice $A' \in M_{n-1}(K)$ telle que $P_{A'}$ ait toutes ses racines dans K, on a $P_{A'}(A') = 0$.

Soit $A \in M_n(K)$, comme on se place dans le cas où P_A a toutes ses racines dans K, la matrice A est trigonalisable. Si u est l'endomorphisme de E de matrice A, il existe une base $(e_1, \dots, e_n)$ de E dans laquelle la matrice de u est égale à la matrice T où

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & & \vdots \\ \vdots & \ddots & \ddots & t_{n-1,n} \\ 0 & \cdots & 0 & t_{nn} \end{pmatrix}.$$

On a alors

$$P_A(X) = P_u(X) = P_T(X) = (t_{11} - X)\cdots(t_{nn} - X).$$

Soit F le sous-espace vectoriel de E engendré par $(e_1, \dots, e_{n-1})$, comme pour $1 \leq i \leq n-1$, $u(e_i) \in F$, le sous-espace F est stable par u. On considère la restriction de u à F, notons la $v = u_{|F}$,

$$\begin{aligned} v : F &\longrightarrow F \\ x &\longmapsto v(x) = u(x). \end{aligned}$$

La matrice de v dans la base $(e_1, \dots, e_{n-1})$ est la matrice triangulaire

$$T' = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1,n-1} \\ 0 & t_{22} & & \vdots \\ \vdots & \ddots & \ddots & t_{n-2,n-1} \\ 0 & \cdots & 0 & t_{n-1,n-1} \end{pmatrix}.$$

Par hypothèse de récurrence, on a $P_{T'}(T') = 0$, ce qui implique $P_v(v) = 0$. Ainsi, pour tout vecteur $x \in F$, on a $P_v(v)(x) = 0 = P_v(u)(x)$. Or, $P_u(X) = (t_{n,n} - X)P_v(X)$, d'où $P_u(u)(x) = 0$ pour tout $x \in F$. Il faut maintenant calculer $P_u(u)(e_n)$ pour conclure.

Compte tenu de la forme de la matrice T, il existe un vecteur $y \in F$ tel que

$$u(e_n) = t_{n,n}e_n + y.$$

c'est le vecteur $y = t_{1,n}e_1 + \cdots + t_{n-1,n}e_{n-1}$.

On en déduit que

$$(t_{n,n}\,\mathrm{id}_E - u)(e_n) = t_{n,n}e_n - t_{n,n}e_n - y = -y \in F,$$

Or, nous avons vu que

$$P_u(X) = (t_{n,n} - X)P_v(X) = P_v(X)(t_{n,n} - X),$$

on peut donc écrire

$$P_u(u)(e_n) = P_v(u) \circ (t_{n,n}\,\mathrm{id}_E - u)(e_n) = P_v(u)(-y) = 0$$

car $y \in F$. Ainsi l'endomorphisme $P_u(u)$ s'annule-t-il sur tous les vecteurs de la base $(e_1, \dots, e_n)$ de E, ce qui prouve que $P_u(u) = 0$. On a donc également $P_A(A) = 0$. En effet, si $P_A(X) = \sum_{k=0}^{r} a_k X^k$, alors $P_A(X) = \sum_{k=0}^{r} a_k A^k$ est la matrice de l'endomorphisme $\sum_{k=0}^{r} a_k u^k$, or, cet endomorphisme est identiquement nul, d'où $P_A(A) = 0$.

Il faut maintenant généraliser ce résultat au cas où le polynôme $P_A = P_u$ n'a plus nécessairement toutes ses racines dans K. Rappelons que le corps K est égal à $\mathbb{R}$ ou $\mathbb{C}$. Si $K = \mathbb{C}$, l'affaire est réglée puisque $\mathbb{C}$ est algébriquement clos. Si $K = \mathbb{R}$, c'est un sous-corps de $\mathbb{C}$, toute matrice $A \in M_n(\mathbb{R})$ peut être considérée comme une matrice de

$M_n(\mathbb{C})$. Le polynôme $P_A(X) = \det(A - XI_n)$ a toutes ses racines dans $\mathbb{C}$, on en déduit donc que $P_A(A) = 0$, d'après ce qui précède. Mais cette égalité $P_A(A) = 0$ reste vraie que l'on décompose le polynôme P_A dans $\mathbb{C}[X]$ ou dans $\mathbb{R}[X]$. D'où le résultat.

Exemple : Soit $A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$, le polynôme caractéristique de A est

$$P_A(X) = \begin{vmatrix} 1-X & -2 \\ 1 & -1-X \end{vmatrix} = (1-X)(-1-X) + 2 = X^2 + 1.$$

Ce polynôme n'a pas de racines dans $\mathbb{R}$, la matrice A n'est donc pas trigonalisable sur $\mathbb{R}$, par contre elle l'est sur $\mathbb{C}$. En effet, on a $P_A(X) = (X - i)(X + i)$ où i désigne la racine de -1. La matrice A admet donc deux valeurs propres complexes qui sont i et $-i$. Déterminons les sous-espaces propres associés. On cherche donc les vecteurs $v = (x, y) \in \mathbb{C}^2$ tels que $Av = iv$ ou $Av = -iv$.

$$\begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = i\begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} x - 2y = ix \\ x - y = iy \end{cases} \iff x = (1+i)y.$$

Le sous-espace E_i est donc la droite vectorielle de $\mathbb{C}^2$ engendrée par le vecteur $e_1 = (1+i, 1)$.

$$\begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -i\begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} x - 2y = -ix \\ x - y = -iy \end{cases} \iff x = (1-i)y.$$

Le sous-espace E_{-i} est donc la droite vectorielle de $\mathbb{C}^2$ engendrée par le vecteur $e_2 = (1 - i, 1)$.
La matrice A est équivalente dans $M_2(\mathbb{C})$ à la matrice diagonale

$$D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

On a bien $P_A(A) = P_D(A) = A^2 + I_2 = (A - iI_2)(A + iI_2) = 0$. L'égalité $A^2 + I_2 = 0$ est évidemment vraie dans $M_2(\mathbb{R})$.

II.7. Polynôme minimal.

Nous venons de démontrer que si u est un endomorphisme et P_u son polynôme caractéristique, alors $P_u(u) = 0$ (et de même $P_A(A) = 0$ pour $A \in M_n(K)$). Nous allons démontrer qu'il existe un plus petit polynôme ayant cette propriété.

Proposition 1. *Soit u un endomorphisme de E, il existe un unique polynôme $Q \in K[X]$ tel que*
1) Q unitaire et $Q(u) = 0$
2) Si $P \in K[X]$ est tel que $P(u) = 0$, alors Q divise P.

Ce polynôme que nous noterons Q_u est appelé polynôme minimal de u.

Démonstration.

Notons $n = \dim E$. Soit u un endomorphisme de E et P_u son polynôme caractéristique. Le polynôme $(-1)^n P_u$ est unitaire et s'annule en u (c'est le théorème de Cayley-Hamilton). Ainsi, l'ensemble des polynômes unitaires vérifiant $P(u) = 0$ n'est pas vide. Choisissons dans cet ensemble un polynôme Q de degré minimal. Il est clair que tout polynôme multiple de Q s'annule également en u. En effet, si $P = TQ$, alors $P(u) = T(u) \circ Q(u) = 0$.

Réciproquement, soit $P \in K[X]$ tel que $P(u) = 0$. On effectue la division euclidienne de P par Q et on obtient $P = TQ + R$ avec $\deg R < \deg Q$, on a

$$P(u) = T(u) \circ Q(u) + R(u) = 0,$$

d'où l'on déduit $R(u) = 0$. Mais, compte tenu du choix de Q, supposé de degré minimal, on obtient $R = 0$, d'où $P = TQ$ c'est-à-dire Q divise P. Vérifions l'unicité d'un tel Q, supposons qu'il existe Q_1 et Q_2 vérifiant les propriétés de la proposition. Les polynômes Q_1 et Q_2 sont alors unitaires et s'annulent en u, on en déduit que Q_1 divise Q_2 et Q_2 divise Q_1, ce qui prouve leur égalité.

Nous allons maintenant démontrer les deux propositions suivantes :

Proposition 2. *Soit $\lambda \in K$, alors λ est valeur propre de u si et seulement si λ est racine de son polynôme minimal.*

Démonstration.

On sait que Q_u divise P_u, ainsi, si λ est racine du polynôme minimal Q_u on a

$$Q_u(\lambda) = 0 = P_u(\lambda),$$

donc λ est une racine du polynôme caractéristique, c'est-à-dire une valeur propre de u.

Réciproquement, supposons que λ soit une valeur propre de u, alors, nous avons vu dans une précédente proposition que pour tout polynôme P, $P(\lambda)$ est valeur propre de l'endomorphisme $P(u)$, en particulier, si Q est le polynôme minimal, $Q(\lambda)$ est valeur propre de $Q(u)$. Or, $Q(u)$ est l'endomorphisme identiquement nul, il n'admet donc pas d'autre valeur propre que 0, on a donc nécessairement $Q(\lambda) = 0$.

Nous arrivons ainsi à une nouvelle condition nécessaire et suffisante de diagonalisation.

Proposition 3. *L'endomorphisme u est diagonalisable si et seulement si son polynôme minimal a toutes ses racines dans K et celles-ci sont simples.*

Démonstration.

Supposons l'endomorphisme u diagonalisable. Notons $\lambda_1, \ldots, \lambda_r$ ses valeurs propres distinctes et Q le polynôme ainsi défini

$$Q(X) = (X - \lambda_1) \cdots (X - \lambda_r).$$

Pour $1 \leq i \leq r$, on note

$$P_i(X) = \frac{Q(X)}{X - \lambda_i} = (X - \lambda_1) \cdots (X - \lambda_{i-1})(X - \lambda_{i+1}) \cdots (X - \lambda_r).$$

On a alors $Q(X) = P_i(X)(X - \lambda_i)$ d'où $Q(u) = P_i(u) \circ (u - \lambda_i \,\mathrm{id}_E)$. Ainsi pour tout $x \in E_{\lambda_i}$, on a $Q(u)(x) = 0$ car $x \in \ker(u - \lambda_i \,\mathrm{id}_E)$. Mais, ceci est vrai pour tout i, c'est-à-dire quelque soit la valeur propre λ_i, et, u est supposée diagonalisable, donc

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r},$$

ce qui prouve que pour tout $x \in E$, $Q(u)(x) = 0$ donc $Q(u) = 0$, mézalors, le polynôme minimal Q_u de u divise Q (par définition), or toute valeur propre de u est racine de Q_u donc $Q_u = Q$ et toutes les racines du polynôme minimal sont dans K et sont simples.

Réciproquement, si toutes les racines de Q_u sont dans K et sont simples, alors, les valeurs propres étant racines, on a

$$Q_u(X) = (X - \lambda_1) \cdots (X - \lambda_r).$$

Or,

$$E = \ker Q_u(u) = \ker(u - \lambda_1 \,\mathrm{id}_E) \oplus \cdots \oplus \ker(u - \lambda_r \,\mathrm{id}_E).$$

Ceci prouve que u est diagonalisable.

(Cette dernière égalité sera démontrée dans le paragraphe suivant sous le nom de lemme des noyaux.)

Application à la diagonalisation : Soit

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

on a

$$P_A(X) = \begin{vmatrix} 2 - X & 1 & -1 \\ -1 & -1 - X & 3 \\ 0 & -1 & 3 - X \end{vmatrix} = (2 - X)(1 - X)^2.$$

Le polynôme minimal Q_A de A divise P_A il est donc égal à $(X-1)(X-2)$ où $(X-1)^2(X-2)$, or,

$$(A - I_3)(A - 2I_3) = \begin{pmatrix} -1 & -1 & 1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{pmatrix} \neq 0,$$

ainsi, le polynôme minimal est égal à $(X - 1)^2(X - 2)$, il admet une racine double, par conséquent A n'est diagonalisable ni dans $\mathbb{R}$ ni dans $\mathbb{C}$.

II.8. Sous-espaces caractéristiques et décomposition de Dunford

Nous avons vu que dans le cas où u est diagonalisable, on a $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$ avec $E_{\lambda_i} = \ker(u - \lambda_i \,\mathrm{id}_E)$ le sous-espace propre associé à la valeur propre λ_i. Nous allons démontrer que lorsque u n'est pas diagonalisable, si son polynôme a toutes ses racines dans le corps K, on peut écrire

$$E = \ker(u - \lambda_1 \,\mathrm{id}_E)^{m_1} \oplus \cdots \oplus \ker(u - \lambda_r \,\mathrm{id}_E)^{m_r},$$

où m_i est la multiplicité de la valeur propre λ_i comme racine du polynôme caractéristique de u.

Commençons par démontrer le lemme suivant :

Lemme des noyaux. *Soient P et Q des polynômes de $K[X]$, premiers entre eux , alors*

$$\ker P(u) \oplus \ker Q(u) = \ker(PQ)(u)$$

Démonstration.

Si P et Q sont premiers entre eux, alors, d'après le théorème de Bézout, il existe des polynômes U et V tels que $UP + VQ = 1$. On a donc

$$U(u) \circ P(u) + V(u) \circ Q(u) = id_E.$$

Ainsi, si $x \in \ker P(u) \cap \ker Q(u)$, on a

$$U(u) \circ \underbrace{P(u)(x)}_{=0} + V(u) \circ \underbrace{Q(u)(x)}_{=0} = x,$$

d'où $\ker P(u) \cap \ker Q(u) = \{0\}$, ce qui prouve que la somme est directe. Considérons maintenant $x \in \ker(PQ)(u)$, on a, toujours en raison du théorème de Bézout,

$$x = \underbrace{U(u) \circ P(u)(x)}_{\in \ker Q(u)} + \underbrace{V(u) \circ Q(u)(x)}_{\in \ker P(u)},$$

car, c'est une propriété des polynômes, $Q(u) \circ U(u) \circ P(u) = U(u) \circ P(u) \circ Q(u)$ et $P(u) \circ V(u) \circ Q(u) = V(u) \circ Q(u) \circ P(u) = V(u) \circ P(u) \circ Q(u)$, et, on a $P(u) \circ Q(u)(x) = 0$ car $x \in \ker(PQ)(u)$. D'où le résultat.

Ce résultat se généralise par récurrence sur n au cas d'un produit de n polynômes premiers entre eux. Si $P = Q_1 ... Q_n$ alors

$$\ker P(u) = \ker Q_1(u) \oplus \ker Q_2(u) \oplus \cdots \oplus \ker Q_n(u).$$

Définition. *Soient u un endomorphisme de E, λ une valeur propre de u et m sa mulptiplicité dans P_u, le sous-espace $N = \ker(u - \lambda \,\mathrm{Id}_E)^m$ est appelé sous-espace caractéristique de u pour la valeur propre λ.*

Proposition 1. *Soit u un endomorphisme de E tel que P_u ait toutes ses racines dans K (u est trigonalisable), notons $P_u(X) = (\lambda_1 - X)^{m_1} \cdots (\lambda_r - X)^{m_r}$ et, pour $1 \leq i \leq r$, N_i le sous-espace caractéristique associé à la valeur propre λ_i, alors*
1) N_i est stable par u
2) $E_{\lambda_i} \subset N_i$
3) $E = N_1 \oplus \cdots \oplus N_r$
4) $\dim N_i = m_i$

Démonstration.

1) Si $x \in N_i$, on a $(u - \lambda_i \operatorname{id}_E)^{m_i}(x) = 0$, or

$$(u - \lambda_i \operatorname{id}_E)^{m_i} \circ u(x) = u \circ (u - \lambda_i \operatorname{id}_E)^{m_i}(x) = 0,$$

d'où $u(x) \in N_i$.

2) $E_{\lambda_i} = \ker(u - \lambda_i \operatorname{id}_E) \subset N_i = \ker(u - \lambda_i \operatorname{id}_E)^{m_i}$, évident.

3) C'est le lemme des noyaux.

$$P_u(X) = (\lambda_1 - X)^{m_i} \cdots (\lambda_r - X)^{m_r},$$

les polynômes $(\lambda_i - X)^{m_i}$ sont premiers entre eux puisque les valeurs propres sont distinctes. Par récurrence, on obtient

$$\ker P_u = N_1 \oplus \cdots \oplus N_r,$$

or, d'après le théorème de Cayley-Hamilton, on a $P_u(u) = 0$, donc $\ker P_u(u) = E$, d'où le résultat.

4) Notons $v_i = u_{|N_i}$ pour $1 \leq i \leq r$. Pour $i \neq j$, $N_i \cap N_j = \{0\}$, or, $E_{\lambda_j} \subset N_j$, donc la seule valeur propre de v_i est λ_i. On a donc $P_{v_i}(X) = (\lambda_i - X)^{\dim N_i}$, (n'oublions pas que P_u, et donc ses diviseurs, ont toutes leur racines dans K). Or,

$$P_u = P_{v_1} \cdots P_{v_r} = (\lambda_1 - X)^{m_1} \cdots (\lambda_r - X)^{m_r},$$

d'où $\dim N_i = m_i$ pour $1 \leq i \leq r$.

Définition. *On dit que l'endomorphisme u (resp. la matrice A) est* nilpotent(e) *si il existe $k \in \mathbb{N}^*$ tel que $u^k = 0$ (resp. $A^k = 0$)*

Proposition 2. *Si u est nilpotent, 0 est son unique valeur propre et on a*

$$P_u(X) = (-1)^n X^n.$$

Démonstration.

$A^k = 0 \Rightarrow \det A = 0$, donc l'endomorphisme u n'est pas bijectif, $\ker u \neq \{0\}$, ce qui prouve que 0 est une valeur propre de u.
Supposons que λ soit une autre valeur propre, il existe $x \neq 0$ tel que $u(x) = \lambda x$, on a donc, par récurrence sur n, pour tout n, $u^n(x) = \lambda^n x$, or, u est supposé nilpotent, il existe donc $k \in \mathbb{N}^*$ tel que $u^k = 0$, d'où, si x est vecteur propre pour la valeur propre λ, $\lambda^k x = 0$, ce qui implique $\lambda^k = 0$, donc $\lambda = 0$. Ainsi, 0 est la seule valeur propre de u, on a donc $P_u(X) = (-1)^n X^n$.

Nous allons démontrer que les endomorphismes nilpotents et les endomorphismes diagonalisables permettent de décrire tous les endomorphismes trigonalisables (c'est-à-dire ceux dont le polynôme caractéristique a toutes ses racines dans K).

Théorème : Décomposition de Dunford. *Soit u un endomorphisme de E tel que P_u ait toutes ses racines dans K. Alors, il existe un unique couple (n, d) d'endomorphismes, avec n nilpotent et d diagonalisable tels que*

$$(i)\ u = n + d\,, \quad (ii)\ n \circ d = d \circ n.$$

Ce que l'on peut encore écrire :

Théorème. *Pour toute matrice $A \in M_n(K)$, trigonalisable, il existe une unique matrice N nilpotente et une unique matrice D diagonalisable telles que*

$$(i)\ A = N + D\,, \quad (ii)\ ND = DN.$$

Avant de démontrer ces théorèmes, nous allons démontrer deux lemmes dont les résultats nous seront utiles.

Lemme 1. *Si u est diagonalisable et V est stable par u alors la restriction de u à V est diagonalisable.*

Démonstration.

Notons $v = u_{|V}$ et $\lambda_1, \ldots, \lambda_r$ les valeurs propres de u. L'endomorphisme u étant supposé diagonalisable, on a

$$E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r},$$

où les E_{λ_i} sont les sous-espaces propres de u. On a, puisque V est stable par u,

$$V = (V \cap E_{\lambda_1}) \oplus \ldots \oplus (V \cap E_{\lambda_r}).$$

Or, $\ker(v - \lambda_i \operatorname{id}_V) = V \cap \ker(u - \lambda_i \operatorname{id}_E)$ donc les valeurs propres de v notées $\mu_1, \ldots, \mu_s$ sont dans l'ensemble $\{\lambda_1, \ldots, \lambda_r\}$. On a donc $V = V_{\mu_1} \oplus \cdots \oplus V_{\mu_s}$, ce qui prouve que v est diagonalisable.

Lemme 2. *Soient u et v deux endomorphismes diagonalisables, on suppose que $u \circ v = v \circ u$, alors il existe une base commune de vecteurs propres.*

Démonstration.

Soient $\lambda_1, \ldots, \lambda_r$ les valeurs propres de v. Notons $V_i = \{x \in E / v(x) = \lambda_i x\}$. On a alors pour $x \in V_i$,

$$v(u(x)) = u(v(x)) = u(\lambda_i x) = \lambda_i u(x),$$

donc $u(x) \in V_i$, V_i est stable par u. D'après le lemme 1, la restriction de u à V_i est donc diagonalisable. On considère dans V_i une base $\mathcal{B}_i$ de vecteurs propres de u, comme v est diagonalisable, on a

$$E = \underbrace{V_1}_{\mathcal{B}_1} \oplus \cdots \oplus \underbrace{V_r}_{\mathcal{B}_r},$$

la base $\mathcal{B} = \mathcal{B}_1 \cup \ldots \mathcal{B}_r$ est donc une base formée de vecteurs qui sont des vecteurs propres de u et de v.

Démonstration du théorème de décomposition de Dunford.

Soit P_u le polynôme caractéristique de u qui a toutes ses racines dans K,

$$P_u(X) = \prod_{i=1}^{r} (\lambda_i - X)^{m_i}.$$

Soient $N_1 \dots, N_r$ les sous-espaces caractéristiques, pour $1 \leq i \leq r$, on a

$$N_i = \ker(u - \lambda_i \operatorname{id}_E)^{m_i} \text{ et } E = N_1 \oplus \cdots \oplus N_r.$$

Nous allons définir les endomorphismes n et d sur chaque N_i de la manière suivante, pour tout $x \in N_i$, on pose

$$d(x) = \lambda_i x \text{ et } n(x) = u(x) - d(x).$$

Pour $1 \leq i \leq r$, on a $d_i = d_{|N_i} = \lambda_i \operatorname{id}_{N_i}$, l'espace vectoriel E étant somme directe des N_i, les endomorphismes ainsi construits conviennent, c'est ce que nous allons vérifier :

1) Par construction, d est diagonalisable. En effet, E étant somme directe des N_i, dans une base de E formée de bases des N_i, $1 \leq i \leq r$, la matrice de d est diagonale.

2) On pose $n_i = n_{|N_i} = u_{|N_i} - \lambda_i \operatorname{id}_{N_i}$, on a donc $n_i^{m_i} = 0$, ainsi, si $m = \sup m_i$, puisque n^m s'annule sur chaque N_i alors $n^m = 0$, ce qui prouve que n est nilpotent.

3) On vérifie que $d \circ n = n \circ d$, soit $x \in E$, pour $1 \leq i \leq r$, il existe $x_i \in N_i$ tels que $x = x_1 + \cdots + x_r$, on a donc

$$d \circ n(x) = d \circ n(x_1) + \cdots + d \circ n(x_r) = n \circ d(x_1) + \cdots + n \circ d(x_r) = n \circ d(x),$$

en effet, sur chaque N_i, $d_{|N_i} = \lambda_i \operatorname{id}_{N_i}$ donc commute avec tout endomorphisme.

4) Il reste à vérifier l'unicité : supposons que l'on ait deux couples (n, d) et (n', d') vérifiant (i) et (ii), on a alors $u = d + n = d' + n'$, d'où $d' \circ u = d' \circ d' + d' \circ n' = u \circ d'$. Ainsi, comme pour tout i, $1 \leq i \leq r$, N_i est stable par d', pour $x \in N_i$ on a

$$(u - \lambda_i \operatorname{id})^{m_i} \circ d'(x) = d' \circ (u - \lambda_i \operatorname{id})^{m_i}(x) = 0.$$

Par ailleurs, $d_{|N_i} = \lambda_i \operatorname{id}_{N_i}$, donc d et d' commutent sur E. D'après le lemmme 2, il existe une base commune de vecteurs propres, ainsi $d - d'$ est diagonalisable. Or, $n = u - d$ et $n' = u - d'$ donc, si d et d' commutent, n et n' commutent également. Les endomorphismes n et n' étant nilpotents, $n - n'$ l'est également, en effet, si p et q sont des entiers tels que $n^p = n'^q = 0$, alors $(n - n')^{p+q} = 0$. Ainsi $d - d' = n - n'$ est un endomorphisme qui est à la fois diagonalisable et nilpotent donc nul puisqu'alors sa seule valeur propre est 0. On a donc $d - d' = n - n' = 0$ ce qui prouve l'unicité.

La décomposition de Dunford est utile dans la mesure où elle permet de calculer les puissances d'une matrice, nous verrons plus en détails comment et pourquoi dans le paragraphe suivant. Pour cela il est nécessaire que les matrices D et N commutent car on peut alors utiliser la formule du binôme de Newton :

$$(N + D)^k = \sum_{i=1}^{k} C_k^i N^i D^{k-i}.$$

A partir d'un certain rang i, les matrices N^i sont nulles, quant aux matrices diagonalisables, on peut calculer leurs puissances d'une manière relativement simple. En effet, si P est la matrice de passage qui exprime la base des vecteurs propres de D dans la base canonique, on a pour tout entier k, $D^k = P\Delta^k P^{-1}$ et, si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres, on a

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

Attention, une matrice triangulaire peut toujours s'écrire comme somme d'une matrice diagonale et d'une matrice nilpotente, mais celles-ci ne commutent pas en général. Etudions un exemple.

Considérons la matrice

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

On peut écrire

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{D} + \underbrace{\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{N}$$

mais

$$DN = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq ND = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Par contre, puisque A est triangulaire, on peut la décomposer de manière unique suivant la décomposition de Dunford. Le polynôme caractéristique P_A est égal à

$$P_A(X) = \begin{pmatrix} 1-X & 1 & 1 \\ 0 & 1-X & 1 \\ 0 & 0 & 2-X \end{pmatrix} = (1-X)^2(2-X).$$

Nous avons donc deux valeurs propres qui sont 1 et 2. L'espace vectoriel E s'écrit comme somme directe

$$E = \ker(A - \mathrm{Id})^2 \oplus \ker(A - 2\,\mathrm{Id}).$$

Déterminons les sous-espaces caractéristiques. Le noyau $\ker(A-2\,\mathrm{Id}) = \{u \in E,\ Au = 2u\}$, si $u = (x, y, z)$, on résout

$$\begin{cases} x + y + z = 2x \\ y + z = 2y \\ 2z = 2z \end{cases} \iff \begin{cases} -x + y + z = 0 \\ -y + z = 0 \end{cases} \iff \begin{cases} y = z \\ x = 2z \end{cases}$$

Le sous-espace $\ker(A - 2\,\mathrm{Id})$ est donc la droite vectorielle engendrée par le vecteur $(2, 1, 1)$. La matrice A n'est pas diagonalisable, en effet $\ker(A - \mathrm{Id}) = \{u = (x, y, z),\ Au = u\}$, on résout

$$\begin{cases} x + y + z = x \\ \quad\ \ y + z = y \\ \qquad\ 2z = z \end{cases} \iff y = z = 0\ ,$$

on obtient la droite vectorielle engendrée par le vecteur $(1, 0, 0)$ sa dimension n'est pas égale à la multiplicité de la racine. Déterminons $\ker(A - \mathrm{Id})^2$,

$$A - \mathrm{Id} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (A - \mathrm{Id})^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ainsi $\ker(A - \mathrm{Id})^2$ est le plan vectoriel engendré par les vecteurs $i = (1, 0, 0)$ et $j = (0, 1, 0)$, notons $e_1 = i$, $e_2 = j$ et $e_3 = 2i + j + k$ où $k = (0, 0, 1)$. Soient

$$N_1 = \ker(A - \mathrm{Id})^2 = \mathbb{R}i + \mathbb{R}j = \mathbb{R}e_1 + \mathbb{R}e_2 \text{ et } N_2 = \ker(A - 2\,\mathrm{Id}) = \mathbb{R}e_3.$$

On pose $d(e_1) = e_1$, $d(e_2) = e_2$ et $d(e_3) = 2e_3$, d'où $d(i) = i$, $d(j) = j$ et

$$d(k) = -2e_1 - e_2 + 2e_3 = -2i - j + 4i + 2j + 2k = 2i + j + 2k.$$

On a donc dans la base (i, j, k) :

$$D = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \text{ et } N = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On peut regarder les choses sous divers points de vue, dans la base (e_1, e_2, e_3), la matrice de l'endomorphisme u est la matrice T dont les colonnes sont $u(e_1) = u(i) = i = e_1$, $u(e_2) = u(j) = i + j = e_1 + e_2$ et $u(e_3) = 2u(i) + u(j) + u(k) = 2e_3$ car $e_3 \in \ker(u - 2\,\mathrm{Id})$. On a $T = N' + D'$ où

$$N' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } D' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

la matrice D' est diagonale et la matrice N' est nilpotente. La matrice de passage P est égale à

$$P = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

on a $PTP^{-1} = A = PN'P^{-1} + PD'P^{-1} = N + D$. La matrice N est nilpotente et la matrice D est diagonalisable.

On a $ND = PN'P^{-1}PD'P^{-1} = PN'D'P^{-1} = PD'N'P^{-1} = DN$, c'est la décomposition de Dunford.

Application à la trigonalisation

Nous avons démontré dans un chapitre précédent qu'un polynôme qui a toutes ses racines dans $\mathbb{R}$ est trigonalisable dans $\mathbb{R}$. Par contre nous avions laissé assez libre le choix de la base de trigonalisation, on ne demandait au troisième vecteur de cette base que d'être linéairement indépendant des deux premiers (qui sont des vecteurs propres). La décomposition de Dunford, que nous venons de voir, permet de choisir la base de trigonalisation de manière à obtenir naturellement la décomposition de Dunford de la matrice triangulaire obtenue.

Soit u un endomorphisme, on suppose que son polynôme caractéristique a toutes ses racines dans $\mathbb{R}$. On écrit
$$P_u(X) = (\lambda_1 - X)^{m_1} \dots (\lambda_r - X)^{m_r}.$$
On a $E = N_1 \oplus \dots \oplus N_r$, où $N_i = \ker(\lambda_i - u)^{m_i}$.
On sait que chaque N_i est stable par u et si on note $\mathcal{B}_i$ une base de N_i, alors, dans la base de E, $\mathcal{B} = \cup \mathcal{B}_i$ la matrice de u est une matrice par blocs
$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}.$$
Où A_i est la matrice de $u_i = u_{|_{N_i}}$. On trigonalise alors chaque A_i. On a $P_{u_i}(X) = (\lambda_i - X)^{m_i}$. On choisit une base de $\ker(u - \lambda_i \operatorname{id}_E)$ et on a
$$\ker(u - \lambda_i \operatorname{id}_E) \subset \ker(u - \lambda_i \operatorname{id}_E)^2 \subset \dots \subset \ker(u - \lambda_i \operatorname{id}_E)^{m_i}.$$

Exemple: Soit A la matrice de l'endomorphisme u suivante
$$A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 3 & -4 \\ 3 & 1 & -2 \end{pmatrix}$$

$P_A(X) = -(X+1)(X-2)^2$. Déterminons les sous-espaces caractéristiques.

- $N_{-1} = F = \ker(u + \operatorname{id}_E)$, c'est le sous-espace propre associé à la valeur propre -1.

- $N_2 = G = \ker(u - 2\operatorname{id}_E)^2 \supset \ker(u - 2\operatorname{id}_E) = E_2$.

L'espace F est la droite vectorielle engendrée par le vecteur $e_1 = (0,1,1)$ et l'espace E_2 est la droite vectiorielle engendrée par le vecteur $e_2 = (1,1,1)$. Si e_3 est un vecteur de G indépendant de e_2 alors, dans le base (e_1, e_2, e_3) la matrice de u sera de la forme
$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & a \\ 0 & 0 & 2 \end{pmatrix}.$$

On choisit un vecteur e_3 dans $N_2 \setminus E_2$. On a

$$(A-2I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -9 & 0 & 9 \\ -9 & 0 & 9 \end{pmatrix},$$

d'où $\ker(A-2I)^2 = \{(x,y,z)/\ x=z$, c'est un plan vectoriel. Le vecteur $e_3 = (1,0,1)$ est dans N_2 mais pas dans E_2 est les vecteurs e_1, e_2, e_3 forment une base de E. On calcule $u(e_3)$ et on identifie

$$u(e_3) = (1,-1,1) = a(1,1,1) + 2(1,0,1)$$

d'où $a = -1$.

La matrice de u dans la base (e_1, e_2, e_3) est triangulaire, elle s'écrit

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

et, si P est la matrice de passage

$$P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

on a $T = P^{-1}AP$.

Cette trigonalisation permet d'obtenir la décomposition de Dunford de A en écrivant

$$T = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{\Delta} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}}_{M}.$$

On a $A = PTP^{-1} = \underbrace{P\Delta P^{-1}}_{D} + \underbrace{PMP^{-1}}_{N}$. La matrice D est diagonalisable, la matrice N est nilpotente et ce sont les matrices de la décomposistion de Dunford.

III. Résolution de systèmes différentiels linéaires

En première année, on apprend à résoudre les équations différentielles du premier ordre

$$x'(t) = a(t)x(t) + b(t)$$

où x, a et b sont des fonctions de la variable réelle t à valeurs dans $\mathbb{R}$, la fonction x étant la fonction inconnue, et les équations différentielles du second ordre à coefficients constants

$$x''(t) + px'(t) + qx(t) = 0,$$

où p et q sont des constantes réelles. Nous allons maintenant étudier le cas où la fonction inconnue ne prend plus ses valeurs dans $\mathbb{R}$ mais dans $\mathbb{R}^n$. C'est-à-dire que nous étudierons les systèmes d'équations

$$\begin{cases} x_1'(t) = a_{1,1}x_1(t) + a_{1,2}x_2(t) + \cdots + a_{1,n}x_n(t) + b_1(t) \\ \qquad \vdots \\ x_n'(t) = a_{n,1}x_1(t) + a_{n,2}x_2(t) + \cdots + a_{n,n}x_n(t) + b_n(t) \end{cases}$$

que l'on peut encore écrire sous forme matricielle $X' = AX + B$, où $A \in M_n(\mathbb{R})$ et B est un vecteur colonne de $\mathbb{R}^n$,

$$A = (a_{i,j})_{1\leq i,j\leq n} \quad \text{et} \quad B = \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix},$$

les $(b_i(t))_{1\leq i\leq n}$ sont des fonctions de la variable réelle t et les $(a_ij)_{1\leq i,j\leq n}$ des réels.

On cherche les solutions de

$$(\star) \qquad X' = AX + B$$

où X est une application inconnue de $\mathbb{R}$ dans $\mathbb{R}^n$, $A \in M_n(\mathbb{R})$ et B une application de $\mathbb{R}$ dans $\mathbb{R}^n$

Rappelons que, dans le cas où $n = 1$, on obtient les solutions de l'équation $x' = ax+b(t)$ $(*)$ en intégrant d'abord l'équation homogène $x' = ax$ dont les solutions sont les fonctions $x(t) = ke^{at}$ où k est une constante réelle. On obtient toutes les solutions de $(*)$ en ajoutant à la solution générale de l'équation homogène une solution particulière de $(*)$ que l'on peut trouver par la méthode de variation de la constante.

Dans le cas des systèmes d'équations, on voit immédiatement apparaître l'intérêt de la trigonalisation (et évidemment de la diagonalisation) qui permettra de résoudre le système de proche en proche. En effet, si $A = T$ est triangulaire on a

$$\begin{cases} x_1' = t_{1,1}x_1 + \cdots + \cdots + t_{1,n}x_n + b_1 \\ x_2' = \qquad t_{2,2}x_2 + \cdots + t_{2,n}x_n + b_2 \\ \quad \vdots \\ x_n' = \qquad\qquad t_{n,n}x_n + b_n \end{cases}$$

On peut alors intégrer la dernière équation puis reporter la solution dans la précédente et ainsi de proche en proche intégrer tout le sytème.

Si $A = D$ est diagonale, c'est encore plus simple on est ramené à la résolution de n équations du type $x' = \lambda x$.

Afin de pouvoir généraliser les méthodes de résolution en dimension 1 aux systèmes, nous allons introduire les exponentielles de matrices.

III.1. Exponentielle de matrices.

L'espace vectoriel $M_n(\mathbb{C})$ est un espace vectoriel de dimension finie, n^2, sur lequel toutes les normes sont équivalentes, on en choisit une que l'on note $\| \ \|$. La série de terme général $\frac{1}{k!}a^k$ étant convergente pour tout $a \in \mathbb{R}$, la série de terme général $\frac{1}{k!}\|A^k\|$ est également convergente pour toute matrice $A \in M_n(\mathbb{C})$. Par conséquent, la série $\displaystyle\sum_{k=0}^{+\infty} \frac{1}{k!}A^k$ est convergente dans $M_n(\mathbb{C})$.

Rappels : Rappelons la définition d'une série, soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels, on appelle série de terme général u_n la suite $S_n = \sum_{k=0}^{n} u_k$, si cette suite admet une limite quand n tend vers l'infini, on dit que la série converge et on note $S = \sum_{k=0}^{+\infty} u_k$ sa limite.

Si $x \in \mathbb{C}$, on peut définir l'exponentielle de x par

$$e^x = \lim_{p\to+\infty}(1 + x + \frac{x^2}{2!} + \ldots + \frac{x^p}{p!} = \sum_{k=0}^{+\infty} \frac{1}{k!}x^k.$$

De même, nous allons définir l'exponentielle d'une matrice.

Définition. *Soit $A \in M_n(\mathbb{C})$, la matrice définie par*

$$\sum_{k=0}^{+\infty} \frac{1}{k!}A^k = \lim_{p\to+\infty}(I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{p!}A^p)$$

est l'exponentielle de la matrice A, on la note $\exp A$.

Remarquons que si A est diagonale

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \quad \text{et} \quad \exp A = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix}.$$

Lorsque A est nilpotente, la somme définissant $\exp A$ est une somme finie.
L'exponentielle de matrices vérifie les propriétés suivantes :

Propriétés.
1) Si A est la matrice nulle, $\exp A = I_n$
2) Si A et $B \in M_n(\mathbb{C})$ vérifient $AB = BA$, alors $\exp(A + B) = (\exp A)(\exp B)$.
3) Pour toute matrice $A \in M_n(\mathbb{C})$, la matrice $\exp A$ *est inversible et* $(\exp A)^{-1} = \exp(-A)$
4) Si A et $P \in M_n(\mathbb{C})$, et P inversible, on a $\exp(P^{-1}AP) = P^{-1}(\exp A)P$

Nous ne démontrerons pas ces propriétés, nous pouvons cependant faire les remarques suivantes : le 1) est évident, le 2) se démontre comme dans le cas de l'exponentielle complexe, le fait que les matrices commutent permet d'utiliser la formule du binôme de Newton. Pour le 3), on remarque que les matrices A et $-A$ commutent d'où

$$\exp(A - A) = \exp 0 = I = (\exp A)(\exp(-A)).$$

Pour le 4), on note que pour tout $k \in \mathbb{N}^*$, on a $P^{-1}A^kP = (P^{-1}AP)^k$ et l'on revient à la définition de l'exponentielle

$$\exp(P^{-1}AP) = \sum_{k=0}^{+\infty} \frac{1}{k!} P^{-1}A^kP = P^{-1} \left(\sum_{k=0}^{+\infty} \frac{1}{k!} A^k \right) P.$$

Le calcul de l'exponentielle se fait alors de la manière suivante, si A est diagonale ou nilpotente, il n'y a pas de problème, sinon on utilise la décomposition de Dunford $A = N + D$ avec D diagonalisable et N nilpotente et $ND = DN$, ce qui permet d'écrire $\exp A = \exp N \exp D$. La matrice D étant diagonalisable, il existe une matrice P inversible telle que $P^{-1}DP$ soit diagonale, $D = P\Delta P^{-1}$ d'où

$$\exp D = \exp(P\Delta P^{-1}) = P(\exp \Delta)P^{-1}.$$

On peut donc toujours calculer l'exponentielle d'une matrice à coefficients dans $\mathbb{C}$.

III. 2. Systèmes différentiels linéaires

Définition. *Soient $A \in M_n(\mathbb{R})$, $I \subset \mathbb{R}$ un intervalle ouvert, $B : I \to \mathbb{R}^n$ une application. Une solution du sytème $X' = AX + B$, est une fonction S définie sur I et à valeurs dans $\mathbb{R}^n$, dérivable telle que pour tout $t \in I$, on ait*

$$S'(t) = AS(t) + B(t).$$

C'est-à-dire

$$\begin{cases} s_1'(t) = a_{1,1}s_1(t) + \cdots + a_{1,n}s_n(t) + b_1(t) \\ s_2'(t) = a_{2,1}s_1(t) + \cdots + a_{2,n}s_n(t) + b_2(t) \\ \qquad \vdots \\ s_n'(t) = a_{n,1}s_1(t) + \cdots + a_{n,n}s_n(t) + b_n(t) \end{cases}$$

Si $I = \mathbb{R}$ et $B = 0$, on dit que le système est homogène.

Nous allons voir comment généraliser les méthodes de résolution étudiées en dimension 1 et comment utiliser les propriétés de l'exponentielle de matrices et la réduction des matrices carrées pour écrire les solutions.

Proposition 1. *Soient $A \in M_n(\mathbb{R})$, $I \subset \mathbb{R}$ un intervalle ouvert, $B : I \to \mathbb{R}^n$ une application. Si S_0 est une solution de $X' = AX + B$, alors toute solution de $X' = AX + B$ s'écrit $S + S_0$, où S est solution du système homogène $X' = AX$.*

Démonstration.

Soit S telle que $S'(t) = AS(t)$ et S_0 telle que $S_0'(t) = AS_0(t) + B(t)$. On a bien

$$(S + S_0)'(t) = A(S(t) + S_0(t)) + B(t)$$

ce qui prouve que $S + S_0$ est solution de $X' = AX + B$.
Réciproquement, soit U une solution de $X' = AX + B$, alors

$$\begin{aligned}(U - S_0)'(t) &= U'(t) - S_0'(t)\\ &= AU(t) + B(t) - AS_0(t) - B(t)\\ &= A(U - S_0)(t).\end{aligned}$$

Ainsi, $U - S_0$ est solution du système homogène et on a $U = S_0 + (U - S_0)$, ce qui démontre la proposition.

III.2.1. Résolution des systèmes homogènes

Ainsi, comme dans le cas $n = 1$, nous allons nous attacher à donner une méthode de résolution des systèmes homogènes. Mais, avant de démontrer le théorème général, nous allons voir quelques propositions.

Proposition 2. *Soient $A \in M_n(\mathbb{R})$, λ une valeur propre de A et V un vecteur propre associé. Alors, la fonction*

$$\begin{aligned}\mathbb{R} &\longrightarrow \mathbb{R}^n\\ t &\longmapsto e^{\lambda t}V\end{aligned}$$

est solution de $X' = AX$.

Démonstration.

Soit $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ et $S(t) = e^{\lambda t}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$. On a alors

$$S'(t) = \lambda e^{\lambda t}V = e^{\lambda t}(\lambda V) = e^{\lambda t}AV = AS(t).$$

Ce qui prouve que S est bien solution du système homogène $X' = AX$.

Exemple : Soit $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$, on a $P_A(X) = (X - 2)^2$, la seule valeur propre de A est donc $\lambda = 2$, déterminons un vecteur propre : soit $V = (x, y) \in \mathbb{R}^2$ tel que $A.V = 2V$, on a alors $x + y = 0$, le vecteur $V = (1, -1)$ est un vecteur propre de A et l'application $S(t) = e^{2t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est une solution du système $S' = AS$.

Proposition 3. *Soit $Y : \mathbb{R} \to \mathbb{R}^n$ une fonction. Soient A et $P \in M_n(\mathbb{R})$, P inversible, alors, PY est solution de $X' = AX$ si et seulement si Y est solution de $X' = (P^{-1}AP)X$.*

Démonstration.

On suppose Y dérivable et PY solution de $X' = AX$. On a $(PY)' = PY'$, d'où

$$(PY)' = A(PY) \iff PY' = APY \iff Y' = (P^{-1}AP)Y.$$

Proposition 4. *Soit $A \in M_n(\mathbb{R})$, pour $t \in \mathbb{R}$, l'application, de $\mathbb{R}$ dans $M_n(\mathbb{R})$, définie par $R(t) = \exp tA$ est dérivable et on a $R'(t) = AR(t)$.*

Démonstration.

On a

$$R(t) = \exp tA = \sum_{k=0}^{+\infty} \frac{1}{k!} t^k A^k = \sum_{k=0}^{+\infty} R_k(t)$$

où $R_k(t) = \frac{1}{k!} t^k A^k$ et, pour tout k, on a

$$R'_k(t) = \frac{1}{(k-1)!} t^{k-1} A^k = AR_{k-1}(t).$$

Pour des raisons de convergence normale, comme dans le cas des séries de fonctions, on a

$$R'(t) = \sum_{k=0}^{+\infty} R'_k(t) = \sum_{k=0}^{+\infty} AR_k(t) = AR(t).$$

On arrive alors au théorème principal.

Théorème. *Soit $A \in M_n(\mathbb{R})$, les solutions du système différentiel homogène $X' = AX$ sont les fonctions*

$$\begin{aligned} S : \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \exp(tA).V \end{aligned}$$

où V est un vecteur de $\mathbb{R}^n$.

Démonstration.

Notons $E_1, \ldots, E_n$ les vecteurs de la base canonique de $\mathbb{R}^n$. Soit $i \in \{1, \ldots, n\}$, on pose $S_i(t) = \exp(tA)E_i$, d'après la proposition précédente, on a $S'_i(t) = AS_i(t)$. Si

$$V = v_1E_1 + \cdots + v_nE_n = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

on pose

$$S(t) = (\exp(tA))V = \sum_{k=1}^{n} v_k(\exp(tA))E_k = \sum_{k=1}^{n} v_k S_k(t),$$

d'où

$$S'(t) = \sum_{k=1}^{n} v_k S_k'(t) = \sum_{k=1}^{n} v_k A S_k(t) = A \sum_{k=1}^{n} v_k S_k(t) = AS(t).$$

Ainsi, la fonction S est bien solution. Il reste à vérifier que toutes les solutions s'écrivent de cette manière.
Soit S une solution de $X' = AX$ et soit

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \exp(-tA)S(t). \end{aligned}$$

On a alors

$$F'(t) = -A\exp(-tA)S(t) + \exp(-tA)S'(t) = -A\exp(-tA)S(t) + A\exp(-tA)S(t) = 0$$

car les matrices A et $\exp(tA)$ commutent. Ceci prouve que l'application F est constante, il existe donc un vecteur $V \in \mathbb{R}^n$ tel que pour tout $t \in \mathbb{R}$, on ait $F(t) = V$. Ainsi, la solution du système s'écrit

$$S(t) = (\exp tA)V.$$

Dans la pratique, que fait-on ?
Il s'agit d'intégrer l'équation $X' = AX$ dont les solutions s'écrivent

$$X(t) = \exp(tA).V,\ V \in \mathbb{R}^n.$$

Soit P une matrice inversible telle que la matrice $B = P^{-1}AP$ s'écrive $B = D + N$ avec D diagonale, N nilpotente et $ND = DN$. D'après la proposition 3, PY est solution de $X' = AX$ si et seulement si Y est solution de $X' = (P^{-1}AP)X = BX$. Les solutions $Y(t)$ sont donc de la forme

$$Y(t) = \exp(tB)V,\ V \in \mathbb{R}^n.$$

On obtient alors

$$PY = X = P\exp(tB)V,\ V \in \mathbb{R}^n,$$

or, $\exp(tB) = \exp(tD + tN) = \exp(tD)\exp(tN)$ et les matrices $\exp(tD)$ et $\exp(tN)$ sont faciles à calculer puisque D est diagonale et N ets nilpotente. Ainsi, les solutions de l'équation $X' = AX$ sont de la forme

$$X(t) = P\exp(tD)\exp(tN)V,\ V \in \mathbb{R}^n$$

et il est inutile de calculer la matrice P^{-1}.

Proposition 5. *Soit $A \in M_n(\mathbb{R})$, l'ensemble des solutions du système homogène $X' = AX$ est un espace vectoriel de dimension n.*

Démonstration.

Nous allons vérifier que c'est un sous-espace vectoriel de l'espace des fonctions de $\mathbb{R}$ dans $\mathbb{R}^n$.
- Il est clair que la fonction nulle est solution du système homogène.
- Si S_1 et S_2 sont deux solutions, a_1 et a_2 deux réels, on a

$$(a_1S_1 + a_2S_2)' = a_1S_1' + a_2S'2 = a_1AS_1 + a_2AS_2 = A(a_1S_1 + a_2S_2),$$

l'ensemble des solutions est donc stable par combinaison linéaire.
Notons $\mathcal{S}$ l'espace des solutions de $X' = AX$ et soit

$$\begin{aligned} u : \mathbb{R}^n &\longrightarrow \mathcal{S} \\ V &\longmapsto S : t \longmapsto (\exp tA)V. \end{aligned}$$

L'application u est linéaire et $\ker u = \{0\}$, en effet

$$\ker u = \{V \in \mathbb{R}^n, S \equiv 0\} = \{V \in \mathbb{R}^n, (\exp tA)V = 0,\ \forall t \in \mathbb{R}\} = \{0\}$$

car la matrice $\exp(tA)$ est inversible. D'après le théorème précédent, cette application est surjective puisque toute solution s'écrit $(\exp tA)V$, l'application u est donc bijective, ce qui prouve que $\dim \mathcal{S} = \dim \mathbb{R}^n = n$.
Les colonnes de la matrice $\exp(tA)$ forment une base de l'espace vectoriel des solutions du système différentiel $X' = AX$.

Voyons un cas particulier

Proposition 6. *Soit $A \in M_n(\mathbb{R})$ une matrice diagonalisable sur $\mathbb{R}$. Notons $(V_1, \ldots, V_n)$ une base de vecteurs propres et $\lambda_1, \ldots, \lambda_n$ les valeurs propres correspondantes. Alors les fonctions $S_i(t) = e^{\lambda_i t}V_i$, $(1 \leq i \leq n)$ forment une base de l'espace des solutions du système $X' = AX$.*

Démonstration.

Montrons que ces solutions sont linéairement indépendantes, en effet si $a_1, \ldots, a_n$ sont des réels tels que

$$a_1S_1(t) + \cdots + a_nS_n(t) = 0,$$

cette égalité étant vraie pour tout t, elle est vraie en particulier pour $t = 0$ où elle devient

$$a_1V_1 + \cdots + a_nV_n = 0,$$

ce qui implique $a_1 = \ldots = a_n = 0$ car les V_i forment une base de $\mathbb{R}^n$. Comme l'espace des solutions est de dimension n, les S_i forment une base.

Soit P la matrice dont les colonnes sont les vecteurs $V_1, \dots, V_n$. Si $X = PY$ est solution de $X' = AX$, alors Y est solution de $Y' = (P^{-1}AP)Y$ où $P^{-1}AP = D$ est diagonale. On a

$$\begin{cases} y_1' = \lambda_1 y_1 \\ \vdots \\ y_n' = \lambda_n y_n \end{cases} \quad \text{d'où} \quad Y(t) = \begin{pmatrix} k_1 e^{\lambda_1 t} \\ \vdots \\ k_n e^{\lambda_n t} \end{pmatrix}$$

d'où

$$PY(t) = k_1 e^{\lambda_1 t} V_1 + \cdots + k_n e^{\lambda_n t} V_n = k_1 S_1(t) + \cdots + k_n S_n(t).$$

Comme corollaire du théorème, nous allons démontrer l'unicité de la solution sous conditions initiales.

Proposition 7. *Soient $A \in M_n(\mathbb{R})$, $t_0 \in \mathbb{R}$ et $X_0 \in \mathbb{R}^n$. Il existe une unique solution S_0 du système différentiel $X' = AX$ qui vérifie $S_0(t_0) = X_0$, c'est la fonction $S_0 : \mathbb{R} \to \mathbb{R}^n$ définie par*

$$S_0(t) = (\exp(t - t_0)A)X_0.$$

Démonstration.

On pose

$$S_0(t) = (\exp(t - t_0)A)X_0 = \exp(tA)\exp(-t_0 A)X_0 = \exp(tA)Y_0.$$

D'après le théorème, S est une solution de $X' = AX$ et

$$S_0(t_0) = \exp(0A)X_0 = I_n X_0 = X_0.$$

S'il existe une autre solution S_1 de $X' = AX$ vérifiant $S_1(t_0) = X_0$, alors on a

$$S_1(t) = \exp(tA)V \text{ et } \exp(t_0 A)V = X_0,$$

c'est-à-dire $V = \exp(-t_0 A)X_0$ donc $V = Y_0$ et $S_1(t) = S_0(t)$ pour tout $t \in \mathbb{R}$.

Nous avons donc démontré que l'espace des solutions est un espace vectoriel de dimension n et que si l'on impose la condition $X(t_0) = X_0$, il y a unicité de la solution.

III.2.2. Cas général, variation de la constante

On cherche maintenant à obtenir la solution générale du système $X' = AX + B$ où

$$A \in M_n(\mathbb{R}) \text{ et } B : I \to \mathbb{R}^n.$$

Connaissant la solution générale du système homogène $X' = AX$, nous allons chercher une solution du système non homogène.

Soit $(S_1, \dots, S_n)$ une base de l'espace vectoriel des solutions de $X' = AX$, toute solution de ce système s'écrit donc

$$S(t) = \sum_{i=1}^{n} \alpha_i S_i(t)$$

où les α_i, $1 \leq i \leq n$ sont des réels.
Nous allons chercher une solution du système $X' = AX + B$ sous la forme

$$S(t) = \sum_{i=1}^{n} \alpha_i(t) S_i(t)$$

où les fonctions α_i sont des fonctions réelles dérivables sur I. On a alors, pour tout $t \in \mathbb{R}$,

$$S'(t) = \sum_{i=1}^{n} \alpha_i(t) S_i'(t) + \sum_{i=1}^{n} \alpha_i'(t) S_i(t) = AS(t) + \sum_{i=1}^{n} \alpha_i'(t) S_i(t)$$

car $S_i'(t) = AS_i(t)$, on identifie alors

$$\sum_{i=1}^{n} \alpha_i'(t) S_i(t) = B(t)$$

Les colonnes de la matrice $R(t) = \exp tA$ forment une base de l'espace des solutions de $X' = AX$, à partir de cette base on cherche une solution de $X' = AX + B$ sous la forme $S(t) = R(t)F(t)$ où

$$\begin{aligned} F : I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto F(t) = \begin{pmatrix} \alpha_1(t) \\ \vdots \\ \alpha_n(t) \end{pmatrix} \end{aligned}$$

on peut écrire sous forme matricielle

$$S'(t) = R'(t)F(t) + R(t)F'(t) = AR(t)F(t) + R(t)F'(t) = AS(t) + R(t)F'(t)$$

ainsi, en identifiant on a

$$R(t)F'(t) = B(t)$$

ce qui nous donne

$$F'(t) = R^{-1}(t)B(t) = \exp(-tA)B(t)$$

Il ne reste plus qu'à intégrer terme à terme $F'(t)$, c'est-à-dire les $\alpha_i(t)$.

III.2.3. Application aux équations différentielles d'ordre n

Nous allons voir, là encore comment des méthodes d'algèbre linéaire permettent de résoudre des problèmes d'analyse.

On considère une équation différentielle d'ordre n à coefficients constants

$$(\star) \qquad y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = 0$$

où la fonction inconnue est une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On introduit les fonctions auxilliaires

$$\begin{cases} y_1 = y \\ y_2 = y_1' = y' \\ \quad\vdots \\ y_{n-1} = y_{n-2}' = y^{(n-2)} \\ y_n = y_{n-1}' = y^{(n-1)} \end{cases}$$

Pour intégrer l'équation ($\star$) on intègre le système

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad\vdots \\ y_{n-1}' = y_n \\ y_n' = -a_1 y_n - a_2 y_{n-1} - \cdots - a_n y_1 \end{cases}$$

c'est-à-dire

$$Y' = AY$$

avec

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

III.3. Exemples en dimension 2

III.3.1 Systèmes homogènes

Soit $A \in M_2(\mathbb{R})$. Les solutions du système $X' = AX$ sont des applications de $\mathbb{R}$ dans $\mathbb{R}^2$ qui peuvent être représentées par des courbes paramétrées. On a

$$X' = AX \iff \begin{cases} x'(t) = a_{11}x(t) + a_{12}y(t) \\ y'(t) = a_{21}x(t) + a_{22}y(t) \end{cases}$$

Si l'application

$$S : \mathbb{R} \longrightarrow \mathbb{R}^2$$
$$t \longmapsto (x(t), y(t))$$

est solution de $X' = AX$, l'ensemble $S(\mathbb{R}) = \{S(t) = (x(t), y(t)) \in \mathbb{R}^2,\ t \in \mathbb{R}\}$ est appelé trajectoire.

Les solutions constantes sont les applications $S(t) = X_0 \in \mathbb{R}^2$ pour tout $t \in \mathbb{R}$, on a donc $S'(t) = 0$ c'est-à-dire $AX_0 = 0$. Le vecteur X_0 est appelé point d'équilibre du système, le point $(0, 0)$ est toujours point d'équilibre.

Proposition 1. *Les trajectoires du système $X' = AX$ sont disjointes ou confondues*

Démonstration.

Soient S_1 et S_2 deux solutions du système homogène $X' = AX$. On a donc, pour tout $t \in \mathbb{R}$, $S_1'(t) = AS_1(t)$ et $S_2'(t) = AS_2(t)$. Alors, ou bien $S_1(\mathbb{R}) \cap S_2(\mathbb{R}) = \emptyset$, ou bien il existe $V \in \mathbb{R}^2$ tel que $V \in S_1(\mathbb{R}) \cap S_2(\mathbb{R})$, nous allons montrer que dans ce cas, les trajectoires sont confondues, c'est-à-dire que l'on a $S_1(\mathbb{R}) = S_2(\mathbb{R})$.
D'après l'étude des systèmes homogènes, on sait que si S_1 et S_2 sont solutions, il existe des vecteurs V_1 et V_2 dans $\mathbb{R}^2$ tels que, pour tout $t \in \mathbb{R}$,

$$S_1(t) = \exp(tA)V_1 \quad \text{et} \quad S_2(t) = \exp(tA)V_2.$$

Si $V \in S_1(\mathbb{R}) \cap S_2(\mathbb{R})$, alors, il existe $t_1 \in \mathbb{R}$ et $t_2 \in \mathbb{R}$ tels que

$$V = S_1(t_1) = S_2(t_2) = (\exp t_1 A)V_1 = (\exp t_2 A)V_2.$$

Or, S_2 est l'unique solution prenant la valeur V en t_2, on a donc pour tout $t \in \mathbb{R}$,

$$S_2(t) = (\exp(t - t_2)A)V.$$

Ainsi pour tout $t \in \mathbb{R}$, on a

$$S_2(t + t_2 - t_1) = (\exp(t - t_1)A)V = (\exp(t - t_1)A)(\exp t_1 A)V_1 = (\exp tA)V_1 = S_1(t).$$

Ce qui prouve que les trajectoires sont confondues, en effet, pour tout $t \in \mathbb{R}$, on a $S_1(t) \in S_2(\mathbb{R})$ et $S_2(t) \in S_2(\mathbb{R})$.

III.3.2. Equations différentielles du second ordre

On souhaite intégrer l'équation

$$x''(t) + px'(t) + qx(t) = b(t)$$

où p et q sont des constantes réelles et où la fonction inconnue x est une fonction réelle à valeurs réelles et la fonction b est définie et continue sur un intervalle I de $\mathbb{R}$. Pour cela on pose $y = x'$ et on intègre le système

$$\begin{cases} x' = y \\ y' = -qx - py + b(t) \end{cases}$$

On est ainsi ramené à l'étude des systèmes du paragraphe 2.

Proposition 2. *La fonction $s : I \to \mathbb{R}$ est solution de l'équation différentielle*

$$x'' + px' + qx = b$$

si et seulement si l'application $S : I \to \mathbb{R}^2$ définie par

$$S(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} s(t) \\ s'(t) \end{pmatrix}$$

est solution du système $X' = AX + B$ avec

$$A = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 \\ b(t) \end{pmatrix}.$$

La démonstration est une vérification immédiate et la matrice A est très simple à étudier.

Son polynôme caractéristique est égal à : $P_A(X) = X^2 + pX + q$, l'espace vectoriel des solutions de l'équation homogène d'ordre 2, $x'' + px' + q = 0$ est de dimension 2, on peut en donner une base en fonction des racines du polynôme caractéristique.

- Si λ_1 et λ_2 sont deux racines distinctes, l'espace des solutions de l'équation homogène est engendré par les fonctions

$$t \mapsto e^{\lambda_1 t} \text{ et } t \mapsto e^{\lambda_2 t}.$$

- Si λ est une racine réelle double il est engendré par

$$t \mapsto e^{\lambda t} \text{ et } t \mapsto te^{\lambda t}.$$

- Si $a + ib$ et $a - ib$ sont deux racines complexes conjuguées, il est engendré par

$$t \mapsto e^{at} \cos bt \text{ et } t \mapsto e^{at} \sin bt.$$

III.4 Etudes d'exemples

Nous allons commencer par un exemple de système linéaire avec second membre par la méthode de variation des constantes.

Exemple 1. Soit le système

$$(\epsilon) \qquad \begin{cases} x' = 3x + y + te^t \\ y' = -x + y + e^t \end{cases}$$

Notons ϵ_0 le système homogène

$$(\epsilon_0) \qquad \begin{cases} x' = 3x + y \\ y' = -x + y \end{cases} \iff X' = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} X = AX$$

Considérons la matrice $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$, son polynôme caractéristique est égal à

$$P_A(X) = (3 - X)(1 - X) + 1 = X^2 - 4X + 4 = (X - 2)^2.$$

D'après le théorème de Hamilton-Cayley, on a $P_A(A) = 0$, c'est-à-dire $(A - 2I_2)^2 = 0$. Posons $N = A - 2I_2$, c'est une matrice nilpotente et on a : $A = N + 2I_2$, comme N et I_2 commutent, c'est la décomposition de Dunford.

La solution générale de l'équation homogène s'écrit

$$X(t) = \exp(tA)V$$

où V est un vecteur de $\mathbb{R}^2$. Calculons $\exp(tA)$.

$$\exp(tA) = \exp(t(N + 2I_2)) = \exp(tN + 2tI_2) = \exp(tN)\exp(2tI_2)$$

or, $(tN)^2 = t^2N^2 = 0$ d'où

$$\exp(tA) = e^{2t}I_2(I + tN) = e^{2t}\begin{pmatrix} t+1 & t \\ -t & 1-t \end{pmatrix}.$$

La solution générale du système (ϵ_0) s'écrit donc

$$X(t) = e^{2t}\begin{pmatrix} t+1 & t \\ -t & 1-t \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = e^{2t}\begin{pmatrix} a(t+1) + bt \\ -at + b(1-t) \end{pmatrix} = ae^{2t}\begin{pmatrix} t+1 \\ -t \end{pmatrix} + be^{2t}\begin{pmatrix} t \\ 1-t \end{pmatrix},$$

avec $V = (a, b) \in \mathbb{R}^2$. Les applications

$$t \mapsto e^{2t}\begin{pmatrix} t+1 \\ -t \end{pmatrix} \quad \text{et} \quad t \mapsto e^{2t}\begin{pmatrix} t \\ 1-t \end{pmatrix}$$

sont deux solutions linéairement indépendantes, elles forment une base de l'espace des solutions de (ϵ_0).

On cherche une solution particulière de l'équation complète (ϵ) en faisant varier le vecteur $V = (a, b)$ c'est-à-dire les constantes a et b que l'on cherche sous forme de fonctions $a(t)$ et $b(t)$. En écrivant les choses sous forme matricielle, on obtient $X(t) = \exp(tA)V(t)$ d'où

$$X'(t) = \underbrace{A\exp(tA)V(t)}_{AX(t)} + \underbrace{\exp(tA)V'(t)}_{B(t)}.$$

Ainsi, on a $V'(t) = \exp(-tA)B(t)$ d'où

$$V'(t) = e^{-2t}\begin{pmatrix} -t+1 & -t \\ t & 1+t \end{pmatrix}\begin{pmatrix} te^t \\ e^t \end{pmatrix} = e^{-t}\begin{pmatrix} 1-t & -t \\ t & 1+t \end{pmatrix}\begin{pmatrix} t \\ 1 \end{pmatrix} = e^{-t}\begin{pmatrix} -t^2 \\ t^2+1+t \end{pmatrix}.$$

On cherche donc des fonctions $a(t)$ et $b(t)$ qui vérifient

$$\begin{cases} a'(t) = -e^{-t}t^2 \\ b'(t) = e^{-t}(t^2+t+1) \end{cases}$$

On les cherche sous la forme $e^{-t}P(t)$ où $P(t)$ est un polynôme de degré 2 et l'on obtient

$$a(t) = (t^2+2t+2)e^{-t} \quad \text{et} \quad b(t) = -(t^2+3t+4)e^{-t}.$$

Ainsi, la solution générale de l'équation complète s'écrit

$$X(t) = \underbrace{e^{2t}\left(\alpha\begin{pmatrix} t+1 \\ -t \end{pmatrix} + \beta\begin{pmatrix} t \\ 1-t \end{pmatrix}\right)}_{\text{solution de } (\epsilon_0)} + \underbrace{e^{2t}\left(a(t)\begin{pmatrix} t+1 \\ -t \end{pmatrix} + b(t)\begin{pmatrix} t \\ 1-t \end{pmatrix}\right)}_{\text{solution de } (\epsilon)}$$

où α et β sont des constantes réelles, $a(t)$ et $b(t)$ les fonctions trouvées ci-dessus. Simplifions la solution particulière de (ϵ)

$$S_1(t) = e^{2t}\left(e^{-t}(t^2+2t+2)\begin{pmatrix} t+1 \\ -t \end{pmatrix} + e^{-t}(-t^2-3t-4)\begin{pmatrix} t \\ 1-t \end{pmatrix}\right) = e^t\begin{pmatrix} 2 \\ -t-4 \end{pmatrix}.$$

La solution générale de l'équation (ϵ) s'écrit donc

$$S(t) = e^t\begin{pmatrix} 2 \\ -t-4 \end{pmatrix} + e^{2t}\left(\alpha\begin{pmatrix} t+1 \\ -t \end{pmatrix} + \beta\begin{pmatrix} t \\ 1-t \end{pmatrix}\right)$$

α et β étant des constantes réelles arbitraires.

Nous allons maintenant étudier quelques exemples simples d'équations homogènes dans $\mathbb{R}^2$ qui nous amèneront à tracer des trajectoires (courbes paramétrées).

Exemple 2. Considérons le système :

$$\begin{cases} x' = x \\ y' = 2y \end{cases} \iff \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

c'est un cas où la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ est diagonale. On résout les deux équations et on obtient

$$\begin{cases} x(t) = ae^t \\ y(t) = be^{2t} \end{cases}$$

où a et b sont des constantes réelles. Etudions plus précisément les trajectoires obtenues selon les valeurs des constantes a et b.

- Si $a = b = 0$, la solution du système obtenue est l'application nulle de $\mathbb{R} \to \mathbb{R}^2$, sa trajectoire est réduite au point $(0,0) \in \mathbb{R}^2$.

- Si $a = 0$ et $b \neq 0$, alors pour tout $t \in \mathbb{R}$, $x(t) = 0$ et $y(t)$ est du signe de b ainsi les trajectoires sont les demi-axes Ox et $-Ox$.

Si $a \neq 0$ et $b \neq 0$, alors pour tout $t \in \mathbb{R}$, on a $y(t) = b\left(\dfrac{x(t)}{a}\right)^2$, les trajectoires sont donc les courbes d'équation $y = b\left(\dfrac{x}{a}\right)^2$, c'est-à-dire des branches de paraboles.

Exemple 3. Considérons maintenant le système :

$$\begin{cases} x' = 0 \\ y' = x \end{cases}$$

On a alors, pour tout $t \in \mathbb{R}$, $x(t) = k$, k étant une constante réelle et donc $y(t) = kt + h$ où h est également une constante réelle. Les trajectoires sont les droites $x = k$, parcourues de $y = -\infty$ vers $y = +\infty$ si $k > 0$ et dans l'autre sens si $k < 0$.

Exemple 4. Nous allons maintenant étudier le système

$$\begin{cases} x' = y \\ y' = -x \end{cases}$$

Avant d'intégrer ce système en applicant les méthodes du cours, on peut remarquer que les applications suivantes sont solution.

$$S_1 : t \longmapsto \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} \quad \text{et} \quad S_1 : t \longmapsto \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

Vérifions que ces deux solutions sont linéairement indépendantes, soient a et b des réels tels que pour tout $t \in \mathbb{R}$,

$$aS_1(t) + bS_2(t) = 0$$

alors, pour $t = 0$ on a

$$aS_1(0) + bS_2(0) = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ce qui implique $a = b = 0$.

On a ainsi toutes les solutions du système comme combinaison linéaire de cette base de l'espace des solutions.

Si on ne pense pas à ces solutions évidentes, alors on résout le système $X' = AX$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. On détermine les racines du polynôme caractéristique de A. On a

$$P_A(X) = X^2 + 1 = (X - i)(X + i).$$

ce polynôme n'a pas de racines réelles, par contre il admet deux racines distinctes dans $\mathbb{C}$, la matrice A est donc diagonalisable dans $\mathbb{C}$, et c'est cette diagonalisation que nous allons utiliser, sachant que les parties réelles et imaginaires des solutions complexes sont les solutions réelles recherchées. En effet si $Z(t) = X(t) + iY(t)$ alors

$$Z'(t) = AZ(t) \Rightarrow X'(t) + iY'(t) = AX(t) + iAY(t)$$

et, comme A est une matrice à coefficients réels, on obtient par identification $X' = AX$ et $Y' = AY$. De plus les solutions complexes sont conjuguées puisque le système est à coefficients réels.

On peut diagonaliser A ou plus simplement déterminer les vecteurs propres V_i et V_{-i} associés aux valeurs propres i et $-i$, on obtient ainsi deux solutions linéairement indépendantes, et la solution générale s'écrit

$$Z(t) = \alpha e^{it} V_i + \beta e^{-it} V_{-i}$$

où α et β sont des contantes complexes. On a

$$AV = iV \iff \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ix \\ iy \end{pmatrix} \iff y = ix$$

et

$$AV = -iV \iff \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -ix \\ -iy \end{pmatrix} \iff y = -ix$$

on prend $V_i = \begin{pmatrix} 1 \\ i \end{pmatrix}$ et $V_{-i} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. La solution générale complexe s'écrit alors

$$Z(t) = \alpha e^{it} \begin{pmatrix} 1 \\ i \end{pmatrix} + \beta e^{-it} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

où α et β sont des constantes complexes. Cette solution s'écrit encore

$$Z(t) = (\alpha + \beta) \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + i(\alpha - \beta) \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}.$$

La solution générale réelle s'écrit donc

$$S(t) = a \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + b \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

où a et b sont des constantes réelles.

Remarquons que comme les solutions complexes sont conjuguées deux à deux, une combinaison linéaire de deux solutions complexes indépendantes fournit une combinaison linéaire réelle de deux solutions réelles indépendantes (et non de quatre).

Etudions maintenant la forme des trajectoires, soit $(a, b) \in \mathbb{R}^2$ considérons la solution S définie par $S(t) = (x(t), y(t))$ et

$$\begin{cases} x(t) = a \cos t + b \sin t \\ y(t) = -a \sin t + b \cos t \end{cases}$$

Alors pour tout $t \in \mathbb{R}$ on a,

$$\begin{aligned} x^2 + y^2 &= (a \cos t + b \sin t)^2 + (-a \sin t + b \cos t)^2 \\ &= a^2 \cos^2 t + b^2 \sin^2 t + 2ab \cos t \sin t + a^2 \sin^2 t + b^2 \cos^2 t - 2ab \sin t \cos t \\ &= (a^2 + b^2)(\cos^2 t + \sin^2 t) = a^2 + b^2. \end{aligned}$$

Ce qui prouve que la trajectoire de S est le cercle de centre $O = (0, 0)$ et de rayon $\sqrt{a^2 + b^2}$.

Exemple 5. Nous allons maintenant, à partir du système étudié dans l'exemple précédent intégrer l'équation du second ordre $x'' + x = 0$. Les fonctions $t \mapsto \sin t$ et $t \mapsto \cos t$ sont deux solutions linéairement indépendantes.

Ecrivons l'équation sous forme de système linéaire

$$\begin{pmatrix} x' \\ x'' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix},$$

ou encore $X' = AX$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ et $X = \begin{pmatrix} x \\ x' \end{pmatrix}$. On retrouve les solutions de notre système précédent via la matrice $R(t) = \exp(tA)$, c'est-à-dire après calculs

$$R(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$

dont les colonnes sont des solutions linéairement indépendantes du système $X' = AX$. La solution générale s'écrit

$$S(t) = a\begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + b\begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

avec $(a, b) \in \mathbb{R}^2$.

Si l'équation a un second membre, c'est-à-dire, si l'on veut intégrer $x'' + x = f(t)$, où f est une fonction réelle continue, les solutions seront cherchées en faisant varier les constantes a et b. Les fonctions $a(t)$ et $b(t)$ étant obtenue par identification en écrivant

$$\begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = R(-t)\begin{pmatrix} 0 \\ f(t) \end{pmatrix}$$

c'est-à-dire

$$\begin{cases} a'(t) = -(\sin t)f(t) \\ b'(t) = (\cos t)f(t) \end{cases}$$

Bibliographie

Ces notes de cours ont été rédigées à l'aide des ouvrages suivants :

- F.Liret, D.Martinet, *Algèbre et géométrie, 2e année*, Dunod.

- J.Lelong-Ferrand, J.M.Arnaudiès, *Cours de Mathématiques, tome 1 : Algèbre*, Dunod.

- X.Gourdon, *Algèbre*, "Les maths en tête", Ellipses

- J.C.Savioz, *Algèbre linéaire*, Vuibert.