Deuxième année de Licence MIASHS

Algèbre 3[1]

Julien GREPAT[2]

Contents

[2] Contact : `julien.grepat@univ-grenoble-alpes.fr`

Par son caractère conclusif dans la formation en algèbre linéaire de cette licence, ce cours permet de prendre un certain recul. Nous proposons des rappels sur les espaces vectoriels et le calcul matriciel de première année dans le cadre restreint des espaces vectoriels $\mathbb{R}^N$. L'enjeu du cours est la réduction des endomorphismes (diagonalisation-trigonalisation de leur matrice).

Ce cours doit vous conforter dans la maîtrise de l'algèbre linéaire pour transférer les résultats aux différents domaines d'utilisation. Ce cours doit également permettre de développer la vision géométrique des concepts abstraits de l'algèbre linéaire et ainsi amener à leur interprétation dans différents cadres.

Part I

Résumé des épisodes précédents...

1 Les espaces vectoriels

Dans le plan

$$\mathbb{R}^2 = \{(x, y) \ ; \ x \in \mathbb{R} \ , \ y \in \mathbb{R}\} \, ,$$

on définit de façon naturelle une addition et un produit de la façon suivante.

- Si $u = (x_1, y_1)$ et $v = (x_2, y_2)$ on pose

$$u + v = (x_1 + x_2, y_1 + y_2).$$

Les propriétés de l'addition des réels montrent que pour tout $u \in \mathbb{R}^2$ et tout $v \in \mathbb{R}^2$ on a $u + v = v + u$. De même, si u, v et w sont des éléments de $\mathbb{R}^2$ alors

$$(u + v) + w = u + (v + w).$$

En notant $-u = (-x_1, -y_1)$, on a, pour tout $u \in \mathbb{R}^2$, $u - u = (0, 0)$. Si on note $0 = (0, 0)$, alors, pour tout $u \in \mathbb{R}^2$ on a $u + 0 = 0 + u = u$.

- On peut aussi définir une multiplication par les réels :

$$\lambda u = (\lambda x_1, \lambda y_1),$$

  pour tout $\lambda \in \mathbb{R}$ et tout $u \in \mathbb{R}^2$.

On a de plus $\lambda(u + v) = \lambda u + \lambda v$ et $(\lambda\mu)u = \lambda(\mu u)$.

Les opérations définies sur les vecteurs dans $\mathbb{R}^3$ en terminale ont les mêmes propriétés. Ce type d'opérations (addition et multiplication par des scalaires) sont très courantes en mathématiques : on peut les définir aussi pour les suites numériques, pour les fonctions définies sur un intervalle etc ... On est donc naturellement amené à définir une structure qui englobe ces objets a priori distincts.

1.1 Espace vectoriel

Definition 1.1 *Un espace vectoriel sur $\mathbb{R}$ est un ensemble E muni d'une addition*

$$+ : E \times E \to E,$$

et d'une multiplication par les scalaires (les réels)

$$\cdot : \mathbb{R} \times E \to E$$

vérifiant les propriétés suivantes.

(i) L'addition est associative : quels que soient u, v, w dans E,

$$u + (v + w) = (u + v) + w.$$

(ii) Neutre de l'addition : il existe un élément de E noté 0 tel que pour tout $u \in E$,

$$u + 0 = 0 + u = u.$$

(iii) Opposé : pour tout $u \in E$, il existe un unique élément de E, noté $-u$ tel que

$$u + (-u) = (-u) + u = 0.$$

(iv) L'addition est commutative : quels que soient u, v dans E,

$$u + v = v + u.$$

(v) Neutre pour la multiplication : pour tout $u \in E$,

$$1 \cdot u = u.$$

(vi) Quels que soient λ, μ dans $\mathbb{R}$ et $u \in E$,

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

(vii) Distributivité : quels que soient $\lambda \in \mathbb{R}$, $u \in E$ et $v \in E$,

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

1.2 Cadre du cours

On restreindra le cours au cas des espaces vectoriels $\mathbb{R}^n$. Sur $\mathbb{R}^n$, on définit une addition et une multiplication par les réels en posant

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots x_n + y_n)$$

et

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

$\mathbb{R}^n$ muni de ces deux opérations est un espace vectoriel.

1.3 Sous espace vectoriel

Une partie F d'un espace vectoriel E est un *sous espace vectoriel* si, muni des lois induites par celles de E, c'est un espace vectoriel.

Proposition 1.2 *Soit E un $\mathbb{R}$ espace vectoriel et F une partie de E. Pour que F soit un sous espace vectoriel il faut et il suffit que F contienne 0 et que quels que soient $u, v \in F$ et $\lambda, \mu \in \mathbb{R}$*

$$\lambda u + \mu v \in F.$$

1.4 Combinaison linéaire

On comprend alors aisément que la structure d'espace vectoriel ou de sous espace vectoriel est une structure stable par combinaison linéaire, *i.e.* toute combinaison linéaire d'éléments de l'espace (ou du sous-espace) doit rester dans l'espace (ou le sous-espace) vectoriel.

Definition 1.3 *Soit E un espace vectoriel et $\{u_1, \cdots, u_k\}$ une famille d'éléments de E. On appelle combinaison linéaire des vecteurs $\{u_i\}$ toute somme finie de la forme $\lambda_1 u_1 + \cdots + \lambda_k u_k$ où les λ_j sont des scalaires (réels).*

Soit $\{u_1, \cdots u_k,\}$ des vecteurs de E. Nous utiliserons les dénominations suivantes :

(i) On dit que la combinaison linéaire $\lambda_1 u_1 + \cdots + \lambda_k u_k$ est nulle si son résultat est le vecteur nul, cest-à-dire si

$$\lambda_1 u_1 + \cdots + \lambda_k u_k = 0.$$

(ii) On dit que la combinaison linéaire $\lambda_1 u_1 + \cdots + \lambda_k u_k$ est triviale si tous les coefficients λ_i sont nuls, cest-à-dire si

$$\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.$$

Il est évident que la combinaison linéaire triviale est nulle.

1.5 Sous espace engendré

Puisque, pour un sous espace vectoriel, toute combinaison linéaire de ses vecteurs doit rester dans le sous espace. On peut alors considérer, par essence, que l'ensemble des combinaisons linéaires d'une famille d'éléments forme un sous espace vectoriel dès lors qu'on remarque qu'une combinaison de combinaisons linéaires d'éléments de cette famille est en fait une combinaison linéaire de ces éléments par le jeu des développements et factorisations. On donne alors le résultat suivant et une caractérisation dans $\mathbb{R}^N$.

Theorem 1.4 *Si X est une partie d'un espace vectoriel E, l'ensemble $Vect(X)$ des combinaisons linéaires des éléments de X est un sous-espace vectoriel de E ; on l'appelle le sous-espace engendré par X.*

Dans $\mathbb{R}^N$, le sous-espace vectoriel engendré par un nombre fini de vecteurs $H = \{u_1, \cdots u_k,\}$ est

$$Vect(H) = \left\{ \sum_{i=1}^{k} \lambda_i u_i, \quad \lambda_i \in \mathbb{R} \right\}.$$

1.6 Famille libre, génératrice, base

On a des espaces vectoriels, des sous espaces vectoriels et le sous espace vectoriel engendré par une famille de vecteurs. C'est bien mais on doit aller plus loin ! On peut se demander, pour un sous espace vectoriel, quelle est la famille qui l'engendre. C'est à dire, avec quels vecteurs on engendre le sous espace vectoriel en question. Cela nous amène à la notion de famille génératrice.

Definition 1.5 *Soit E un espace vectoriel. Une famille de vecteur $H = \{u_i\}_i$ est dite* génératrice si tout élément de E est combinaison linéaire des u_i :

$$E = Vect(H) := \langle u_1, \cdots, u_k \rangle$$

.

Note : Le signe := signifie qu'on introduit une notation.

Imaginons, maintenant, que dans cette famille H, un des vecteur soit combinaison linéaire des autres. Il serait bien inutile pour générer notre espace. On dit que la famille serait liées :

$$\exists \lambda_1, \cdots \lambda_k \text{ non tous nuls : } \qquad u_j = \sum_{i \neq j} \lambda_i u_i.$$

Cela revient immédiatement à la défnition suivante.

Definition 1.6 *Dans un espace vectoriel E on dit qu'un système de vecteurs $(u_1, \ldots, u_k)$ est un système lié s'il existe des scalaires $\lambda_1, \cdots, \lambda_k$ non tous nuls tels que*

$$\lambda_1 u_1 + \cdots + \lambda_k u_k = 0.$$

Un système qui n'est pas lié est dit libre*.*

Pour que le système $(u_1, \ldots, u_k)$ soit libre, il faut et il suffit que

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = 0 \Longrightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.$$

Chouette ! On peut retirer tout vecteur combinaison linéaire des autres de la famille tout en gardant son caractère générateur. Lorsque tout vecteur linéairement lié est retiré de la famille génératrice, on obtient une base du sous espace vectoriel.

Definition 1.7 *On dit que $(u_i)_i$ est une* base *de E si c'est un système libre et générateur.*

Dire que (u_i) est une base revient à dire que tout élément u de E peut s'écrire *de façon unique* comme combinaison linéaire des u_i. Les coefficients de cette combinaison linéaires sont les *coordonnées* de u dans la base (u_i).

Note : Dans le strict cas de deux vecteurs, deux vecteurs sont linéairement liés s'ils sont colinéaires. Dans le cas de trois vecteurs ou plus, ils peuvent être liés sans qu'aucun couple ne soit colinéaire. C'est le cas dans l'exemple suivant :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad u = e_1 + e_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

ne comporte aucun couple de vecteurs colinéaires.

1.7 Dimension

Definition 1.8 *Un espace vectoriel non nul E est dit de dimension finie s'il existe une famille génératrice finie $(e_1, \cdots, e_n)$. On appelle alors dimension de E et on note $dim(E)$ le plus petit des entiers n tels qu'il existe effectivement une famille génératrice $(e_1, \cdots, e_n)$ de cardinal n.*

Remarque : Par convention, nous posons $dim\{0\} = 0$.

On peut montrer que tout espace vectoriel admet une base et que si un espace vectoriel admet une base finie, toute base de E est finie et a le même nombre d'éléments, soit la dimension de E. Il vient

Proposition 1.9 *Soit E un espace vectoriel de dimension n et soit $(u_1, \cdots u_p)$ une famille de vecteurs.*

(i) Si $(u_1, \cdots, u_p)$ est libre, alors $p \leq n$

(ii) Si $(u_1, \cdots u_p)$ est libre et si $p = n$, alors c'est une base de E.

(iii) Si une famille de n vecteurs $(u_1, \cdots u_n)$ est génératrice, c'est une base.

1.8 Dimension et équations cartésiennes

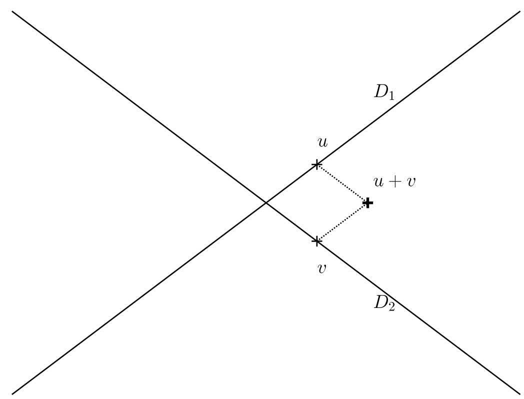
On peut définir un sous espace vectoriel de $\mathbb{R}^n$ par une équation cartésienne. On obtient alors un hyperplan de dimension $n-1$. On peut aisément vérifier que l'intersection de sous-espaces vectoriels est un sous espace vectoriel. Cela revient à considédérer un système d'équations cartésiennes. Un système de k équations cartésiennes linéairement indépendantes (*i.e.* non combinaisons linéaires les unes des autres – *cf* pivot de Gauss) décrit un sous espace vectoriel de dimension $n - k$.

2 Somme directe de sous espaces vectoriels

2.1 Union d'espaces vectoriels

La remarque fondamentale à faire en entamant ce paragraphe est que, si F et G sont deux sous espaces vectoriels d'un espace vectoriel E, alors $F \cup G$ n'est presque jamais un sous-espace vectoriel. En fait, en dehors des cas très particuliers $F \subset G$, ou $G \subset F$, $F \cup G$ n'est jamais un sous-espace vectoriel. Voici un contre-exemple élémentaire.

Dans $\mathbb{R}^2$ la réunion de la droite D_1 d'équation $x - y = 0$ et de la droite D_2 d'équation $x + y = 0$ *n'est pas* un sous-espace vectoriel. En effet, $u = (1, 1) \in D_1$, $v = (1, -1) \in D_2$ mais la somme $u + v$ n'appartient ni à D_1 ni à D_2.



2.2 Somme de deux sous espaces vectoriels

On peut définir la somme de deux sous espaces vectoriels de la façon suivante.

Proposition 2.1 *Soit F et G deux sous-espaces vectoriels d'un espace vectoriel E. L'ensemble*

$$F + G = \{x_F + x_G ; x_F \in F, x_G \in G\}$$

est un sous-espace vectoriel de E.

En fait, il s'agit du plus petit sous-espace vectoriel contenant à la fois F et G, c'est-à-dire

$$F + G = Vect(F \cup G).$$

Proof. Soit $x, y \in F + G$ et $\lambda \in \mathbb{R}$. Il existe $x_F, y_F \in F$ et $x_G, y_G \in G$ tels que $x = x_F + x_G$ et $y = y_F + y_G$. Par conséquent, nous avons

$$x + \lambda y = \underbrace{(x_F + \lambda y_F)}_{\in F} + \underbrace{(x_G + \lambda y_G)}_{\in G} \in F + G.$$

Comme par ailleurs, $0 \in F + G$, il en résulte que $F + G$ est bien un sous espace vectoriel de E. □

2.3 Somme directe de deux sous espaces vectoriels

En général, il est possible qu'un élément de $F+G$ puisse s'écrire de plusieurs façons possibles comme somme d'un élément de F et d'un élément de G.

Prenons un exemple, considérons l'espace vectoriel $E = R^3$ avec sa base canonique (e_1, e_2, e_3) et les sous espaces $F = Vect(e_1, e_2)$, $G = Vect(e_2, e_3)$. Le vecteur e_2 peut s'écrire par exemple

$$e_2 = \underbrace{0}_{F} + \underbrace{e_2}_{G}$$

ou

$$e_2 = \underbrace{e_2/2}_{F} + \underbrace{e_2/2}_{G}.$$

Ce qui rend cette double écriture possible, c'est que le vecteur e_2 appartient à la fois à F et à G. Ceci nous pousse à la définition suivante.

Definition 2.2 *Soit F et G deux sous-espaces vectoriels d'un espace vectoriel E. On dit que F et G sont en somme directe lorsque $F \cap G = \{0\}$. La somme de F et G se note alors $F \oplus G$ à la place de $F+G$ et on parle de la somme directe de F et G.*

Et on définit les projection de la façon suivante.

Proposition 2.3 *Soit F et G deux sous-espaces vectoriels d'un espace vectoriel E. On suppose que la somme de F et G est directe : $F \oplus G$. Alors tout élément x de $F+G$ s'écrit de manière unique sous la forme $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$. On dit alors que x_F (resp. x_G) est le projeté de x sur F (resp. G) parallèlement à G (resp. F).*

On peut d'ailleurs noter que x_F et x_G forment une famille libre.

2.4 Espaces supplémentaires

Definition 2.4 *On dit que deux sous-espaces vectoriels F et G d'un espace vectoriel E sont supplémentaires lorsque $F \oplus G = E$.*

Cela signifie que F et G sont en somme directe et que leur somme est égale à l'espace tout entier. Cela revient encore au même de dire que tout vecteur de l'espace E se décompose de manière unique comme somme d'un élément de F et d'un élément de G. Vient alors immédiatement la proposition suivante.

Proposition 2.5 *Soit deux sous espaces vectoriels supplémntaire F et G d'un espace vectoriel E : $F \oplus G = E$. Si $(e_1, \cdots e_p)$ et une base de F et $(\varepsilon_1, \cdots \varepsilon_q)$ et une base de G alors,*

- *$(e_1, \cdots, e_p, \varepsilon_1, \cdots, \varepsilon_q)$ est une base de E ;*
- *$dim(E) = dim(F) + dim(G)$.*

2.5 Généralisation à la somme directe de plusieurs sous espaces vectoriels

Les notions précédentes se généralisent de la façon suivante lorsqu'on considère plusieurs sous espaces vectoriels.

Definition 2.6 *Soient $E_1, \dots, E_n$ des sous espaces d'un espace vectoriel E. La somme des sous-espaces E_i est le sous-espace de E engendré par*

$$\bigcup_i E_i.$$

On note ce sous espace $\sum E_i$, c'est l'ensemble des combinaisons linéaires d'éléments des E_i. On a donc

$$\sum E_i = \{u_1 + \dots + u_n\, ;\ u_i \in E_i\}\,.$$

La somme $F = \sum_1^n E_i$ est dite *directe* si tout élément de F s'écrit de façon *unique*

$$u = u_1 + \dots + u_n,$$

avec $u_i \in E_i$. On note alors $F = E_1 \oplus E_2 \dots \oplus E_n$.

Proposition 2.7 *Soient $E_1, \dots, E_n$ des sous-espaces vectoriels de E Pour que la somme $\sum E_i$ soit directe il faut et il suffit que pour tout $i \in \{1, \dots, n\}$ on ait*

$$E_i \cap \left(\sum_{j \neq i} E_j\right) = \{0\}.$$

Les résultats sur le recollement des bases et la somme des dimensions de la Proposition 2.5 se généralisent directement dans le cadre $E = E_1 \oplus E_2 \dots \oplus E_n$.

3 Application linéaire

On a vu que la combinaison linéaire était l'outil clé des espaces vectoriels. L'application linéaire est le morphisme[3] entre espaces vectoriels. Il doit donc nécessairement respecter les combinaisons linéaires. On introduit donc la définition suivante.

3.1 Définition

Definition 3.1 *Soient E et F deux espaces vectoriels. Une application φ de E dans F est dite* linéaire *si $\varphi(0) = 0$ et pour tout $u, v \in E$ et tout $\lambda, \mu \in \mathbb{R}$ on a*

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

[3]En algèbre, le morphisme est une application entre deux structures algébriques de même espèce, c'est-à-dire des ensembles munis de lois, qui respectent les propriétés liées à la structure en passant d'un espace à l'autre.

Il suit qu'il suffit de connaître les images des vecteurs de la base de E pour connaître φ entièrement. En effet, si $u \in E$ s'écrit (de manière unique)

$$u = \sum_i \lambda_i e_i, \qquad \lambda_i \in \mathbb{R},$$

où $(e_i)_i$ est la base de E, alors

$$\begin{aligned} \varphi(u) &= \varphi\left(\sum_i \lambda_i e_i\right) \\ &= \sum_i \varphi(\lambda_i e_i) \\ &= \sum_i \lambda_i \varphi(e_i). \end{aligned}$$

Et, réciproquement, on a le résultat suivant.

Proposition 3.2 *Si $\{e_1, \ldots, e_n\}$ est une base de E et $u_1, \cdots, u_n$ sont des éléments de F, il existe une application linéaire φ unique telle que $\varphi(e_1) = u_1, \ldots, \varphi(e_n) = u_n$.*

3.2 Lien avec les matrices

Toute application linéaire de $\varphi : E = \mathbb{R}^n \to F = \mathbb{R}^p$ peut être représentée par une matrice A de la façon suivante. En posant $(e_1, \cdots, e_n)$ une base de E, pour tout $u = u_1 e_1 + \cdots + u_n e_n \in E$

$$\varphi(u) = A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},$$

avec A la matrice portant en colonne les images $\varphi(e_i)$ des vecteurs de base de E :

$$A = \begin{array}{c} \begin{matrix} \varphi(e_1) & \cdots & \varphi(e_n) \end{matrix} \\ \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,n} \\ & \vdots & \\ a_{p,1} & \cdots & a_{p,n} \end{pmatrix} \end{array}.$$

Il suffit alors de finir de se persuader en vérifiant que

$$\varphi(e_1) = A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{p,1} \end{pmatrix}$$

puis

$$\varphi(e_2) = A \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{p,2} \end{pmatrix} \cdots$$

Note : On pourra utiliser sans distinction la matrice ou l'application linéaire. On ne fera pas de différence entre les deux objets mathématiques.

3.3 Formules de calcul matriciel

Si $A = (a_{i,j})$ et $B = (b_{i,j})$ sont deux matrices $n \times p$, on définit la *somme* $A + B$ par

$$A + B = (a_{i,j} + b_{i,j}).$$

Le produit d'une matrice $A = (a_{i,j})$ de type $n \times p$ par un scalaire λ est la matrice

$$\lambda \cdot A = (\lambda a_{i,j}).$$

L'ensemble des matrices de type $n \times p$ est donc muni d'une structure naturelle d'espace vectoriel de dimension np.

Pour le produit entre matrcies, c'est un peu plus compliqué. Si $A = (a_{i,j})$ est une matrice $n \times p$ et $B = (b_{k,l})$ une matrice $p' \times q$, A et B sont multipliables si et seulement si $p = p'$. Leur produit $C = AB$ est la matrice $n \times q$ définie par

$$C = (c_{i,j}), \quad c_{i,j} = a_{i,1}b_{1,j} + \cdots + a_{i,p}b_{p,j} = \sum a_{i,k}b_{k,j}.$$

En utilisant les notations précédentes on a donc

$$\begin{pmatrix} c_{1,1} & \cdots & c_{1,q} \\ \cdots & \mathbf{c_{i,j}} & \cdots \\ c_{n,1} & \cdots & c_{n,q} \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \mathbf{a_{i,1}} & \mathbf{a_{i,k}} & \mathbf{a_{i,p}} \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \cdot \begin{pmatrix} b_{1,1} & \cdots & \mathbf{b_{1,j}} & \cdots & b_{1,q} \\ \cdots & \cdots & \mathbf{b_{k,j}} & \cdots & \cdots \\ b_{n,1} & \cdots & \mathbf{b_{p,j}} & \cdots & b_{p,q} \end{pmatrix}.$$

Le produit est associatif et distributif par rapport à l'addition : étant données trois matrices A, B et C on a

$$\begin{aligned} (AB)C &= A(BC), \\ A(B + C) &= AB + AC, \\ (A + B)C &= AC + BC, \end{aligned}$$

dès que les sommes et les produits écrits dans ces formules ont un sens.

Note : On appelle matrice *carrée* toute matrice de type $n \times n$.

3.4 Lien avec les applications linéaires

On notera que, des bases de E et F étant fixées, si φ et ψ sont les applications linéaires définies par A et B respectivement, $A + B$ n'est autre que la matrice de $\varphi + \psi$ dans les bases données et λA est celle de $\lambda \cdot \varphi$ dans ces bases. D'autre part, on a le théorème suivant sur la composition des applications linéaires.

Theorem 3.3 *Soient E, F et G trois espaces vectoriels de dimension n, p et q. Si $\varphi : E \to F$ de matrice A et $\psi : F \to G$ de matrice B sont deux applications linéaires, la composée $\psi \circ \varphi$ est encore une application linéaire dont la matrice est BA.*

3.5 Noyau d'une application linéaire

Definition 3.4 *Soient E , F deux espaces vectoriels et φ une application linéaire de E dans F. Le* noyau *de φ, noté* $\ker\varphi$*, est l'ensemble des vecteurs de E d'image nulle:*

$$\ker\varphi = \{u \in E;\ \varphi(u) = 0\}.$$

Il est facile de voir que $\ker\varphi$ est un sous-espace vectoriel de E.

3.6 Rang d'une application linéaire

3.6.1 Définition

L'ensemble des images par φ des éléments de E est un sous-espace de F que l'on note $Im(\varphi)$. La dimension de l'image de φ est le rang de φ noté $Rg(\varphi)$.

3.6.2 Détermination pratique du rang d'une matrice

On considère la matrice

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 5 & 8 & 3 \end{pmatrix}$$

de l'application linéaire $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$. On peut utiliser la méthode du pivot de Gauss pour déterminer $Rg(M)$.

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 5 & 8 & 3 \end{pmatrix} \quad \begin{array}{c} \sim \\ L2-2L1 \\ L3-5L1 \end{array} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \end{pmatrix} \quad \begin{array}{c} \sim \\ \\ L3-2L2 \end{array} \begin{pmatrix} \boxed{1} & 2 & 1 \\ 0 & \boxed{\text{-1}} & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Il ne reste que deux pivots non nuls, on en déduit que $Rg(M) = 2$.

3.7 Injectivité, surjectivité, bijectivité

Rappelons qu'une application entre *ensembles* $\varphi : E \to F$ est

- *injective* si quels que soient x et x' dans E, $x \neq x'$ entraîne $\varphi(x) \neq \varphi(x')$.
- *surjective* si l'image de φ, *i.e.* l'ensemble $\{\varphi(x);\ x \in E\}$ est égal à F.
- *bijective* si elle est à la fois injective et surjective.

Dans le cas des applications linéaires (entre espaces vectoriels) on a les résultats suivants.

Proposition 3.5
- *L'application linéaire $\varphi : E \to F$ est injective si et seulement si* $\ker\varphi = \{0\}$.
- *L'application linéaire $\varphi : E \to F$ est surjective si $Rg(\varphi) = dimF$.*

Une application linéaire bijective est appelée **isomorphisme**.

3.8 Théorème du rang

Theorem 3.6 [**Théorème du rang**] *Soient E et F deux espaces vectoriels de dimension finie. Si $\varphi : E \to F$ est une application linéaire, alors*

$$dim\ker\varphi + dim\ Im\varphi = \dim E.$$

3.9 Endomorphisme

Une application linéaire d'un espace vectoriel E dans lui-même est appelée endomorphisme de E. Sa matrice est carrée d'ordre $dimE$.

En appliquant directement le théorème du rang, on s'aperçoit que si φ est un endomorphisme de E dans E, les propriétés suivantes sont équivalentes:

(i) φ est injective

(ii) φ est surjective

(iii) φ est bijective.

Un endomorphisme bijectif est appelé **automorphisme**.

4 Déterminant d'une matrice d'endomorphisme

On propose de rappeler les principales méthodes pour le calcul du déterminant d'une matrice carrée. On insiste sur le fait qu'on ne peut calculer un déterminant que sur une matrice carrée. Le déterminant est un nombre calculé d'après les coefficients de la matrice avec des propriétés particulières. Par exemple, nous savons déjà que s'il est non nul, la matrice est inversible. Des applications bien plus fondamentales sont à venir dans les chapitres suivants.

4.1 Cas des matrices 2×2

On considère la matrice

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Son déterminant est donné par la règle suivante :

$$det(M) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb.$$

Visuellement, on a compté positivement les produits sur la diagonale descendante, et négativement les produits sur la diagonale montante :

$$det(M) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \underline{ad} - \underline{cb}.$$

4.2 Cas des matrices 3×3

On considère la matrice

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Son déterminant est donné par la règle dite de Sarrus :

$$det(M) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + dhc - gec - ahf - idb.$$

Visuellement, on a compté positivement les produits sur les diagonales descendantes, et négativement les produits sur les diagonales montantes :

$$det(M) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \underline{aei} + \underline{bfg} + \underline{dhc} - \underline{gec} - \underline{ahf} - \underline{idb}.$$

Malheureusement, ces règles de produits sur les diagonales ne se généralisent pas au delà de la dimension 3.

4.3 Cas des matrices triangulaires

Si M est une matrice triangulaire, $\det M$ est le produit des éléments diagonaux :

$$\begin{vmatrix} \lambda_1 & \dots & \star \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{vmatrix} = \lambda_1 \cdots \lambda_n.$$

4.4 Règles générales sur le déterminant d'une matrice

On a les règles suivantes avec le déterminant d'une matrice.

Proposition 4.1 *(i)* $\det I = 1$

(ii) Considérons la matrice M écrite sous forme de colonnes : $M = (V_1, \cdots, V_n)$. Le déterminant est linéaire par rapport aux colonnes

$$\det(V_1, \dots, \lambda V_i + \mu W_i, \dots, V_n) = \lambda \det(V_1, \dots, V_i, \dots, V_n) + \mu \det(V_1, \dots, W_i, \dots, V_n)$$

On en déduit que si M est une matrice carrée s'ordre n, alors $det(\lambda M) = \lambda^n det(M)$, avec $\lambda \in \mathbb{R}$.

(iii) $\det$ *est alternée i.e. change de signe lorsque l'on permute deux colonnes :*

$$\det(V_1, \dots, V_i, \dots, V_j, \dots, V_n) = -\det(V_1, \dots, V_j, \dots, V_i, \dots, V_n).$$

(iv) *On peut montrer que* $\det M = \det{}^t M$ *(où* ${}^tM$ *est la transposée de* M*) : ce que l'on vient de dire pour les colonnes est donc aussi vrai pour les lignes.*

(v) *Si* A *et* B *sont deux matrices carrées* $n \times n$*, alors* $\det(AB) = \det(A)\det(B)$.

(vi) *Il résulte de ce théorème et de la propriété (i) ci-dessus que si* A *est inversible,* $\det A \neq 0$. *Réciproquement on peut montrer que si* $\det A \neq 0$*, alors* A *est inversible.*

(vii) *On ne change pas la valeur d'un déterminant en ajoutant à une colonne (resp. une ligne)* V_i *une combinaison linéaire des autres colonnes (resp. lignes).*

4.5 Méthode du pivot de Gauss

De cette dernière propriété, on peut lancer une méthode du Pivot de Gauss pour se ramener à une matrice triangulaire supérieure et en déduire le déterminant par le produit des termes diagonaux.

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 5 & 8 & 3 \end{vmatrix} \quad \begin{matrix} = \\ L2-2L1 \\ L3-5L1 \end{matrix} \begin{vmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \end{vmatrix} \quad \begin{matrix} = \\ \\ L3-2L2 \end{matrix} \begin{vmatrix} \boxed{1} & 2 & 1 \\ 0 & \boxed{\text{-1}} & -1 \\ 0 & 0 & 0 \end{vmatrix} = 1 \times (-1) \times 0 = 0.$$

Au passage, on remarquera que la matrice n'est pas inversible.

4.6 Développement d'un déterminant

Soit $M = (m_{i,j})$ une matrice d'ordre n. Si l'on se fixe une ligne i, on a

$$\det M = \sum_{j=1}^{j=n} (-1)^{i+j} m_{i,j} D_{i,j} \tag{4.1}$$

où $D_{i,j}$ est le déterminant de la matrice obtenue en supprimant de la matrice M la i-ème ligne et la j-ème colonne. On dit que l'on a développé $\det M$ suivant la i-ème ligne. On a de même, pour tout j,

$$\det M = \sum_{i=1}^{i=n} (-1)^{i+j} m_{i,j} D_{i,j} \tag{4.2}$$

c'est le développement de $\det M$ suivant la j-ème colonne.

Sur une matrice 3×3, développons sur la première ligne :

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} \text{-} a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

Reprenons cet exemple de manière illustrée avec une matrice de signes

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}.$$

Visuellement, nous avons

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & \cancel{a_2} & \cancel{a_3} \\ \cancel{b_1} & b_2 & b_3 \\ \cancel{c_1} & c_2 & c_3 \end{vmatrix} \quad \begin{vmatrix} \cancel{a_1} & a_2 & \cancel{a_3} \\ \boxed{b_1} & \cancel{b_2} & \boxed{b_3} \\ \boxed{c_1} & \cancel{c_2} & \boxed{c_3} \end{vmatrix} \quad \begin{vmatrix} \cancel{a_1} & \cancel{a_2} & a_3 \\ b_1 & b_2 & \cancel{b_3} \\ c_1 & c_2 & \cancel{c_3} \end{vmatrix}$$

Il résulte de ce qui précède que l'on a tout intérêt à avoir, sur une même ligne ou une même colonne, le plus grand nombre possible de coefficients nuls.

4.7 Matrice triangulaires par blocs

Les résultats de la sous-section 4.3 se généralisent aux matrices triangulaires par blocs.

$$\left|\begin{array}{c|c} \text{A} & \text{B} \\ \hline 0 & \text{D} \end{array}\right| = det(A)detD).$$

On peut, par exemple calculer le déterminant suivant très rapidement :

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 10 & 11 \end{vmatrix} = \left|\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & 9 & 10 \\ & & 11 & 12 \end{array}\right| = \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \times \begin{vmatrix} 9 & 10 \\ 11 & 12 \end{vmatrix}$$
$$= (1 \times 6 - 5 \times 2) \times (9 \times 12 - 11 \times 10) = (-4) \times (-2) = 8.$$

Attention, il est nécessaire que le bloc sous ou sur diagonale (B) soit nul, sinon le résultat ne tient plus.

Part II

Réduction des endomorphismes

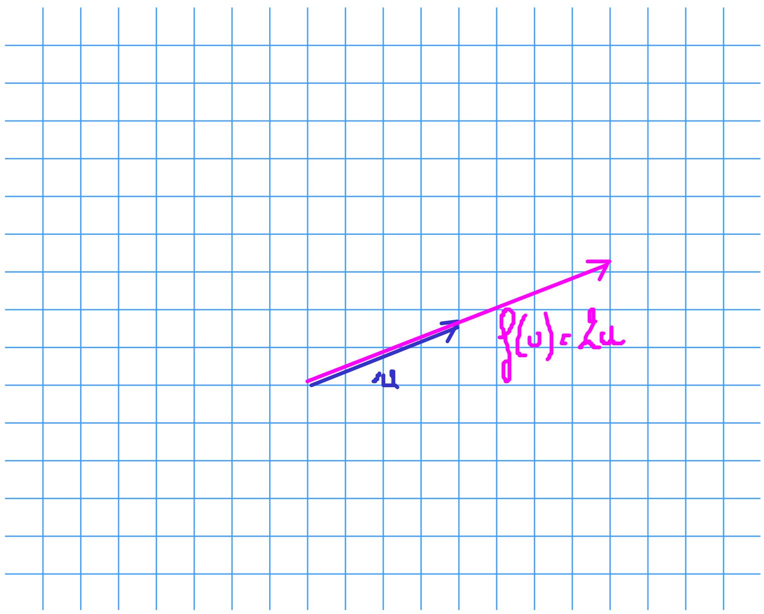
5 Introduction

Considérons, pour faire simple, un endomorphisme $f : \mathbb{R}^2 \to \mathbb{R}^2$, de matrice M.

Supposons, par pure fiction, qu'il existe un coefficient λ, par exemple $\lambda = 2$, et un vecteur u, non nul, tel que

$$f(u) = \lambda u \quad \Longleftrightarrow \quad f(u) = 2u.$$

Pour ce vecteur u, l'endomorphisme u se comporte comme une dilatation de rapport 2 :

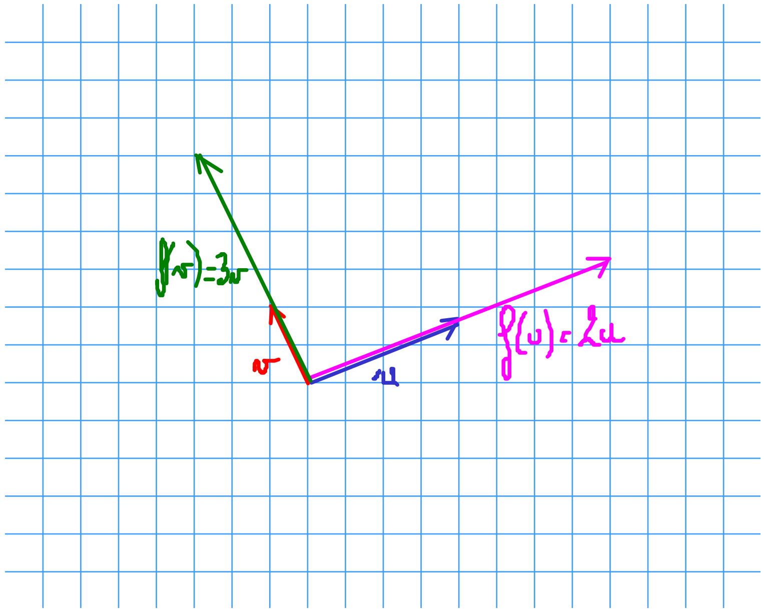


On aurait alors

$$Mu = 2u \quad \Longleftrightarrow \quad Mu - 2u = 0 \quad \Longleftrightarrow \quad (M - 2I)u = 0 \quad \Longleftrightarrow \quad u \in Ker(M - 2I).$$

Ce qu'on pouvait prévoir par la linéarité de l'endomorphisme f, est que cet effet dilatation sera vérifé pour tous les vecteurs du sous espace vectoriel $Ker(M - 2I)$. On nommera ce vecteur u *vecteur propre* associé à la *valeur propre* 2. D'autre part, le sous espace vectoriel $E_2 = Ker(M - 2I)$ sera nommé *sous espace propre* ou *sous espace stable* associé à la valeur propre 2.

Supposons maintenant qu'il existe une seconde valeur propre $\lambda = 3$ pour qui f agit comme une dilatation de rapport 3 sur le vecteur propre v :



On a alors $v \in Ker(M-3I)$ et $f(v) = 3v$. Il semble assez clair que u et v ne peuvent être colinéaires. En effet, les sous espaces propres étant des sous espaces vectoriels, si u et v sont colinéaires, u et v sont dans le même sous espace propre et f se comporterait comme une dilatation de même rapport pour u et pour v, ce qui est absurde. La famille (u, v) forme donc une base de $\mathbb{R}^2$. Écrivons la matrice D de f dans cette base (u, v) de $\mathbb{R}^2$. On rappelle que D porte en colonne l'image de u dans la base (u, v), c'est-à-dire le vecteur $(2, 0)$, suivi de l'image de v dans la base (u, v), c'est-à-dire le vecteur $(0, 3)$. Ainsi, f a pour matrice, dans la base (u, v), la matrice diagonale

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

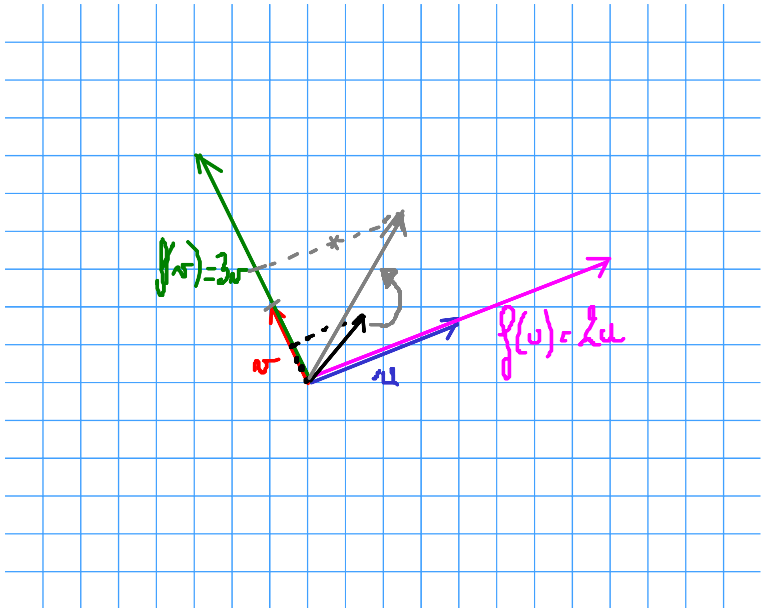
En donnant les matrices de changement de base P et P^{-1}, on trouvera alors la décomposition de la matrice M suivante

$$M = PDP^{-1}.$$

On a diagonalisé notre matrice M, et, par son caractère linéaire, l'endomorphisme f transformera toute combinaison linéaire de vecteurs u et v : $\lambda u + \mu v$ en combinaison de vecteurs $2u$ et $3v$:

$$f(\lambda u + \mu v) = 2\lambda u + 3\mu v.$$

Visuellement :

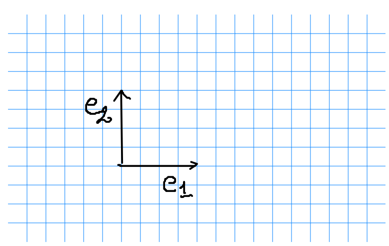


Trois questions se posent :

- Comment trouver les valeurs propres λ ?
- Existent-elles toujours ?
- Peut-on toujours trouver une base de vecteur propre pour diagonaliser la matrice ?

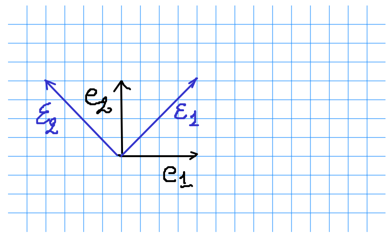
6 Rappel : changement de base

Nous allons rappeler la façon dont on peut exprimer la matrice d'un endomorphisme dans une base différente de la base canonique. Nous partons d'un exemple dans $\mathbb{R}^2$. Nous nommons $e = (e_1, e_2)$ cette base canonique.



Nous souhaitons changer de base vers la base $\varepsilon = (\varepsilon_1, \varepsilon_2)$ avec

$$\begin{aligned} \varepsilon_1 &= (1,1), \\ \varepsilon_2 &= (1,-1). \end{aligned}$$



Soit le vecteur $v = (a, b)_e$ exprimé dans la base e. On pose $v = (\alpha, \beta)_\varepsilon$ ses coordonnée dans la base ε. On a

$$ae_1 + be_2 = \alpha\varepsilon_1 + \beta\varepsilon_2.$$

Matriciellement,

$$a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} = \alpha\begin{pmatrix}1\\1\end{pmatrix} + \beta\begin{pmatrix}1\\-1\end{pmatrix}.$$

Autrement dit,

$$\begin{cases} a &= \alpha + \beta \\ b &= \alpha - \beta \end{cases}.$$

Il vient alors la relation matricielle suivante,

$$\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}\alpha\\\beta\end{pmatrix}$$

En notant V_e et V_ε les coordonnés du vecteur v dans les bases e et ε,

$$V_e = \begin{pmatrix}a\\b\end{pmatrix}, \qquad V_\varepsilon = \begin{pmatrix}\alpha\\\beta\end{pmatrix},$$

on a

$$V_e = PV_\varepsilon, \qquad P = \begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}.$$

Ainsi, de coordonnées V_ε, on obtient les coordonnées V_e ne multipliant les coordonnés V_ε par la matrice P. La matrice P est appelée matrice de passage de e dans ε. Il est important de voir qu'elle porte les vecteurs de ε en colonne exprimés dans le base e. On aurait aussi

$$P^{-1}V_e = V_\varepsilon.$$

On voit alors que la matrice P^{-1} est la matrice de passage de ε à e.

En toute généralité, on pose la définition suivante.

Definition 6.1 *Si $B = \{e_1, \ldots, e_n\}$ et $B' = \{e'_1, \ldots, e'_n\}$ sont deux bases de E on appelle matrice de passage de la base B à la base B' la matrice P ayant pour vecteurs colonnes les vecteurs e'_i écrits dans la base B.*

Nous venons d'expliquer la proposition suivante.

Proposition 6.2 *Si $B = \{e_1, \ldots, e_n\}$ et $B' = \{e'_1, \ldots, e'_n\}$ sont deux bases de E et P la matrice de passage de la base B à la base B'.*
Si $u = \sum_1^n x_i e_i = \sum_1^n x'_i e'_i$ on a

$$X = PX' \quad \textit{soit encore} \quad X' = P^{-1}X,$$

où

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Il vient alors la relations suivante pour les matrices.

Proposition 6.3 *Si $B = \{e_1, \ldots, e_n\}$ et $B' = \{e'_1, \ldots, e'_n\}$ sont deux bases de E et P la matrice de passage de la base B à la base B'.*
Soit $u = \sum_1^n x_i e_i = \sum_1^n x'_i e'_i$. On pose alors

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Soit maintenant M la matrice de f dans la base B et M' celle de f dans la base B'. Notant Y la matrice colonne de $f(u)$ dans la base B et Y' celle de ce vecteur dans la base B'. On a :

- *$Y = MX$ et $Y' = M'X'$,*
- *$PM'P^{-1}X = MX$ pour tout X,*
- *$M' = P^{-1}MP$, ou $M = PM'P^{-1}$.*

7 Diagonalisation des matrices

Nous allons filer dans la suite du cours l'exemple de l'endomorphisme de $\mathbb{R}^3$ de matrice

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

7.1 Polynôme caractéristique

Definition 7.1 *Le polynôme caractéristique est le polynôme unitaire (dont le coefficient du terme de plus haut degré est 1) $P_M(X) = det(XI - M)$.*

Dans notre exemple, le polynôme caractéristique de A est donné par

$$P_M(X) = \det(XI - A) = \begin{vmatrix} X-1 & 1 & 0 \\ 1 & X-2 & 1 \\ 0 & 1 & X-1 \end{vmatrix}$$

En utilisant directement la règle de Sarrus, on obtient

$$P_M(X) = (X-1)(X-2)(X-1) - (X-1) - (X-1).$$

En le factorisant par $(X-1)$,

$$P_M(X) = (X-1)[(X-2)(X-1) - 2] = (X-1)(X^2 - 3X) = (X-1)X(X-3).$$

7.2 Valeurs propres

Definition 7.2

- *Les valeurs propres (***VaP***) de M sont les racines du polynôme caractéristique $P_M(X)$.*
- *On nomme spectre de M, noté $Sp(M)$ l'ensemble des valeurs propres de M.*

Dans l'exemple, avec $P_M(X) = (X-1)X(X-3)$, les valeurs propres sont 0, 1 et 3 :

$$Sp(M) = \{0, 1, 3\}.$$

7.3 Sous espace propre

Definition 7.3 *Considérons une valeur propre $\lambda \in Sp(M)$. Le sous espace propre (***SeP***) E_λ associé à la valeur propre λ est défini par*

$$E_\lambda = Ker(M - \lambda I).$$

Dans notre exemple, les sous espaces propres sont :

- $E_0 = Ker(M)$,
- $E_1 = Ker(M - I)$,
- $E_3 = Ker(M - 3I)$.

7.4 Vecteur propre

Definition 7.4 *Considérons une valeur propre $\lambda \in Sp(M)$. Un vecteur propre (***VeP***) u_λ associé à la valeur propre λ est un vecteur de E_λ.*

Note : En pratique, on cherchera une base de vecteurs propres pour chaque sous espace propre.

Dans notre exemple...

- Déterminons $E_0 = Ker(M)$. On cherche donc le vecteur $u_0 = (x, y, z)$ tel que

$$
\begin{aligned}
Mu_0 = 0 \qquad &\Longleftrightarrow \qquad \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \\
&\Longleftrightarrow \qquad \begin{cases} x & -y & & = 0 \\ -x & +2y & -z & = 0 \\ & -y & +z & = 0 \end{cases} \\
&\Longleftrightarrow \qquad \begin{cases} x & = y \\ y & = z \end{cases}.
\end{aligned}
$$

  E_0 est donc le sous espace engendré par $u_0 = (1, 1, 1)$.

- Déterminons $E_1 = Ker(M - I)$. On cherche donc le vecteur $u_1 = (x, y, z)$ tel que

$$
\begin{aligned}
(M - I)u_1 = 0 \qquad &\Longleftrightarrow \qquad \begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \\
&\Longleftrightarrow \qquad \begin{cases} & -y & & = 0 \\ -x & +y & -z & = 0 \\ & -y & & = 0 \end{cases} \\
&\Longleftrightarrow \qquad \begin{cases} y & = 0 \\ x & = -z \end{cases}.
\end{aligned}
$$

  E_1 est donc le sous espace engendré par $u_1 = (1, 0, -1)$.

- On trouve de même que $E_3 = Ker(M - 3I)$ est engendré par le vecteur $u_3 = (1, -2, 1)$.

7.5 Base de vecteurs propres et matrice diagonale

On a vu dans l'introduction que nécessairement, les sous espaces propres sont en somme directe, ils ne peuvent avoir de vecteurs communs non nuls. Nous le vérifions dans le résultat suivant.

Proposition 7.5 *Soient* $\{u_1, \dots, u_k\}$ *des vecteurs associés à des valeurs propres* $\lambda_1, \dots, \lambda_k$ distinctes. $\{u_1, \dots, u_k\}$ *est un système libre.*

Proof. Si u_1 et u_2 sont associés à des valeurs propres distinctes, et si $\alpha_1 u_1 + \alpha_2 u_2 = 0$, alors appliquant f on obtient

$$f(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 = 0;$$

comme λ_1 et λ_2 ne sont pas simultanément nuls, on peut supposer $\lambda_1 \neq 0$. On obtient donc, en multipliant la première équation par λ_1 et en soustrayant la seconde, $\alpha_2(\lambda_2 - \lambda_1)u_2 = 0$ ce qui entraîne (puisque $\lambda_2 - \lambda_1 \neq 0$) $\alpha_2 = 0$. On a donc aussi $\alpha_1 = 0$ et la proposition est démontrée pour $k = 2$.

Supposons le résultat vrai pour $k < n$. Si $u = \alpha_1 u_1 + \cdots + \alpha_n u_n = 0$, on a $f(u) = 0$ de sorte que

$$\alpha_1 \lambda_1 u_1 + \cdots + \alpha_n \lambda_n u_n = 0.$$

L'un au moins des λ_i étant non nul, on peut supposer que $\lambda_1 \neq 0$. Éliminant u_1 on obtient $\sum_2^n \alpha_j(\lambda_j - \lambda_1)u_j = 0$. Il suffit alors d'appliquer l'hypothèse de récurrence pour conclure. □

On peut alors énoncer le théorème suivant.

Theorem 7.6 *Considérons l'endormorphisme f de $\mathbb{R}^n$ de matrice M. Avec les notations en cours, les assertions suivantes sont équivalentes :*

- *On peut trouver une base de vecteurs propres $(e_1, \cdots e_n)$.*
- $\mathbb{R}^n = \bigoplus_i E_{\lambda_i}$.
- *Dans la base de vecteurs propres $(e_1, \cdots e_n)$, la matrice de f est diagonale.*

Cela signifie qu'on peut écrire M sous la forme

$$PDP^{-1}$$

avec D une matrice diagonale portant les valeurs propres sur sa diagonale.

La matrice P est la matrice de passage de la base canonique dans la base de vecteurs propres, elle porte les vecteurs propres en colonne dans l'ordre des valeurs propres sur la diagonale de D. On obtient, si nécessaire la matrice P^{-1} en inversant P.

Dans notre exemple, la matrice de f dans la base $\{u_1, u_2, u_3\}$ n'est autre que $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Si P est la matrice de passage de la base initiale à $\{u_1, u_2, u_3\}$, on a

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

et donc $M = PDP^{-1}$. On peut alors rappeler le calcul de l'inverse par la méthode du pivot de Gauss.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \quad \begin{array}{c} \sim \\ L2 - L1 \\ L3 - L1 \end{array} \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -3 & -1 & 1 & 0 \\ 0 & -2 & 0 & -1 & 0 & 1 \end{array}\right)$$

$$\begin{array}{c} \sim \\ L2 \leftrightarrow L3/(-2) \end{array} \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & -1 & -3 & -1 & 1 & 0 \end{array}\right)$$

$$\begin{array}{c} \sim \\ \\ -(L3 + L2)3 \end{array} \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & 0 & 1 & 1/6 & -1/3 & 1/6 \end{array}\right)$$

$$\begin{array}{c} L1 - L2 - L3 \\ \sim \end{array} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & 0 & 1 & 1/6 & -1/3 & 1/6 \end{array}\right)$$

Ainsi,

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

7.6 Diagonalisation en pratique

On considère l'endomorphisme f de matrice M. On souhaite diagonaliser f.

(i) On calcule le polynôme caractéristique $Det(XI - M)$.

(ii) On en déduit les valeurs propres qui sont ses racines.

(iii) On cherche un vecteur propre associé à chaque valeur propre λ. C'est à dire, pour chaque valeur propre λ, on trouve un vecteur non nul de $Ker(M - \lambda I)$.

(iv) Si on obtient une base de vecteurs propres, on peut diagonaliser (sinon les ennuis commencent).

(v) On écrit M sous la forme PDP^{-1} avec D une matrice diagonale portant les valeurs propres sur sa diagonale. La matrice P porte les vecteurs propres en colonne dans l'ordre des valeurs propres sur la diagonale de D. On obtient, si nécessaire la matrice P^{-1} en inversant P.

8 Applications

8.1 Puissance d'une matrice

On a vu que l'endomorphisme que l'endomorphisme de notre exemple de matrice

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

pouvait s'écrire sous la forme Ainsi,

$$M = PDP^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ 1/6 & -1/3 & 1/6 \end{pmatrix}.$$

On peut en déduire que

$$M^N = (PDP^{-1})^N = \underbrace{PDP^{-1} \times \cdots \times PDP^{-1}}_{n\ fois} = PD^N P^{-1}. \tag{8.3}$$

pour tout $N \in \mathbb{N}$. Il est assez immédiat que

$$D^N = \begin{pmatrix} 0^N & 0 & 0 \\ 0 & 1^N & 0 \\ 0 & 0 & 3^N \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^N \end{pmatrix}.$$

(On pourrait le montrer par récurrence...) Si on le souhaite, on peut aisément expliciter les calculs :

$$\begin{aligned}
M^N &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^N \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ 1/6 & -1/3 & 1/6 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 3^N \\ 0 & 0 & -2 \times 3^N \\ 0 & -1 & 3^N \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ 1/6 & -1/3 & 1/6 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1+3^{N-1}}{2} & -3^{N-1} & \frac{-1+3^{N-1}}{2} \\ 3^{N-1} & 2 \times 3^{N-1} & -3^{N-1} \\ \frac{-1+3^{N-1}}{2} & -3^{N-1} & \frac{1+3^{N-1}}{2} \end{pmatrix}.
\end{aligned}$$

Bien entendu, on évitera d'expliciter les calculs s'il n'y en a pas besoin.

8.2 Inverse d'une matrice

On rappelle que l'inverse d'un produit de matrice vérifie

$$(AB)^{-1} = B^{-1}A^{-1}.$$

On en déduit qu'alors

$$M^{-1} = (PDP^{-1})^{-1} = P(PD)^{-1} = PD^{-1}P^{-1}.$$

Il devient alors claire que le formule (8.3) s'étend à $N \in \mathbb{Z}$.

8.3 Exponentielle d'une matrice

Pour définir l'exponentielle d'une matrice, on utilisera le développement en série entière de celle-ci,

$$e^A = \sum_{i \geq 0} \frac{A^i}{i!}.$$

Puisque $M^i = PD^iP^{-1}$, on en déduit que

$$e^M = \sum_{i \geq 0} \frac{M^i}{i!} = \sum_{i \geq 0} \frac{PD^iP^{-1}}{i!} = P \left(\sum_{i \geq 0} \frac{D^i}{i!} \right) P^{-1}.$$

On a

$$\sum_{i \geq 0} \frac{D^i}{i!} = \sum_{i \geq 0} \frac{1}{i!} \begin{pmatrix} 0^i & 0 & 0 \\ 0 & 1^i & 0 \\ 0 & 0 & 3^i \end{pmatrix} = \begin{pmatrix} e^0 & 0 & 0 \\ 0 & e^1 & 0 \\ 0 & 0 & e^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e^3 \end{pmatrix} = e^D.$$

Il vient que

$$e^M = Pe^DP^{-1}.$$

8.4 Application aux systèmes différentiels

8.4.1 Le cas scalaire

Soit $\lambda \in \mathbb{R}$ un nombre fixé. L'équation différentielle $x' = \lambda x$ est une équation dont l'inconnue est une fonction. Une solution de cette équation est une fonction dérivable $x : t \mapsto x(t)$ définie sur $\mathbb{R}$ et telle que, pour tout $t \in \mathbb{R}$, on ait $x'(t) = \lambda x(t)$. Il ne faut pas se laisser abuser par l'écriture (guidée par l'usage) : ici x est une fonction et non pas un nombre.

Si $x : \mathbb{R} \to \mathbb{R}$ est une fonction dérivable quelconque, alors la dérivée de la fonction $t \mapsto e^{-\lambda t}x(t)$ est

$$t \mapsto (x'(t) - \lambda x(t))e^{-\lambda t}.$$

Par conséquent, la fonction x est solution de l'équation différentielle $x' = \lambda x$ si, et seulement si, la fonction $t \mapsto e^{-\lambda t}x(t)$ a une dérivée nulle donc est constante. Si nous notons c cette constante, il vient

$$x(t) = ce^{\lambda t}, \qquad \forall t \in \mathbb{R}.$$

Nous voyons alors que $x(0) = c$. Par conséquent, nous retenons le résultat suivant.

Theorem 8.1 *Soit λ un nombre réel fixé. Quel que soit le nombre réel c, il existe une unique fonction dérivable $x : \mathbb{R} \to \mathbb{R}$ solution de l'équation différentielle $x' = \lambda x$ et vérifiant la condition initiale $x(0) = c$. Cette fonction est donnée par la formule*

$$x(t) = ce^{\lambda t}, \qquad \forall t \in \mathbb{R}.$$

8.4.2 Le cas multidimensionnel

On souhaite résoudre le système d'équations différentielles

$$\begin{cases} x_1'(t) &= & x_1(t) - x_2(t) \\ x_2'(t) &= & -x_1(t) + 2x_2(t) - x_3(t) \\ x_3'(t) &= & -x_2(t) + x_3(t) \end{cases}, \qquad \begin{cases} x_1(0) &= & 1 \\ x_2(0) &= & 2 \\ x_3'(t) &= & -1 \end{cases} \tag{8.4}$$

(i) En posant

$$X'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix}, \qquad X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix},$$

On peut écrire le système sous la forme $X' = MX$, avec

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

(ii) On a diagonalisé M. En effet, $M = PDP^{-1}$, où D est la matrice diagonale

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

et P une matrice de passage rappelée ultérieurement.

(iii) On va effectuer le changement de variable $Y = P^{-1}X$. Il est à noter qu'il n'est pas nécessaire d'expliciter P^{-1}.

Partons du système
$$X' = MX = PDP^{-1}X.$$
On a naturellement, avec le changement de variable $Y = P^{-1}X$,
$$X' = PDY.$$
En multipliant à gauche par P^{-1}
$$P^{-1}X' = P^{-1}PDY = DY.$$
Il reste à observer que $Y' = (P^{-1}X)' = P^{-1}X'$. En effet,
$$\begin{aligned}(P^{-1}X(t))' = \left(\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}\right)' &= \left(\begin{pmatrix} ax_1(t) + bx_2(t) + cx_3(t) \\ dx_1(t) + ex_2(t) + fx_3(t) \\ gx_1(t) + hx_2(t) + ix_3(t) \end{pmatrix}\right)' \\ &= \begin{pmatrix} ax_1'(t) + bx_2'(t) + cx_3'(t) \\ dx_1'(t) + ex_2'(t) + fx_3'(t) \\ gx_1'(t) + hx_2'(t) + ix_3'(t) \end{pmatrix} = P^{-1}X'(t).\end{aligned}$$
En fait, ce résultat est assez évident dès qu'on se rappelle que la dérivée d'une combinaison linéaire est la combinaison linéaire des dérivées.

(iv) On sait résoudre le système $Y' = DY$. En effet, on l'écrit
$$\begin{cases} y_1'(t) &= 0 \\ y_2'(t) &= y_2(t) \\ y_3'(t) &= 3y_3(t) \end{cases}.$$
Coordonnée par coordonnée, en appliquant le Théorème 8.1,
$$\begin{cases} y_1(t) &= C_1 \\ y_2(t) &= C_2e^t \\ y_3(t) &= C_3e^{3t} \end{cases}.$$

(v) On peut donc en déduire la solution X du système initial en se rappelant que si $Y = P^{-1}X$ alors $X = PY$. On n'aura pas besoin d'inverser P (même si c'est déjà fait). On a donc
$$X(t) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} Y(t) = \begin{pmatrix} C_1 + C_2e^t + C_3e^{3t} \\ C_1 - 2C_3e^{3t} \\ C_1 - C_2e^t + C_3e^{3t} \end{pmatrix}.$$
On utilise les conditions initiales (8.4) pour en déduire les constantes C_1, C_2, C_3.
$$\begin{cases} C_1 + C_2 + C_3 &= 3 \\ C_1 - 2C_3 &= -1 \\ C_1 - C_2 + C_3 &= 1 \end{cases},$$
De solution $C_1 = C_2 = C_3 = 1$.

La solution du système (8.4) est donc

$$X(t) = \begin{pmatrix} 1 + e^t + e^{3t} \\ 1 - 2e^{3t} \\ 1 - 2e^t + e^{3t} \end{pmatrix}.$$

8.5 Applications aux suites récurrentes

On souhaite déterminer les expressions de u_p, v_p, et w_p définies par le système

$$\begin{cases} u_{p+1} &= u_p - v_p \\ v_{p+1} &= -u_p + 2v_p - w_p \\ w_{p+1} &= -v_p + w_p \end{cases},$$

Avec des premiers termes

$$\begin{cases} u_0 \\ v_0 \\ w_0 \end{cases}.$$

(i) On a $U_{p+1} = MU_p$ avec

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

(ii) La formule explicite de U_N est donc

$$U_N = M^N U_0, \qquad N \in \mathbb{N},$$

(par analogie directe avec les suites géométriques classiques).

(iii) Nous avons calculé explicitement la matrice M^N au paragraphe (8.1),

$$M^N = \begin{pmatrix} \frac{1+3^{N-1}}{2} & -3^{N-1} & \frac{-1+3^{N-1}}{2} \\ 3^{N-1} & 2 \times 3^{N-1} & -3^{N-1} \\ \frac{-1+3^{N-1}}{2} & -3^{N-1} & \frac{1+3^{N-1}}{2} \end{pmatrix}.$$

On a donc

$$U_N = \begin{pmatrix} \frac{1+3^{N-1}}{2} & -3^{N-1} & \frac{-1+3^{N-1}}{2} \\ 3^{N-1} & 2 \times 3^{N-1} & -3^{N-1} \\ \frac{-1+3^{N-1}}{2} & -3^{N-1} & \frac{1+3^{N-1}}{2} \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} = \begin{pmatrix} \frac{1+3^{N-1}}{2} u_0 - 3^{N-1} v_0 + \frac{-1+3^{N-1}}{2} w_0 \\ 3^{N-1} u_0 + 2 \times 3^{N-1} v_0 - 3^{N-1} w_0 \\ \frac{-1+3^{N-1}}{2} u_0 - 3^{N-1} v_0 \frac{1+3^{N-1}}{2} w_0 \end{pmatrix}$$

9 Autres réductions

On a vu, dans le théorême 7.6 qu'une matrice est diagonalisable si on peut trouver une base de vecteurs propres. Il y a toujours au moins un vecteur propre par valeur propre et donc, si nous avons n valeurs propres distinctes dans $\mathbb{R}^n$, nous sommes sûrs de pouvoir diagonaliser la matrice. Il arrive qu'il y ait moins de valeurs propres que prévu, comme dans l'exemple suivant.

Considérons l'endomorphisme de $\mathbb{R}^3$ de matrice

$$M = \begin{pmatrix} 5 & -7 & 3 \\ 0 & 1 & 0 \\ -3 & 5 & -1 \end{pmatrix}.$$

On calcule son polynôme caractéristique.

$$\begin{array}{rcl}
P_M(X) = det(XI - M) & = & \begin{vmatrix} X-5 & 7 & -3 \\ 0 & X-1 & 0 \\ 3 & -5 & X+1 \end{vmatrix} \\
C1 + C2 + C3 & = & \begin{vmatrix} X-1 & 7 & -3 \\ X-1 & X-1 & 0 \\ X-1 & -5 & X+1 \end{vmatrix} \\
 & = & (X-1) \begin{vmatrix} 1 & 7 & -3 \\ 1 & X-1 & 0 \\ 1 & -5 & X+1 \end{vmatrix} \\
Sarrus\ ! & = & (X-1)\big((X-1)(X+1) + 15 + 3(X-1) - 7(X+1)\big) \\
 & = & (X-1)(X^2 - 4X + 4) = (X-1)(X-2)^2.
\end{array}$$

Il n'y a que deux valeurs propres : 1 et 2. La valeur propre 2 a son terme muni d'une puissance 2, qu'on nommera multiplicité algébrique.

Definition 9.1 *La multiplicité algébrique d'une valeur propre est la puissance de son terme dans le polynôme caractéristique sous sa forme factorisée.*

On a le résultat suivant qui nous permet que connaître la dimension maximale d'un sous espace propre.

Proposition 9.2 *Soit α_λ la multiplicité de la valeur propre λ. On a*

$$1 \leq \dim E_\lambda \leq \alpha_\lambda,$$

où E_λ est le sous espace propre associé à la valeur propre λ.

D'après le théorême 7.6, on trouvera donc une base de vecteurs propres, et la matrice sera diagonalisable, si $\dim E_\lambda = \alpha_\lambda$, pour toutes les valeurs propres.

- Ici, cherchons le vecteur propre associé à la valeur propre 1. Soit $u_1 = (x, y, z) \in E_1$. On a

  $$u_1 \in E_1 \qquad \Longleftrightarrow \qquad \left\{ \begin{array}{rcl} 4x - 7y + 3z & = & 0 \\ -3x + 5y - 2z & = & 0 \end{array} \right. .$$

  On pourra prendre $u_1 = (1, 1, 1)$.

- Cherchons le ou les vecteurs propres associés à 2. Soit $u_2 = (x, y, z) \in E_2$. On a

$$\begin{array}{rcl} u_1 \in E_1 & \iff & \left\{\begin{array}{rcl} 3x - 7y + 3z & = & 0 \\ -y & = & 0 \\ -3x + 5y - 3z & = & 0 \end{array}\right. \\ & \iff & \left\{\begin{array}{rcl} 3x + 3z & = & 0 \\ y & = & 0 \\ -3x - 3z & = & 0 \end{array}\right. \\ & \iff & \left\{\begin{array}{rcl} 3x + 3z & = & 0 \\ y & = & 0 \end{array}\right. . \end{array}$$

On trouve $u_2 = (1, 0, -1)$.

Avec deux équations cartésiennes, le sous espace propre E_2 est de dimension 1. Il n'y aura pas de base de vecteurs propres. La matrice n'est pas diagonalisable. Pour autant, complétons les vecteurs propres avec un vecteur linéairement indépendant arbitraire (mais simple) : $e_3 = (0, 0, 1)$. On observe que la famille (u_1, u_2, e_3) est une base. Dans cette base, notre endomorphisme porte les images des vecteurs de la base (dans cette base) et a pour matrice

$$T = \begin{pmatrix} 1 & 0 & a \\ 0 & 2 & b \\ 0 & 0 & c \end{pmatrix}.$$

Puisque la trace (somme des éléments diagonaux) est un invariant de similitude (ne change pas malgré le changement de base), on en déduit que $c = 2$. On peut donc trigonaliser la matrice ! On va donc écrire $M = PTP^{-1}$ et identifier nos coefficients manquants.

$$\begin{pmatrix} 5 & -7 & 3 \\ 0 & 1 & 0 \\ -3 & 5 & -1 \end{pmatrix} = PTP^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & a \\ 0 & 2 & b \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & -2 & 1 \end{pmatrix}.$$

(Nous laissons le calcul de P^{-1} aux soins du lecteur). Calculons...

$$\begin{pmatrix} 5 & -7 & 3 \\ 0 & 1 & 0 \\ -3 & 5 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & a+b \\ 1 & 0 & a \\ 1 & -2 & a-b+2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 2+a+b & -1-2a-2b & a+b \\ a & 1-2a & a \\ a-b & -1-2a+2b & a-b+2 \end{pmatrix}.$$

On observe alors que $a = 0$ et

$$\begin{pmatrix} 5 & -7 & 3 \\ 0 & 1 & 0 \\ -3 & 5 & -1 \end{pmatrix} \begin{pmatrix} 2+b & -1-2b & b \\ 0 & 1 & 0 \\ -b & -1+2b & -b+2 \end{pmatrix}.$$

On garde $b = 3$ et

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

On peut observer que $e_3 \in Ker(M-2I)^2$. En effet

$$(M-2I)^2 = \begin{pmatrix} 3 & -7 & 3 \\ 0 & -1 & 0 \\ -3 & 5 & -3 \end{pmatrix} \begin{pmatrix} 3 & -7 & 3 \\ 0 & -1 & 0 \\ -3 & 5 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

et $(M-2I)^2 e_3 = 0$.

En fait, pour trigonaliser une matrice, on pourra toujours chercher des vecteurs dans les puissances successives i de $Ker(M-\lambda)^i$, pour $i \leq \alpha_\lambda$, la multiplicité algébrique de la valeur propre λ.

10 Critères de diagonalisabilité

La lecture du paragraphe 9 est indispensable à la compréhension des prochains éléments sur la possibilité de diagonaliser une matrice.

10.1 Multiplicités

Pour l'instant, le Théorême 7.6 nous dit que nous pouvons diagonaliser une matrice si nous pouvons trouver une base de vecteurs propres. Posons la définition suivante.

Definition 10.1 *Soit λ une valeur propre. Notons E_λ son sous espace propre associé $E_\lambda = Ker(M-\lambda I)$. On appellera multiplicité géométrique la dimension de E_λ.*

Ainsi, avec l'apport du paragraphe 9, nous pouvons rajouter au théorême 7.6 l'assertion donnant légalité entre les multiplicités géométriques et algébriques α_λ.

Theorem 10.2 *Considérons l'endormorphisme f de $\mathbb{R}^n$ de matrice M. Avec les notations en cours, les assertions suivantes sont équivalentes :*

- *On peut trouver une base de vecteurs propres $(e_1, \cdots e_n)$.*
- $\mathbb{R}^n = \bigoplus_i E_{\lambda_i}$.
- $\alpha_\lambda = dim E_\lambda$, $\forall \lambda \in Sp(M)$.
- *Dans la base de vecteurs propres $(e_1, \cdots e_n)$, la matrice de f est diagonale.*

10.2 Étude du polynôme minimal

Il est à noter que le polynôme caractéristique d'une matrice M est annulateur de M, c'est à dire $P_M(M) = 0$. Nous l'avons défini pour qu'il soit unitaire (le coefficient devant le terme de plus grand degré vaut 1). Supposons que le polynôme caractéristique s'écrire

$$P_M(X) = \prod_i (X-\lambda_i)^{\alpha_{\lambda_i}}. \tag{10.5}$$

On dira qu'il est scindé puisqu'il n'a que des racines réelles. (Il est à noter qu'un polynôme est toujours scindé sur $\mathbb{C}$.)

Theorem 10.3 *Si le polynôme caractéristique est scindé, la matrice est trigonalisable.*

Bien entendu, ce n'est pas toujours le cas sur $\mathbb{R}$. Prenons la matrice

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

de polynôme caractériqtique $P_M(X) = X^2 + 1$. Cette matrice n'est pas trigonalisable dans $\mathbb{R}$.

Si le polynôme caractéristique est scindé, d'écriture (10.6), on peut trouver un polynôme minimal, au sens des multiplicité, annulateur de M :

$$P_m(X) = \prod_i (X - \lambda_i)^{\beta_{\lambda_i}}, \tag{10.6}$$

où tous les $\beta_{\lambda_i} \leq \alpha_{\lambda_i}$ sont minimaux.

Par exemple, la matrice

$$M = \begin{pmatrix} 3 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

a pour polynôme caractéristique $P_M(X) = (X-1)(X-2)^2$. Le polynôme minimal de M est alors soit $P_m(X) = (X-1)(X-2)$, soit $P_m(X) = (X-1)(X-2)^2$. Il suffit de voir si le premier est annulateur.

$$(M - I)(M - 2I) = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{pmatrix} = 0.$$

Donc M a pour polynôme minimal $P_m(X) = (X-1)(X-2)$. Il est à noté que ce polynôme est scindé et que toutes les multiplicités des racines sont 1. On dira qu'il est simplement scindé. On annonce le résultat suivant et son corollaire.

Theorem 10.4 *Si le polynôme minimal d'un endomorphisme est simplement scindé, sa matrice est diagonalisable.*

Corollary 10.5 *Si un endomorphisme de $\mathbb{R}^n$ admet n valeurs propres distinctes, il est diagonalisable.*

11 Autres décompositions

11.1 Décomposition LU et calcul de déterminant

La décomposition LU est une méthode de décomposition d'une matrice en une matrice triangulaire inférieure L (comme "Low", bas) et une matrice triangulaire supérieure U (comme "Up", haut). Cette décomposition est utilisée en analyse numérique pour résoudre des systèmes d'équations linéaires.

Definition 11.1 *Soit A une matrice inversible. La matrice A peut être décomposée ainsi :*

$$P^{-1}A = LU; \qquad A = PLU,$$

où P est une matrice de permutation (de même pour P^{-1}), L est une matrice triangulaire inférieure et U une matrice triangulaire supérieure. Parfois, on peut ramener la matrice de passage P à la matrice identité. Dans ce cas, la décomposition devient $A = LU$.

Note : Une matrice de permutation est une matrice carrée qui vérifie les propriétes suivantes :

- les coefficients sont 0 ou 1,
- il n'y a qu'un seul 1 par ligne,
- il n'y a qu'un seul 1 par colonne.

Par exemple, la matrice suivante est une matrice de permutation :

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

Dès lors que le déterminant du produit d'une matrice est le produit des déterminant des matrices, l'intérêt de cette décomposition pour le calcul du déterminant est transparent.

11.2 Matrices orthogonales et décomposition QR

Definition 11.2 *Une matrice orthogonale est une matrice dont la norme des vecteurs colonne (la racine de la somme des carrés des coefficients) vaut 1 et dont le produit scalaire des vecteurs colonne différents (somme des produits coordonnée par coordonné) vaut 0.*

Par exemple, on peut noter que la matrice

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

est orthogonale.

Proposition 11.3 *Pour toute matrice orthogonale Q, on a*

- $Q^tQ = I$ *(ou ${}^tQ$ est la transposée de Q),*
- $det(Q) = \pm 1$,
- $Q^{-1} = {}^tQ$,
- *si Q est une matrice de passage d'une base e à une base ε, la matrice de passage de ε à e est la matrice ${}^tQ$.*

Theorem 11.4 *Soit A une matrice. On peut écrire la décomposition QR de cette matrice A. Il s'agit de l'écriture sous la forme*

$$A = QR$$

où Q est une matrice orthogonale, et R une matrice triangulaire supérieure.

Proposition 11.5 *Si la matrice A est inversible, on peut demander à ce que les coefficients diagonaux de R soient strictement positif, dans la décomposition QR. Dans ce cas, cette décomposition QR est unique.*

11.3 Matrices symétriques

Definition 11.6 *La matrice M est dite symétrique si*

$$M = {}^tM.$$

Un exemple fondamental de matrice symétrique est la matrice de covariance. Pour un vecteur de variables statistiques $X = (X_1, X_2, \cdots X_d)$, la matrice de covariance porte en $i$ème position sur la diagonale les variances des X_i et sur la coordonnée i, j, la covariance de X_i, X_j notée $cov(X_i, X_j)$.

$$M = \big(cov(X_i, X_j)\big)_{i,j \leq d},$$

en rappelant que

$$cov(X_i, X_j) = \frac{1}{n}\sum_{k=1}^{n}(X_{ik}X_{jk} - \bar{X}_i\bar{X}_j), \qquad var(X_i) = cov(X_i, X_i).$$

On note que la covariance est symétrique et par conséquent, M aussi.

11.3.1 Diagonalisabilité

Theorem 11.7 *Toute matrice symétrique réelle est diagonalisable. On peut trouver une matrice Q orthogonale telle que*

$$M = QD^tQ.$$

11.3.2 Application à l'ACP

On souhaite étudier le vin vendu dans les caves de plusieurs producteurs indépendants à la suite de dégustations aux particuliers. Sept variables sont retenues pour l'étude :

- Prix : renseigne sur le prix de la bouteille,
- QT : quantité de vin servie dans le verre lors de la dégustation,
- Alc : il s'agit du degré d'alcool du vin,
- An : renseigne sur l'âge du vin,
- Coul : évalue la couleur du vin,

- EQ : évalue l'équilibre du vin,
- Gout : évalue le goût du vin.

On suppose les variables standardisées (on les a centrées, soustrait la moyenne, et réduites, divisées par leur écart-type), afin d'éviter un rapport de force dû à une variable numériquement plus forte. On va effectuer une analyse en composantes principales pour savoir si l'ensemble des statistiques peuvent se résumer à un nombre inférieurs de facteurs, *i.e.* d'échelles.

Voici les covariances des variables,

	Prix	QT	Alc	An	Coul	EQ	Gout
Prix	1.00	0.83	0.77	-0.41	0.02	-0.05	-0.06
QT	0.83	1.00	0.90	-0.39	0.18	0.10	0.03
Alc	0.77	$0,90$	$1,00$	$-0,46$	0.07	0.04	0.01
An	-0.41	$-0,39$	$-0,46$	$1,00$	-0.37	-0.44	-0.44
Coul	0.02	0.18	0.07	-0.37	1.00	0.91	0.91
EQ	-0.05	0.10	0.04	-0.44	0.91	1.00	0.97
Gout	-0.06	0.03	0.01	-0.44	0.91	0.87	1.00

On peut par exemple lire que la variance de Prix est 1, ou que la covariance de Prix et QT est 0.83. Autrement dit, la matrice de covariance est

$$M = \begin{pmatrix} 1.00 & 0.83 & 0.77 & -0.41 & 0.02 & -0.05 & -0.06 \\ 0.83 & 1.00 & 0.90 & -0.39 & 0.18 & 0.10 & 0.03 \\ 0.77 & 0.90 & 1.00 & -0.46 & 0.07 & 0.04 & 0.01 \\ -0.41 & -0.39 & -0.46 & 1.00 & -0.37 & -0.44 & -0.44 \\ 0.02 & 0.18 & 0.07 & -0.37 & 1.00 & 0.91 & 0.91 \\ -0.05 & 0.10 & 0.04 & -0.44 & 0.91 & 1.00 & 0.97 \\ -0.06 & 0.03 & 0.01 & -0.44 & 0.91 & 0.87 & 1.00 \end{pmatrix}.$$

Diagonalisons la matrice M avec le logiciel `R`.

```
> eigen(M)
eigen() decomposition
$values
[1] 3.32824457+0.00000000i 2.63115971+0.00000000i 0.57823373+0.00000000i
[4] 0.23767625+0.00000000i 0.08762700+0.01513348i 0.08762700-0.01513348i
[7] 0.04943174+0.00000000i

$vectors
               [,1]            [,2]           [,3]            [,4]
[1,]  0.2929959+0i  0.46109369+0i 0.07650880+0i  0.78708219+0i
[2,]  0.3570511+0i  0.42458846+0i 0.31531215+0i -0.17397121+0i
[3,]  0.3355518+0i  0.44215431+0i 0.09211891+0i -0.57766476+0i
[4,] -0.3996285+0i -0.05865946+0i 0.88521527+0i  0.02798035+0i
[5,]  0.4235084+0i -0.34131085+0i 0.29850956+0i  0.08092940+0i
[6,]  0.4212389+0i -0.38004355+0i 0.10845588+0i -0.06523426+0i
[7,]  0.3969490+0i -0.38373129+0i 0.04197676+0i  0.07036397+0i
                       [,5]                    [,6]           [,7]
[1,]  0.2337009+0.1025735i  0.2337009-0.1025735i -0.22876164+0i
[2,] -0.4057999-0.2042638i -0.4057999+0.2042638i  0.59950531+0i
[3,]  0.2909895+0.1840052i  0.2909895-0.1840052i -0.40516810+0i
[4,]  0.1671658+0.0996641i  0.1671658-0.0996641i -0.09387706+0i
[5,] -0.3917276-0.2227041i -0.3917276+0.2227041i -0.11350532+0i
[6,]  0.5380058+0.0000000i  0.5380058+0.0000000i -0.44955098+0i
[7,] -0.0208365+0.2812884i -0.0208365-0.2812884i  0.44755767+0i
```

Notons que `R` voit tout dans le monde des complexes et que la matrice de vecteurs propres est orthogonale (c'est Q dans la décomposition précédente).

Observons les valeurs propres. Elles nous donnent l'importance de l'effet de M sur le vecteur propre. Ainsi, si une valeur propre est proche de zéro, l'effet de M sur le vecteur propre est insignifiant. Il devient peut important dans notre problème.

```
$values
[1] 3.32824457+0.00000000i 2.63115971+0.00000000i 0.57823373+0.00000000i
[4] 0.23767625+0.00000000i 0.08762700+0.01513348i 0.08762700-0.01513348i
[7] 0.04943174+0.00000000i
```

On peut donc choisir de ne conserver que les deux premières valeurs propre, supérieures à 1. D'ailleurs, elles représentent

$$3,33 + 2,63 = 5,99$$

unités de variance, dans un problème où les 7 variables standardisées représentent 7 unités de variance. Bref nos deux vecteurs propres associés restitueront

$$\frac{5,99}{7} \times 100 \approx 86\%$$

de la variance du problème. C'est pas mal !

Nos deux vecteurs propres associés à nos valeurs propres sont

```
$vectors
              [,1]         [,2]
[1,]  0.2929959+0i  0.46109369+0i
[2,]  0.3570511+0i  0.42458846+0i
[3,]  0.3355518+0i  0.44215431+0i
[4,] -0.3996285+0i -0.05865946+0i
[5,]  0.4235084+0i -0.34131085+0i
[6,]  0.4212389+0i -0.38004355+0i
[7,]  0.3969490+0i -0.38373129+0i
```

Ainsi, pour bien connaître un vin, je demande au producteur de me calculer et me transmettre les deux scores aux facteurs (vecteurs propres)

$$
\begin{aligned}
F_1 &= 0.29 \times Prix + 0.36 \times QT + 0.34 \times Alc - 0.4 \times An + 0.42 \times Coul + 0.42 \times EQ + 0.4 \times Gout, \\
F_2 &= 0.46 \times Prix + 0.42 \times QT + 0.44 \times Alc - 0.06 \times An - 0.34 \times Coul - 0.38 \times EQ - 0.38 \times Gout.
\end{aligned}
$$

On fait les observations suivantes.

- Le facteur F_1 est un facteur de qualité globale.
- Le facteur F_2 oppose *prix, quantité, alcool*, des indicateurs de quantité, à *l'année, la couleur, l'équilibre et le goût*, qui sont des facteurs de qualité.

Note : Il ne faut pas oublier que, dans ce problème, nous avons standardisé les variables. Nous avons les variances et moyennes de chaque variable pour calculer les facteurs dans le monde réel.

Nous avons fait une ACP... sans fioriture !