Algèbre linéaire

Guillaume Chauvet

9 octobre 2008

Table des matières

Notations

On utilisera généralement les notations suivantes :

- $u, v, w, \ldots$ désigneront des **endomorphismes**
- $x, y, z, \ldots$ désigneront des **vecteurs**
- $X, Y, Z, \ldots$ désigneront des **vecteurs** (ou **matrices**) **colonnes**
- $\alpha, \beta, \gamma, \ldots$ désigneront des **scalaires**
- $A, B, C, \ldots$ désigneront des **matrices** (généralement carrées)

Un vecteur d'un espace vectoriel E admet de multiples décompositions, selon la base de E considérée. Par abus, on notera parfois

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \textbf{au lieu de} \quad M_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mais on réservera cette notation aux coordonnées sur la base canonique $\mathcal{B}$.

Chapitre 1

Réduction des endomorphismes

Dans ce chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ des réels ou le corps $\mathbb{C}$ des complexes. Sauf mention particulière, E désigne un $\mathbb{K}$-ev de dimension finie égale à n.

1.1 L'analyse en composantes principales

Au sein de l'Analyse de Données, les méthodes factorielles visent à fournir des représentations synthétiques (généralement sous forme graphique) lorsque l'on dispose de vastes tableaux de données numériques. Différentes techniques peuvent être utilisées, selon la nature des données disponibles (variables qualitatives ou quantitatives). Nous nous intéresserons ici au cas de variables quantitatives.

On suppose qu'un ensemble de p variables quantitatives $y^1, \ldots, y^p$ a été observé sur un échantillon de taille n. Les données recueillies sont synthétisées sous la forme d'une matrice $A = (Y^1, \ldots, Y^p)$, où le vecteur colonne $Y^i = (y^i_1, \ldots, y^i_j, \ldots, y^i_n)^t$ donne les valeurs relevées pour la variable y^i sur l'ensemble des individus de l'échantillon. On peut par exemple penser aux données relevées lors d'une enquête auprès des ménages par l'Insee (Enquête Logement, Enquête Budget Des Familles, Enquête Emploi, ...), ou lors d'une enquête quelconque.

On peut, au moins théoriquement, obtenir une représentation graphique de ces données. Si on se place dans l'**espace des individus** $\mathbb{R}^n$, chaque colonne Y^i peut être représentée sous forme d'un point appelé **point variable**, dont chacune des composantes correspond à la valeur prise par la variable y^i sur un des individus de l'échantillon. Si la taille n d'échantillon est importante, cette représentation est difficilement visualisable, et fournit peu d'informa-

tion sur les données et les liens entre les variables. On peut donc chercher une **représentation simplifiée** des données, visant à conserver un maximum d'information tout en rendant les données visualisables. On va pour cela définir des **variables synthétiques** appelées les **composantes principales**, et projeter les données sur le sous-espace engendré par ces variables synthétiques.

L'information contenue dans les données est synthétisée par l'**inertie** calculée, dans l'espace des individus, comme la somme des distances (au carré) entre les points variable. Chercher la première composante principale consiste à rechercher la droite sur laquelle, après projection des données, l'inertie conservée reste maximale. La recherche du reste des composantes principales se fait de manière analogue, en complétant les m composantes déjà déterminées par une nouvelle droite permettant de maximiser l'inertie conservée, parmi tous les sous-espaces de projection de dimension $m+1$ possibles. On peut montrer que ces composantes principales s'obtiennent en calculant les valeurs propres de la matrice $A\ A^t$ (rangées par ordre décroissant), et en prenant les vecteurs propres associés.

1.2 Valeurs propres et sous-espaces propres

1.2.1 Cas d'un endomorphisme

Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. On dit que $\lambda \in \mathbb{K}$ est une **valeur propre** de u si et seulement si il existe un vecteur x de E, non nul, tel que :

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

Un vecteur x vérifiant la propriété précédente est dit **vecteur propre** de u associé à la valeur propre λ.

Comme l'énonce la proposition suivante, il existe de multiples façons de caractériser une valeur propre.

Propriété 1.1 *:*

Soit $\lambda \in \mathbb{K}$ et $u \in \mathcal{L}_{\mathbb{K}}(E)$. Les propositions suivantes sont équivalentes :

1. *λ est une valeur propre de u*
2. *$Ker(u - \lambda\ Id) \neq \{0_E\}$*

3. *$u - \lambda\ Id$ n'est pas bijective*
4. *$det(u - \lambda\ Id) = 0$ (critère qui nous sera particulièrement utile dans la suite)*

Exemple 1.1 *:*

Soit $E = \mathbb{R}^3$, $\mathcal{B} = (e_1, e_2, e_3)$ sa base canonique et $u \in \mathcal{L}_{\mathbb{K}}(E)$ l'endomorphisme de E défini par :

$$\begin{cases} u(e_1) = e_1 + 3e_2 + 2e_3 \\ u(e_2) = -e_1 + e_2 + 4e_3 \\ u(e_3) = 3e_1 - e_2 - 3e_3 \end{cases}$$

Alors 3 est valeur propre de u. En effet :

$$u(e_1 + e_2 + e_3) = 3(e_1 + e_2 + e_3)$$

On a montré du même coup que $e_1 + e_2 + e_3 = (1, 1, 1)^t$ est vecteur propre associé à 3.

Si λ est une valeur propre de u, $Ker(u - \lambda\ Id)$ est donc un sev de E non vide ; on l'appelle **sous-espace propre** de u associé à la valeur propre λ. S'il n'y pas de confusion possible avec un autre endomorphisme, on le notera E_λ. L'ensemble des valeurs propres de u est appelé **spectre** de u.

Exemple 1.2 *:*

Soit E un $\mathbb{K}$-ev quelconque.
On note Id l'application identité de E dans lui-même, et 0 l'application nulle sur E.

$\forall x \in E\ Id(x) = x = 1.x$, donc 1 est valeur propre associée à l'application identité et le sous-espace propre associé est E tout entier.

$\forall x \in E\ 0(x) = 0 = 0.x$, donc 0 est valeur propre associée à l'application identité et le sous-espace propre associé est E tout entier.

De façon plus générale, supposons que u est une homothétie de E, c'est à dire une application de E dans lui-même telle que $\exists \lambda \in \mathbb{K}\ \forall x \in E\ u(x) = \lambda\ x$. En particulier, u est un endomorphisme de E, λ est sa seule valeur propre est l'espace propre associé est E tout entier.

Propriété 1.2 *:*
Les sous-espaces propres d'un endomorphisme sont en somme directe.

Démonstration 1 *:*

Soient λ_1 et λ_2 deux valeurs propres distinctes de u, et E_{λ_1} et E_{λ_2} les deux sous-espaces propres associés.

Si $x \in E_{\lambda_1} \cap E_{\lambda_2}$, alors :

$$u(x) = \lambda_1 x \text{ et } u(x) = \lambda_2 x$$
$$\Rightarrow 0 = u(x) - u(x) = (\lambda_1 - \lambda_2)x$$
$$\Rightarrow x = 0$$

E_{λ_1} et E_{λ_2} sont donc en somme directe.

1.2.2 Cas d'une matrice

Comme pour un endomorphisme, on peut définir la notion de valeur propre d'une matrice. Si $A \in M_n(\mathbb{K})$, on dit que $\lambda \in \mathbb{K}$ est une **valeur propre** de A si et seulement si il existe un vecteur colonne $X \in \mathbb{K}^n$, non nul, tel que :

$$AX = \lambda\ X$$

Un vecteur colonne X vérifiant la propriété précédente est dit **vecteur propre** de A associé à la valeur propre λ.
On appelle encore **spectre** de A l'ensemble des valeurs propres d'une matrice A.

Il existe un lien entre valeur propre d'un endomorphisme et valeur propre d'une matrice :

Propriété 1.3 *Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$ et $\mathcal{B}$ une base quelconque de E. Alors λ est valeur propre de u si et seulement si λ est valeur propre de $M_{\mathcal{B}}(u)$.*

Démonstration 2 *Si λ est valeur propre de u, alors il existe un vecteur x de E, non nul, tel que : $u(x) = \lambda\ x$. Matriciellement, on a alors*

$$\begin{array}{rccc} & M_{\mathcal{B}}(u(x)) & = & M_{\mathcal{B}}(\lambda\ x) \\ \Rightarrow & M_{\mathcal{B}}(u)\ M_{\mathcal{B}}(x) & = & \lambda\ M_{\mathcal{B}}(x) \\ \Rightarrow & A\ X & = & \lambda\ X \end{array}$$

Réciproquement, si X est un vecteur colonne non nul tel que $A\ X = \lambda\ X$, alors en prenant x le vecteur de E tel que $M_{\mathcal{B}}(x) = X$, on a $u(x) = \lambda\ x$.

Les éléments propres (valeurs propres et vecteurs propres) d'un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(E)$ sont indépendants de la base $\mathcal{B}$ de E choisie. En revanche, les coordonnées d'un vecteur propre dépendent de la base utilisée.

Exemple 1.3 *:*

Soit $E = \mathbb{R}^2$, $\mathcal{B} = (e_1, e_2)$ *la base canonique de* E *et* u *l'endomorphisme de* E *défini par :* $\begin{cases} u(e_1) = e_1 \\ u(e_2) = e_1 + 2\ e_2 \end{cases}$.

1 *et* 2 *sont les valeurs propres de* u*, et* e_1 *et* $e_1 + e_2$ *sont des vecteurs propres qui leur sont respectivement associés (on verra plus loin une méthode effective de recherche des éléments propres d'un endomorphisme). On a donc :*

$$E_1 = Vect(e_1) \text{ et } E_2 = Vect(e_1 + e_2)$$

Cela implique que, matriciellement, les vecteurs colonne $M_{\mathcal{B}}(e_1) = (1, 0)^t$ *et* $M_{\mathcal{B}}(e_1 + e_2) = (1, 1)^t$ *sont des vecteurs propres de la matrice*

$$A = M_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

associés respectivement aux valeurs propres 1 et 2.

Si on se place maintenant dans la base $\mathcal{C} = (2\ e_1 + e_2, e_1 + 4\ e_2)$*, les vecteurs colonne* $M_{\mathcal{C}}(e_1) = (4/7, -1/7)^t$ *et* $M_{\mathcal{C}}(e_1 + e_2) = (3/7, 1/7)^t$ *sont des vecteurs propres de la matrice*

$$B = M_{\mathcal{C}}(u) = \frac{1}{7}\begin{pmatrix} 10 & 12 \\ 1 & 11 \end{pmatrix}$$

associés respectivement aux valeurs propres 1 et 2.

Remarque importante : si on travaille sur une matrice à coefficients réels, il est très souvent utile de se plonger dans $\mathbb{C}$ pour déterminer ses valeurs propres, notamment parce qu'il n'est pas toujours possible d'en trouver des réelles.

Exemple 1.4 *:*

Soit $A = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}$. *Cette matrice n'admet aucune valeur propre réelle. Si c'était le cas, on pourrait trouver un réel* λ *tel que* $A - \lambda\ I_2$ *soit non inversible, et on aurait donc :*

$$det(A - \lambda\ I_2) = 0$$

$$\Leftrightarrow (1-\lambda)(-1-\lambda) + 2 = \lambda^2 + 1 = 0$$

ce qui est bien sûr impossible. Donc ici $Sp_{\mathbb{R}}(u) = \varnothing$.

En revanche, i *et* $-i$ *sont solutions de l'équation précédente, donc en tant que matrice complexe,* A *admet* i *et* $-i$ *comme valeurs propres. On a donc* $Sp_{\mathbb{C}}(u) = \{-i, i\}$.

Une recherche directe (via la résolution d'un système de 2 équations à 2 inconnues) nous donne que le vecteur $x = \begin{pmatrix} 1 \\ -1+i \end{pmatrix}$ *est vecteur propre de* A *associé à la valeur propre* i. *On peut mettre en oeuvre une recherche directe pour un vecteur propre associé à* $-i$, *ou remarquer que comme* A *est une matrice réelle, le conjugué* $\bar{x} = \begin{pmatrix} 1 \\ -1-i \end{pmatrix}$ *du vecteur* x *est un vecteur propre associé à la valeur propre* $\bar{i} = -i$ *(voir la propriété suivante).*

Propriété 1.4 *Soit* $A \in M_n(\mathbb{R})$. *Alors* $\lambda \in Sp(A) \Rightarrow \bar{\lambda} \in Sp(A)$. *D'autre part, si* x *est un vecteur propre de* A *associé à la valeur propre* λ, *alors* $\bar{x}$ *est un vecteur propre de* A *associé à la valeur propre* $\bar{\lambda}$.

Démonstration 3 *Soit* λ *valeur propre de* A *et* X *un vecteur propre associé. Alors :*

$$\begin{array}{rrcl} & A\ X & = & \lambda\ X \\ \Rightarrow & \overline{A\ X} & = & \overline{\lambda\ X} \\ \Rightarrow & \bar{A}\ \bar{X} & = & \bar{\lambda}\ \bar{X} \\ \Rightarrow & A\ \bar{X} & = & \bar{\lambda}\ \bar{X} \end{array}$$

1.3 Recherche des valeurs propres

1.3.1 Utilisation des polynômes

Polynômes d'endomorphismes

Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. Si k est un entier positif, on note u^k l'endomorphisme composé avec lui-même k fois. Par convention, on note $u^0 = Id$. On a donc :

$$u^0 = Id$$
$$u^1 = u$$
$$u^2 = u \circ u$$
$$\ldots$$

On peut dès lors définir un polynôme d'endomorphisme : soient $a_0, \ldots, a_p \in \mathbb{K}^{p+1}$ et $P(X) = a_0 + a_1 X + \ldots + a_p X^p$ un polynôme de $\mathbb{K}[X]$. Alors par substitution :

$$P(u) = a_0 Id + a_1 u + \ldots + a_p u^p$$

Comme $\mathcal{L}_{\mathbb{K}}(E)$ est un $\mathbb{K}$ espace vectoriel, $P(u) \in \mathcal{L}_{\mathbb{K}}(E)$. De façon plus générale $K[u] = \{P(u); P \in \mathbb{K}[X]\}$ est un sous espace vectoriel de $\mathcal{L}_{\mathbb{K}}(E)$. On a de plus la propriété suivante :

Propriété 1.5 *:*

$$\forall (P, Q) \in \mathbb{K}[X]^2 \ P(u) \circ Q(u) = Q(u) \circ P(u) = (PQ)(u) = (QP)(u)$$

Démonstration 4 *:*

Soient $P = \sum_{i=1}^{p} a_i X^i$ *et* $Q = \sum_{j=1}^{q} b_j X^j$ *deux polynômes de* $\mathbb{K}[X]$*. Soit* u *un endomorphisme de* E*. Alors :*

$$P(u) \circ Q(u) = (\sum_{i=1}^{p} a_i u^i) \circ (\sum_{j=1}^{q} b_j u^j)$$
$$= \sum_{i=1,\ldots,p \ j=1,\ldots,q} a_i b_j u^{i+j}$$
$$= (\sum_{i=1,\ldots,p \ j=1,\ldots,q} a_i b_j X^{i+j})(u)$$
$$= (PQ)(u) = (QP)(u)$$

Par un raisonnement analogue, $Q(u) \circ P(u) = (PQ)(u) = (QP)(u)$*.*

Exemple 1.5 *Soit* $u \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^2)$ *défini par :*

$$\begin{cases} u(e_1) &= e_1 - e_2 \\ u(e_2) &= e_1 + 2\ e_2 \end{cases}$$

où $\mathcal{B} = (e_1, e_2)$ désigne la base canonique de $\mathbb{R}^2$. Soit $P(X) = 2X - 1$ et $Q(X) = 3X^2 + X + 1$. Les endomorphismes $Q(u)$ et $(PQ)(u)$ sont notés respectivement v et w. On a :

$$v(e_1) = (2u - Id)(e_1) = 2u(e_1) - e_1 = e_1 - 2e_2$$

$$v(e_2) = (2u - Id)(e_2) = 2u(e_2) - e_2 = 2e_1 + 3e_2$$

ce qui détermine entièrement l'endomorphisme v, vu que $\mathcal{B} = (e_1, e_2)$ est une base de $\mathbb{R}^2$. On en déduit que :

$$\begin{aligned}
w(e_1) &= ((PQ)(u))(e_1) = ((QP)(u))(e_1) \\
&= Q(u)(P(u)(e_1)) \\
&= Q(u)(e_1 - 2e_2) \\
&= 3u^2(e_1 - 2e_2) + u(e_1 - 2e_2) + (e_1 - 2e_2) \\
&= 3u(-e_1 - 5e_2) + (-e_1 - 5e_2) + (e_1 - 2e_2) \\
&= -3(e_1 - e_2) - 15(e_1 + 2\ e_2) + (-e_1 - 5e_2) + (e_1 - 2e_2) \\
&= -18e_1 - 34e_2
\end{aligned}$$

$$\begin{aligned}
w(e_2) &= ((PQ)(u))(e_2) = ((QP)(u))(e_2) \\
&= Q(u)(P(u)(e_2)) \\
&= Q(u)(2e_1 + 3e_2) \\
&= 3u^2(2e_1 + 3e_2) + u(2e_1 + 3e_2) + (2e_1 + 3e_2) \\
&= 3u(5e_1 + 4e_2) + (5e_1 + 4e_2) + (2e_1 + 3e_2) \\
&= 15(e_1 - e_2) + 12(e_1 + 2\ e_2) + (5e_1 + 4e_2) + (2e_1 + 3e_2) \\
&= 34e_1 + 16e_2
\end{aligned}$$

Polynômes de matrices

Soit $A \in M_n(\mathbb{K})$. On peut de façon analogue à ce qui précède définir A^k pour k entier positif comme le produit de A avec lui-même k fois :

$$A^0 = I$$

$$A^1 = A$$

$$A^2 = A \times A$$

$$\ldots$$

Si $P(X) = a_0 + a_1X + \ldots + a_pX^p$ un polynôme de $\mathbb{K}[X]$, on définit par linéarité :

$$P(A) = a_0I + a_1A + \ldots + a_pA^p$$

qui est une matrice de $M_n(\mathbb{K})$. De façon plus générale $K[A] = \{P(A); P \in \mathbb{K}[X]\}$ est un sous espace vectoriel de $M_n(\mathbb{K})$. La démonstration du résultat suivant est analogue à celle du résultat similaire obtenu pour les polynômes d'endomorphismes.

Propriété 1.6 *:*

$$\forall (P,Q) \in \mathbb{K}[X]^2 \; P(A) \times Q(A) = Q(A) \times P(A) = (PQ)(A) = (QP)(A)$$

1.3.2 Les polynômes annulateurs

Définition 1.1 *Soit $A \in M_n(\mathbb{K})$ et P un polynôme de $\mathbb{K}[X]$. On dit que P est un polynôme annulateur de A si et seulement si $P(A) = 0$.*

On définit de la même façon la notion de polynôme annulateur d'un endomorphisme. La propriété suivante va nous permettre de limiter le champ des recherches pour les valeurs propres d'une matrice.

Propriété 1.7 *Soit $A \in M_n(\mathbb{K})$, et λ une valeur propre de A. Alors :*

– *$\forall k \in \mathbb{N}$ λ^k est valeur propre de A^k*
– *$\forall P \in \mathbb{K}[X]$ $P(\lambda)$ est valeur propre de $P(A)$*

Soit P un polynome annulateur de A. D'après ce qui précède, $P(\lambda)$ est une valeur propre de $P(A) = 0$. Comme la seule valeur propre de la matrice nulle est 0, on en déduit que $P(\lambda) = 0$. D'où le résultat suivant :

Propriété 1.8 *:*

Soit $A \in M_n(\mathbb{K})$. Si P est un polynôme annulateur de A, alors les valeurs propres de A sont à chercher parmi les racines de P dans $\mathbb{K}$.

Nous allons établir encore quelques résultats sur les polynômes annulateurs qui nous serviront dans la suite.

Propriété 1.9 *:*

Toute matrice $A \in M_n(\mathbb{K})$ admet un polynôme annulateur non nul.

Démonstration 5 *:*

On raisonne par l'absurde : si A n'admet pas de polynôme annulateur, alors en particulier $\mathcal{F} = (I_n, A, \ldots, A^{n^2})$ est une famille libre à $n^2 + 1$ éléments. Or c'est une famille d'éléments de $M_n(\mathbb{K})$, $\mathbb{K}$-ev de dimension n^2. On a donc contradiction.

Toute matrice A admettant un polynôme annulateur, on peut définir l'ensemble $\mathcal{I}(A) = \{P \in \mathbb{K}[X] \ ; \ P(A) = 0\}$, forcément non vide. C'est un idéal de $\mathbb{K}[X]$, c'est à dire un sous-ensemble de $\mathbb{K}[X]$ vérifiant les propriétés suivantes :

$$\forall (P, Q) \in \mathcal{I}(A)^2 \ P + Q \in \mathcal{I}(A)$$

$$\forall P \in \mathcal{I}(A) \ \forall Q \in \mathbb{K}[X] \ PQ \in \mathcal{I}(A)$$

Comme on peut montrer que $\mathcal{I}(A)$ est également un sev de $\mathbb{K}[X]$, on peut résumer ces propriétés ainsi :

- Toute combinaison linéaire de polynômes annulateurs de A est un polynôme annulateur de A
- Le produit d'un polynôme annulateur de A et d'un polynôme quelconque est un polynôme annulateur de A

Définition-Propriété 1.2 *:*

Il existe un unique polynôme Q unitaire, de degré minimal, appartenant à $\mathcal{I}(A)$. Il vérifie :

$$\forall P \in \mathcal{I}(A) \quad Q \text{ divise } P \quad (\text{on note } Q/P)$$

Q est appelé ***polynôme minimal*** *de A.*

Démonstration 6 *: (difficile)*

Soit $\mathcal{D} = \{deg(P); P \in \mathcal{I}(A)\}$. $\mathcal{D}$ est une famille non vide d'éléments de $\mathbb{N}$, donc elle admet un plus petit élément p. Soit Q polynôme de $\mathcal{I}(A)$ de degré p ; quitte à le diviser par son coefficient dominant, on le choisit unitaire. De par son choix, Q est bien de degré minimal.
Si Q n'est pas unique, il existe Q_1 unitaire et de même degré que Q appartenant à $\mathcal{I}(A)$. Alors, $(Q - Q_1) \in \mathcal{I}(A)$ et est de degré $< deg(Q)$, ce qui est absurde. Q est donc unique.
Enfin, soit $P \in \mathcal{I}(A)$. On effectue la division euclidienne de P par Q ; il existe deux polynômes R et S tels que :

$$P = RQ + S \text{ avec } deg(S) < deg(Q)$$

Si $S \neq 0$, $S = P - RQ \in \mathcal{I}(A)$ alors que $deg(S) < deg(Q)$, ce qui est absurde. On en déduit que $S = 0$, et que Q divise P.

1.3.3 Le polynôme caractéristique

On a donné précedemment une caractérisation pour les valeurs propres d'une matrice A : $\lambda \in \mathbb{K}$ est une valeur propre de $A \Leftrightarrow det(A - \lambda\ I_n) = 0$. Notons $\chi_A(X) = det(A - XI_n)$. En appliquant la définition d'un déterminant :

$$\chi_A(X) = \sum_{\sigma \in \Sigma_n} (A - XI_n)_{1\sigma(1)} \dots (A - XI_n)_{n\sigma(n)}$$

Il s'agit donc d'un polynôme : on l'appelle le **polynôme caractéristique** de A. Il est de degré n car dans la somme précédente le terme obtenu pour $\sigma = Id$ est de degré n, et les autres termes de la somme sont de degré strictement inférieur à n.

λ est donc une valeur propre de A si et seulement si λ est racine du polynôme caractéristique de A. On peut alors se demander si $\chi_A(X)$ est un polynôme annulateur de A : c'est le théorême suivant, dit de Cayley-Hamilton, et dont nous admettrons la démonstration.

Théorème 1.10 *:*
Soit $A \in M_n(\mathbb{K})$. Alors $\chi_A(X)$ est un polynôme annulateur de A.

Le polynôme caractéristique est invariant par similitude. En effet, si $P \in GL_n(\mathbb{K})$:

$$\chi_{P^{-1}AP}(X) = det(P^{-1}AP - \lambda\ I_n) = det(P^{-1}(A - \lambda\ I_n)P)$$

$$= det(P^{-1})det(A - \lambda\ I_n)det(P) = det(A - \lambda\ I_n) = \chi_A(X)$$

On peut donc également parler du polynôme caractéristique d'un endomorphisme : il s'agira du polynôme caractéristique de la matrice de u dans une base quelconque.

Exemple 1.6 *:*
Reprenons l'exemple du paragraphe 1.1.3. Soit u l'endomorphisme de $\mathbb{R}^2$ défini par :

$$\begin{cases} u(e_1) = e_1 \\ u(e_2) = e_1 + 2\ e_2 \end{cases}$$

dans la base canonique $\mathcal{B} = (e_1, e_2)$. Dans cette base :

$$A = M_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

et

$$\chi_A(X) = (1-X)(2-X) = X^2 - 3X + 2$$

Soit maintenant une autre base $\mathcal{C} = (2\ e_1 + e_2, e_1 + 4\ e_2)$. *Dans cette base :*

$$B = M_{\mathcal{C}}(u) = \frac{1}{7}\begin{pmatrix} 10 & 12 \\ 1 & 11 \end{pmatrix}$$

On a

$$\begin{aligned}\chi_B(X) &= \left(\frac{10}{7} - X\right)\left(\frac{11}{7} - X\right) - \frac{12}{49} \\ &= X^2 - \frac{21}{7}X + \frac{98}{49} \\ &= X^2 - 3X + 2\end{aligned}$$

Le polynôme caractéristique nous permet d'énoncer une autre propriété qui simplifie parfois grandement la recherche des valeurs propres.

Propriété 1.11 *:*
Soit $A \in M_n(\mathbb{K})$. *Si* A *est triangulaire (donc en particulier si elle est diagonale), alors les coefficients diagonaux sont les valeurs propres de* A *dans* $\mathbb{K}$.

Démonstration 7 *:*
Si A *est triangulaire, la matrice* $A - XI_n$ *l'est aussi. On sait que son déterminant peut alors s'écrire de la façon suivante :*

$$\chi_A(X) = det(A - XI_n) = (\lambda_1 - X)\dots(\lambda_n - X)$$

en notant $\lambda_1, \dots, \lambda_n$ *les coefficients diagonaux de* A *(non nécessairement distincts).*

1.4 Diagonalisabilité

1.4.1 Rappel : matrices semblables

Soient A et B deux matrices de $M_n(\mathbb{K})$. On dit que A et B sont semblables si et seulement si $\exists P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Nous allons montrer qu'en fait, deux matrices A et B sont semblables si et seulement si elles représentent le même endomorphisme de E dans deux bases

différentes.

Condition suffisante : soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de E et u un endomorphisme de E tels que :

$$A = M_{\mathcal{B}}(u) \text{ et } B = M_{\mathcal{C}}(u)$$

Alors en utilisant les formules de changement de base :

$$M_{\mathcal{C}}(u) = P_{\mathcal{C}\mathcal{B}} M_{\mathcal{B}}(u) P_{\mathcal{B}\mathcal{C}} \Leftrightarrow B = P^{-1}AP \text{ avec } P = P_{\mathcal{B}\mathcal{C}}$$

Condition nécessaire : on suppose qu'il existe $P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$. On note A et P sous forme de vecteurs colonnes :

$$A = (A_1, \ldots, A_n) \text{ et } P = (P_1, \ldots, P_n)$$

Soit $\mathcal{B}$ une base de E. On note $\mathcal{C} = (f_1, \ldots, f_n)$ la famille de vecteurs telle que le vecteur f_i admet P_i pour vecteur de coordonnées dans $\mathcal{B}$. Comme P est inversible, on définit ainsi une autre base de E, et on a : $P = P_{\mathcal{B}\mathcal{C}}$.

On définit maintenant u, endomorphisme de E, de la façon suivante : l'image par u du $i$ème vecteur e_i de $\mathcal{B}$ admet A_i pour vecteur de coordonnées dans $\mathcal{B}$. De façon évidente, on définit bien ainsi un endomorphisme et $A = M_{\mathcal{B}}(u)$.

En résumé, on a donc prouvé que :

$$B = P^{-1}AP = P_{\mathcal{B}\mathcal{C}}^{-1} M_{\mathcal{B}}(u) P_{\mathcal{B}\mathcal{C}} = P_{\mathcal{C}\mathcal{B}} M_{\mathcal{B}}(u) P_{\mathcal{B}\mathcal{C}} = M_{\mathcal{C}}(u)$$

c'est à dire que B est la matrice de u dans la base $\mathcal{C}$, et A est la matrice de u dans la base $\mathcal{B}$.

Exemple 1.7 *:*

Soient $A = \begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix}$ *et* $B = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$*. Alors* A *et* B *sont semblables. En effet, elles représentent le même endomorphisme de* $\mathbb{R}^2$ *: dans la base canonique* $\mathcal{B} = (e_1, e_2)$ *de* $\mathbb{R}^2$ *pour* A*, dans la base* $\mathcal{C} = (e_1 + e_2, e_1 - e_2)$ *pour* B*.*

1.4.2 Définition de la diagonalisabilité

Soit $A \in M_n(\mathbb{K})$. A est dite diagonalisable si et seulement si elle est semblable à une matrice diagonale, autrement dit si il existe

$$P \in GL_n(\mathbb{K}) \text{ et } D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \text{ où } \lambda_1, \dots, \lambda_n \in \mathbb{K}^n$$

tels que $A = P^{-1}DP$.

Attention au corps utilisé : une matrice réelle peut très bien être non diagonalisable vue comme une matrice réelle, mais diagonalisable vue comme une matrice complexe. En particulier, si toutes les valeurs propres d'une matrice réelle ne sont pas réelles, cette matrice ne peut être diagonalisable dans $\mathbb{R}$.

Exemple 1.8 *:*

Reprenons l'exemple du paragraphe 1.1.3. On prend

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

Cette matrice n'admet pas de valeurs propres réelles, donc $\nexists P \in Gl_n(\mathbb{R})$ *telle que* $P^{-1}AP$ *soit diagonale.*

En revanche, i *et* $-i$ *sont valeurs propres et on a montré précedemment que* $f_1 = \begin{pmatrix} 1 \\ i-1 \end{pmatrix}$ *et* $f_2 = \begin{pmatrix} 1 \\ -i-1 \end{pmatrix}$ *sont des vecteurs propres qui leur sont respectivement associés.*

Soit $u \in \mathcal{L}_{\mathbb{C}}(\mathbb{C}^2)$*, défini par :*

$$\begin{aligned} u(e_1) &= e_1 - 2e_2 \\ u(e_2) &= e_1 - e_2 \end{aligned}$$

où $\mathcal{B} = (e_1, e_2)$ *désigne la base canonique de* $\mathbb{C}^2$*. Soit* $\mathcal{C}$ *la base constituée de* f_1 *et de* f_2*. D'après ce qui précède, on a :*

$$u(f_1) = if_1 \quad et \quad u(f_2) = (-i)f_2$$

On a donc

$$M_{\mathcal{C}}(u) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$P = P_{\mathcal{BC}} = \begin{pmatrix} 1 & 1 \\ i-1 & -i-1 \end{pmatrix}$$

est la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, donc :

$$\begin{aligned} M_{\mathcal{C}}(u) &= P_{\mathcal{CB}}\ M_{\mathcal{B}}(u)\ P_{\mathcal{BC}} \\ \Rightarrow \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} &= P^{-1}AP \end{aligned}$$

ce qui prouve que A est diagonalisable, et donne sa matrice diagonalisée et la matrice de changement de base.

Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. **L'endomorphisme u est dit diagonalisable si et seulement si il existe une base $\mathcal{B}$ de E dans laquelle $\mathbf{M_{\mathcal{B}}(u)}$ est une matrice diagonale**. On peut également caractériser la diagonalisabilité d'un endomorphisme de la façon suivante :

Propriété 1.12 *:*
Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. u est diagonalisable si et seulement si E admet une base formée de vecteurs propres de u.

Démonstration 8 *:*

Soit $\mathcal{B}$ une base de E dans laquelle on a :

$$M_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

(les λ_i ne sont pas nécessairement distincts). Alors $\forall i\ u(e_i) = \lambda_i e_i$ donc $\mathcal{B}$ est formée de vecteurs propres de u.

Réciproquement, si $\mathcal{B}$ est une base de vecteurs propres de u, alors $\forall e_i \in \mathcal{B}\ \exists \lambda \in \mathbb{K}\ u(e_i) = \lambda\ e_i$, et $M_{\mathcal{B}}(u)$ est diagonale.

Soit u un endomorphisme et $Sp(u) = \{\lambda_1, \dots, \lambda_p\}$. On a vu précedemment que les sous-espaces propres de u sont en somme directe, donc que :

$$E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$$

Si u est diagonalisable, E admet une base $\mathcal{B}$ de vecteurs propres de E. Mais comme $\mathcal{B}$ est en particulier une famille libre de vecteurs de $E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p}$, on a :

$$dim(E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p}) \geq Card(\mathcal{B}) = n$$

On en déduit la propriété suivante :

Propriété 1.13 *u est diagonalisable $\Leftrightarrow$ E est somme directe des sous-espaces propres de u.*

Propriété 1.14 *:*
Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. Si u admet n valeurs propres distinctes, alors elle est diagonalisable.

Démonstration 9 *:*
On note $\lambda_1, \ldots, \lambda_n$ les n racines distinctes de u. Les sous-espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_n}$ sont en somme directe. $F = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_n}$ est donc un sous-espace vectoriel de E de dimension

$$dim(F) = \sum_{i=1}^{n} dim(E_{\lambda_i}) \geq n = dim(E)$$

On en déduit que $F = E$: u est donc diagonalisable.

On en déduit également que si une matrice $A \in M_n(\mathbb{K})$ admet n valeurs propres distinctes, elle est diagonalisable.

1.4.3 Le théorème des noyaux et ses conséquences

Le théorême des noyaux permet d'énoncer plusieurs critères pratiques de diagonalisation. Bien que sa portée soit plus vaste, nous retiendrons surtout un de ses corollaires, le théorême de Shreier.

Définition 1.3 *Soient P et Q deux polynômes de $\mathbb{K}[X]$. On dit que P et Q sont premiers entre eux si et seulement si il n'existe pas de polynôme R non constant (i.e. de degré ≥ 1 tel que R divise P et R divise Q.*
Autrement dit, deux polynômes P et Q sont premiers entre eux si et seulement si ils n'admettent pas de diviseur commun non constant.

Théorème 1.15 *:*

Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$, $Q_1, \ldots, Q_p$ des polynômes de $\mathbb{K}[X]$ deux à deux premiers entre eux. Notons $v = Q_1(u) \circ \ldots \circ Q_p(u)$. Alors :

$$Ker(v) = Ker(Q_1(u) \ldots Q_p(u)) = \oplus_{i=1}^{p} Ker(Q_i(u))$$

Démonstration 10 *: (difficile)*

Nous allons montrer ce résultat par récurrence :

p=1 : *Pas de problème*

p=2 :

Soient Q_1 et Q_2 premiers entre eux. Par le théorême de Bezout, il existe deux polynômes A et B dans $\mathbb{K}[X]$ tels que :

$$1 = AQ_1 + BQ_2$$

On a donc :

$$Id = A(u) \circ Q_1(u) + B(u) \circ Q_2(u)$$

Soit $x \in KerQ_1(u) \cap KerQ_2(u)$. A l'aide de ce qui précède :

$$x = A(u)(Q_1(u))(x) + B(u)(Q_2(u))(x) = 0$$

On en déduit que $KerQ_1(u) \oplus KerQ_2(u)$.

De façon évidente, $KerQ_1(u) \subset Ker(Q_1(u) \circ Q_2(u))$ et $KerQ_2(u) \subset Ker(Q_1(u) \circ Q_2(u))$, d'où :

$$KerQ_1(u) \oplus KerQ_2(u) \subset Ker(Q_1(u) \circ Q_2(u))$$

Soit $x \in Ker(Q_1(u) \circ Q_2(u))$. Alors en utilisant la formule donnée par Bezout :

$$x = A(u)(Q_1(u))(x) + B(u)(Q_2(u))(x) = y + z$$

avec :

$$Q_2(u)(y) = A(u)(Q_1(u) \circ Q_2(u))(x) = 0$$
$$Q_1(u)(z) = B(u)(Q_1(u) \circ Q_2(u))(x) = 0$$

Donc, $x \in KerQ_1(u) \oplus KerQ_2(u)$.

En recollant tous les morceaux, on en tire que :

$$KerQ_1(u) \oplus KerQ_2(u) = Ker(Q_1(u) \circ Q_2(u))$$

p $\Rightarrow$ p+1 :

On suppose la propriété vérifiée pour tout entier $\leq p$. Alors, soient $Q_1, \ldots, Q_{p+1}$ des polynômes premiers entre eux. On note $R_1 = Q_1 \ldots Q_p$ et $R_2 = Q_{p+1}$. Ces deux polynômes sont premiers entre eux, donc en appliquant l'hypothèse de récurrence à R_1 et R_2 :

$$Ker(Q_1(u) \circ \ldots \circ Q_{p+1}(u)) = Ker(R_1(u)) \oplus Ker(R_2(u))$$

et en appliquant l'hypothèse de récurrence à $Q_1, \ldots, Q_p$:

$$Ker(Q_1(u) \circ \ldots \circ Q_{p+1}(u)) = \oplus_{i=1}^{p+1} Ker(Q_i(u))$$

Exemple 1.9 *:*

Soit u un endomorphisme quelconque, et P un polynôme annulateur de u, de la forme $P = Q_1 \ldots Q_p$ produit de p polynômes premiers entre eux. Alors :

$$Ker(P(u)) = E = \oplus_{i=1}^{p} Ker(Q_i(u))$$

Autrement dit, on a décomposé E en une somme directe de sous-espaces.

Exemple 1.10 *:*

Soit $u \in \mathcal{L}(\mathbb{R}^3)$ de matrice dans la base canonique :

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

En développant par rapport à la première colonne, on a $\chi_A(X) = -X((1 - X)(-X)) = -X^3 + X^2$. $P(X) = X^2(X - 1)$ est un polynôme annulateur de A donc de u. Par le théorême des noyaux :

$$E = Ker(u^2) \oplus Ker(u - Id)$$

Théorème 1.16 ***Théorême de Shreier***
Soit $u \in \mathcal{L}_{\mathbb{K}}(E)$. Alors u est diagonalisable si et seulement si u annule un polynôme scindé à racines simples dans $\mathbb{K}$ (i.e. P qui s'écrive sous la forme $(X - \lambda_1) \ldots (X - \lambda_p)$ avec $\lambda_1, \ldots, \lambda_p$ distincts et éléments de $\mathbb{K}$).

Démonstration 11 *:*

Condition suffisante :
Soit $P(X) = (X - \lambda_1) \ldots (X - \lambda_p)$ un polynôme scindé de $\mathbb{K}[X]$. On suppose que u annule P. Alors :

$$0 = (u - \lambda_1 Id) \circ \ldots \circ (u - \lambda_p Id)$$

Par le théorême des noyaux :

$$E = Ker(u - \lambda_1 Id) \oplus \ldots \oplus Ker(u - \lambda_p Id)$$

E est somme de sous-espaces propres de u, donc u est diagonalisable.

Condition nécessaire :
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E dans laquelle u est diagonale. Quitte à changer l'ordre des vecteurs de la base, on la suppose de plus de la forme :

$$M_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \lambda_p & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \lambda_p \end{pmatrix}$$

avec les $\lambda_1, \ldots, \lambda_p$ deux à deux distincts. On pose $P(X) = (X - \lambda_1) \ldots (X - \lambda_n)$.

Soit e_i un vecteur quelconque de $\mathcal{B}$. D'après la forme de $M_{\mathcal{B}}(u)$, $\exists j$ tel que $u(e_i) = \lambda_j \, e_i$. On a donc :

$$P(u)(e_i) = ((u - \lambda_1 Id) \circ \ldots \circ (u - \lambda_p Id))(e_i)$$

$$= ((u - \lambda_1 Id) \circ \ldots \circ (u - \lambda_{j-1} Id) \circ (u - \lambda_{j+1} Id) \circ (u - \lambda_p Id))(u - \lambda_j Id)(e_i)$$

$$= ((u - \lambda_1 Id) \circ \ldots \circ (u - \lambda_{j-1} Id) \circ (u - \lambda_{j+1} Id) \circ (u - \lambda_p Id))(0) = 0$$

Donc $\forall i = 1 \ldots n \; P(u)(e_i) = 0$ et par linéarité $\forall x \in E \; P(u)(x) = 0$. On en déduit que P est bien un polynôme annulateur de u.

Comme précedemment, on étend facilement ce résultat au cas matriciel : une matrice $A \in M_n(\mathbb{K})$ est diagonalisable sur $\mathbb{K}$ si et seulement si elle annule un polynôme scindé à racines simples sur $\mathbb{K}$.

1.4.4 Applications

Au final, à quoi peut servir la réduction d'endomorphismes et de matrices ? Par exemple à simplifier des calculs qui effectués "brutalement" seraient très lourds.

Calcul de puissance

Soient A et B deux matrices de $M_n(\mathbb{K})$ semblables. Il existe donc $P \in GL_n(\mathbb{K})$ telle que : $B = P^{-1}AP$. Alors, pour tout entier k strictement positif :

$$B^k = P^{-1}A^kP$$

Le cas $k = 1$ est évident. Pour $k = 2$:

$$B^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P$$

On conclut par récurrence.

Or, la puissance d'une matrice diagonale se calcule très simplement puisqu'il suffit d'élever les termes diagonaux à la puissance en question :

$$\text{Si } D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \text{ alors } \forall k \text{ entier} \geq 2 \ D^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix}$$

Si une matrice est diagonalisable, il est donc facile de calculer une puissance quelconque de cette matrice : en élevant à la puissance en question la matrice diagonale correspondante, puis en effectuant le changement de base.

Calcul d'inverse

De la même façon, il est très simple d'inverser une matrice diagonale ; il suffit d'inverser un à un ses termes diagonaux (en supposant bien sûr qu'ils soient tous non nuls).

$$\text{Si } D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \text{ alors } D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\lambda_n} \end{pmatrix}$$

Si une matrice est diagonalisable, on peut obtenir son inverse de la façon suivante : on inverse la matrice diagonale correspondante, puis on effectue le changement de base.

Etude détaillée d'un exemple

Soit la matrice

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

On commence par déterminer ses valeurs propres en calculant son polynôme caractéristique. On trouve :

$$\chi_A(X) = -(X^3 - 5X^2 + 4X) = -X(X-1)(X-4)$$

Ce polynôme admet trois racines réelles distinctes : Sp(A)=$\{0,1,4\}$. A est donc diagonalisable dans $\mathbb{R}$. 0 est valeur propre, donc A n'est pas inversible. Cherchons une base de vecteurs propres.

On note $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ un vecteur. Alors :

$$AX = 0$$

$$\Leftrightarrow \quad \begin{cases} x + y + 2z = 0 \\ 2x + y + z = 0 \\ y + 3z = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x - z = 0 \\ y + 3z = 0 \end{cases}$$

$$\Leftrightarrow \quad X \in Vect \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}$$

Donc $E_0 = Vect \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}$.

$$AX = X$$

$$\Leftrightarrow \quad \left\{ \begin{array}{c} x+y+2z=x \\ 2x+y+z=y \\ y+3z=z \end{array} \right.$$

$$\Leftrightarrow \quad \left\{ \begin{array}{c} y=-2z \\ 2x=-z \end{array} \right.$$

$$\Leftrightarrow \quad X \in Vect \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$$

Donc $E_1 = Vect \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$.

$$AX = 4X$$

$$\Leftrightarrow \quad \left\{ \begin{array}{c} x+y+2z=4x \\ 2x+y+z=4y \\ y+3z=4z \end{array} \right.$$

$$\Leftrightarrow \quad \left\{ \begin{array}{c} -7x+7z=0 \\ x-y=0 \end{array} \right.$$

$$\Leftrightarrow \quad X \in Vect \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Donc $E_4 = Vect \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

On utilise alors la base de vecteurs propres

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -3 & 4 & 1 \\ 1 & -2 & 1 \end{pmatrix}$$

On peut inverser cette matrice par opérations élémentaires sur les lignes :

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -3 & 4 & 1 \\ 1 & -2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$L_2 \leftarrow L_2 + 3L_1$$
$$L_3 \leftarrow L_3 - L_1$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 7 & 4 \\ 0 & -3 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$L_2 \leftrightarrow L_3$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 0 & 7 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix}$$

$$L_2 \leftarrow -1/3L_2$$
$$L_3 \leftarrow L_3 - 7L_2$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 0 & -1/3 \\ 2/3 & 1 & 7/3 \end{pmatrix}$$

$$L_3 \leftarrow 1/4L_3$$
$$L_1 \leftarrow L_1 - L_2 - L_3$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} 1/2 & -1/4 & -1/4 \\ 1/3 & 0 & -1/3 \\ 1/6 & 1/4 & 7/12 \end{pmatrix}$$

En effectuant le changement de base, on a :

$$P^{-1}AP = D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

On en déduit les puissances de A :

$$A^p = PD^pP^{-1} = \frac{1}{12}\begin{pmatrix} 4+2.4^p & 3.4^p & -4+7.4^p \\ 16+2.4^p & 3.4^p & -16+7.4^p \\ -8+2.4^p & 3.4^p & 8+7.4^p \end{pmatrix}$$

1.4.5 Et si la matrice n'est pas diagonalisable ?

Plongée dans le corps des complexes

Si A est une matrice réelle, elle peut être non diagonalisable dans $\mathbb{R}$ mais diagonalisable dans $\mathbb{C}$. Par exemple, imaginons que nous voulions réduire la matrice

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

Son polynôme caractéristique est égal à

$$X^3 - X^2 + X - 1 = (X-1)(X^2+1)$$

Les racines de ce polynôme sont 1, i et $-i$. A admet trois valeurs propres distinctes, mais comme elles ne sont pas toutes réelles A n'est pas diagonalisable dans $\mathbb{R}$. Elle l'est en revanche en tant que matrice complexe. Autrement dit il existe une matrice $P \in Gl_n(\mathbb{C})$ telle que :

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$$

Trigonalisation

Il n'est pas toujours sûr qu'une matrice, même vue comme à coefficients complexes, soit diagonalisable. En revanche, on a le résultat (admis) suivant :

Théorème 1.17 *:*

Soit $A \in M_n(\mathbb{C})$. Alors A est trigonalisable dans $\mathbb{C}$, c'est à dire qu'il existe une matrice $P \in GL_n(\mathbb{C})$ telle que $P^{-1}AP$ soit triangulaire.

Il n'est pas aussi facile d'inverser ou de calculer une puissance de matrice triangulaire que pour une matrice diagonale. Une forme triangulaire peut cependant être utile, notamment pour la résolution de systèmes d'équations (méthode du pivot de Gauss).

Utilisation des polynômes annulateurs pour le calcul de puissance

Soit A une matrice quelconque de $M_n(\mathbb{K})$ et P un polynôme annulateur de A (par exemple, le polynôme caractéristique). Supposons que l'on veuille calculer les puissances de A à un ordre quelconque k. Alors par division euclidienne, il existe deux polynômes Q_k et R_k dans $\mathbb{K}[X]$ tels que :

$$X^k = P(X)Q_k(X) + R_k(X) \text{ avec } d^\circ(R_k) < d^\circ(P)$$

On en déduit que :

$$A^k = P(A)Q_k(A) + R_k(A) = R_k(A)$$

c'est à dire que que toute puissance de A peut se calculer comme une simple combinaison linéaire de puissance de A inférieures à n. Ce résultat est surtout utile si on connaît un polynôme annulateur de faible degré (et donc en particulier si on travaille sur des matrices de petite taille).

Exemple 1.11 *:*
Soit

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

On va retrouver la forme générale d'une puissance de A. On a vu que le polynôme caractéristique était :

$$\chi_A(X) = -X(X-1)(X-4)$$

A n'admet pas de polynôme annulateur de plus petit degré (pourquoi ?) ; on va donc travailler avec le polynôme $P(X) = X(X-1)(X-4)$.

En effectuant la division euclidienne de X^k *par* $X(X-1)(X-4)$ *:*

$$X^k = X(X-1)(X-4)^2Q_k(X)+R_k(X) \text{ avec } d^\circ(R_k) < d^\circ(X(X-1)(X-4)) = 3$$

$$\Rightarrow A^k = R_k(A)$$

$R_k(X)$ *est de degré au plus égal à* 2, *et il dépend de* k *; on le note sous la forme :*

$$R_k(X) = a_k + b_kX + c_kX^2$$

Pour calculer ces deux coefficients, on revient à la formule :

$$X^k = X(X-1)(X-4)^2Q_k(X) + a_k + b_kX + c_kX^2$$

On applique cette formule en les valeurs propres de A (i.e. en les racines du polynôme $X(X-1)(X-4)$), ce qui permet d'obtenir les trois équations :

$$\left\{\begin{array}{rcccccc} 0 & = & a_k & & & & \\ 1 & = & a_k & + & b_k & + & c_k \\ 4^k & = & a_k & + & 4b_k & + & 16c_k \end{array}\right.$$

Après résolution, on obtient :

$$a_k = 0 \quad b_k = \frac{1}{3}(4^{k-1}-1) \quad c_k = \frac{4}{3}(1-4^{k-2})$$

On en déduit la forme générale de A^k :

$$A^k = \frac{1}{3}(4^{k-1}-1)\ A + \frac{4}{3}(1-4^{k-2})\ A^2$$

$$\Rightarrow A^k = KD^kK^{-1} = \frac{1}{12}\begin{pmatrix} 4+2.4^k & 3.4^k & -4+7.4^k \\ 16+2.4^k & 3.4^k & -16+7.4^k \\ -8+2.4^k & 3.4^k & 8+7.4^k \end{pmatrix}$$

Exemple 1.12 *:*

Soit

$$A = \begin{pmatrix} -2 & -2 & -3 \\ -4 & 0 & -3 \\ 8 & 4 & 8 \end{pmatrix}$$

Le polynôme caractéristique de A est $\chi_A(X) = -(X-2)^3$. On en déduit d'ores et déjà que A ne peut être diagonalisable (pourquoi ?).

$\chi_A(X)$ est un polynôme annulateur de A. Peut-on en trouver un de plus petit degré ? On sait que le polynôme minimal est de coefficient dominant égal à 1 et qu'il divise $\chi_A(X)$, donc on doit le chercher parmi les polynômes $\{X-2, (X-2)^2, (X-2)^3\}$. Il est facile de vérifier que $(X-2)^2$ est en fait le polynôme minimal.

On désire calculer les puissances de A à un ordre k quelconque. En effectuant la division euclidienne de X^k par $(X-2)^2$:

$$X^k = (X-2)^2Q_k(X) + R_k(X) \text{ avec } d^\circ(R_k) < d^\circ((X-2)^2) = 2$$

$$\Rightarrow A^k = R_k(A)$$

$R_k(X)$ est de degré au plus égal à 1, et il dépend de k ; on le note sous la forme :

$$R_k(X) = a_k + b_k X$$

Pour calculer ces deux coefficients, on revient à la formule :

$$X^k = (X-2)^2 Q_k(X) + a_k + b_k X$$

En la dérivant :

$$kX^{k-1} = (X-2)(2Q_k(X) + (X-2)Q'_k(X)) + b_k$$

On applique ces formules pour $X = 2$ et on obtient :

$$2^k = a_k + 2b_k$$

$$k2^{k-1} = b_k$$

On en déduit que $b_k = k2^{k-1}$ et $a_k = 2^k(1-k)$, et donc que :

$$A^k = 2^{k-1}(kA + 2(1-k)I_3)$$

Utilisation des polynômes annulateurs pour le calcul d'inverse

Même si une matrice est diagonalisable, la diagonaliser en vue de l'inverser n'est pas toujours la solution la plus simple, surtout si on lui connaît un polynôme annulateur "pas trop compliqué". En effet, soit $A \in GL_n(\mathbb{K})$ une matrice inversible. On a vu précedemment qu'on pouvait toujours lui trouver un polynôme annulateur

$$P(X) = a_0 + a_1 X + \ldots + a_n X^n$$

Par exemple, son polynôme caractéristique. A étant inversible, on peut même le choisir tel que $a_0 \neq 0$ (pourquoi ?). En appliquant le théorème de Cayley-Hamilton :

$$0 = a_0 I_n + a_1 A + \ldots + a_n A^n$$

$$\Rightarrow A^{-1} = \frac{-1}{a_0}(a_1 I_n + \ldots + a_n A^{n-1})$$

L'inverse de A se calcule donc comme une combinaison linéaire de puissances de A inférieures à $n-1$.

Exemple 1.13 *:*

Reprenons l'exemple précédent avec

$$A = \begin{pmatrix} -2 & -2 & -3 \\ -4 & 0 & -3 \\ 8 & 4 & 8 \end{pmatrix}$$

Son polynôme minimal est $(X-2)^2$. 0 *n'est pas racine de ce polynôme, donc* A *est inversible. On a :*

$$\begin{aligned} A^2 - 4A + 4I_3 &= 0 \\ \Rightarrow A - 4I_3 + 4A^{-1} &= 0 \\ \Rightarrow A^{-1} &= \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & \frac{3}{4} \\ 1 & 1 & \frac{3}{4} \\ -2 & -1 & -1 \end{pmatrix} \end{aligned}$$

Chapitre 2

Equations de récurrence linéaire

Nous allons maintenant passer à l'étude des équations de récurrence linéaire. Elles sont notamment utilisées en Séries Temporelles, où on cherche à expliquer des variables à l'aide de leur passé.
On se limitera à l'étude des équations de récurrence linéaire à coefficients réels.

2.1 Définition et premières propriétés

On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ vérifie une **équation de récurrence linaire** d'ordre p (que l'on note ERL) s'il existe des **réels** $a_1, \ldots, a_p$ et une suite $(v_n)_{n\in\mathbb{N}}$ tels que :

$$\forall n \; u_{n+p} - a_1 u_{n+p-1} - \ldots - a_p u_n = v_n \quad (1)$$

et si les p premiers termes $u_0, \ldots, u_{p-1}$ sont donnés. L'équation de récurrence linéaire

$$u_{n+p} - a_1 u_{n+p-1} - \ldots - a_p u_n = 0 \quad (1')$$

obtenue en supprimant le second membre v_n est appelée **équation de récurrence linéaire homogène** (ERLH) associée à l'ERL.

Par extension, on dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ vérifie une équation de récurrence linéaire homogène si elle vérifie une relation de type (1) avec $\forall n \; v_n = 0$.

Exemple :

Soit l'ERL :

$$u_{n+2} - 2u_{n+1} + u_n = 3 \text{ avec } u_0 = 0 \; u_1 = 1$$

Elle n'est pas homogène, car ici le terme v_n n'est pas égal à 0. L'ERLH associée est :

$$u_{n+2} - 2u_{n+1} + u_n = 0$$

Propriété 2.1 *L'ensemble S_0 des solutions d'une ERLH d'ordre p est un sous-espace vectoriel de dimension p de $\mathcal{S}(\mathbb{R})$ (ensemble des suites réelles).*

Démonstration 12 *On utilise la forme générale (1') de l'ERLH donnée précédemment. L'ensemble S_0 des solutions de (1') est un sous-ensemble de $\mathcal{S}(\mathbb{R})$, non vide car la suite nulle lui appartient. On montre facilement que toute combinaison linéaire d'éléments de S_0 est encore un élément de S_0 : c'est donc bien un sev de $\mathcal{S}(\mathbb{R})$.*

Considérons maintenant l'application

$$\phi : \begin{array}{ccc} S_0 & \longrightarrow & \mathbb{R}^p \\ (u_n)_{n\in\mathbb{N}} & \mapsto & (u_0, \ldots, u_{p-1}) \end{array}$$

On montre facilement que ϕ est une application linéaire de $\mathcal{S}(\mathbb{R})$ dans $\mathbb{R}^p$. Elle est injective (car si $u_0 = u_1 = \ldots = u_{p-1} = 0$ la relation de récurrence implique que $u_n = 0 \; \forall n$), surjective (on peut toujours trouver une suite respectant la relation de récurrence et avec $p-1$ premiers termes fixés a priori), donc bijective. C'est donc un isomorphisme de S_0 dans $\mathbb{R}^p$, et $dim(S_0) = p$.

On retiendra de cette propriété que, S_0 étant un espace vectoriel de dimension p, il nous faut trouver p suites linéairement indépendantes et vérifiant l'équation homogène pour entièrement caractériser S_0.

Propriété 2.2 *Soit S l'ensemble des solutions de l'ERL (1) et $(a_n)_{n\in\mathbb{N}}$ un élément quelconque de S. Alors :*

$$S = (a_n)_{n\in\mathbb{N}} + S_0$$

où S_0 désigne l'ensemble des solutions de l'ERLH associée (1').

Autrement dit, **la solution générale de S s'obtient en ajoutant à la solution générale de S_0 une solution particulière de S.**

On va donc diviser la résolution d'une équation de récurrence linéaire en 3 étapes :

1. On trouve une base de l'ensemble S_0 des solutions de l'ERLH associée, ce qui nous donne la solution générale $(u_n^0)_{n\in\mathbb{N}}$ de l'ERLH
2. On trouve une solution de l'ERL $(a_n)_{n\in\mathbb{N}}$
3. On en déduit la solution générale (i.e. sans tenir compte des conditions initiales) de l'équation de l'ERL, $(u_n^0 + a_n)_{n\in\mathbb{N}}$. La solution particulière s'obtient à l'aide des conditions initiales.

Dans la suite du chapitre, on continuera à désigner par S l'ensemble des solutions d'une ERL, et par S_0 l'ensemble des solutions de l'ERLH associée.

2.2 Résolution d'une Equation de Récurrence Linéaire Homogène

2.2.1 ERLH d'ordre 1

La forme générale d'une ERLH d'ordre 1 est

$$u_{n+1} - au_n = 0$$

où a est un réel.

Cette équation définit une suite géométrique de raison a. On a donc le résultat suivant :

Propriété 2.3 *La solution générale de cette ERLH est $u_n = a^n u_0$, où $u_0 \in \mathbb{R}$.*

Exemple 2.1

Soit à résoudre l'équation :

$$\forall n \in \mathbb{N} \quad u_{n+1} - 4u_n = 0$$

La solution de cette ERLH est la suite de terme général :

$$u_n = u_0\ 4^n$$

2.2.2 ERLH d'ordre 2

La forme générale d'une ERLH d'ordre 2 est

$$\forall n \in \mathbb{N} \quad u_{n+2} - a_2\, u_{n+1} - a_1\, u_n = 0$$

où a_1 et a_2 sont des réels.

Définition 2.1 *Soit l'équation de degré* 2

$$x^2 - a_2 x - a_1 = 0 \quad (E)$$

(E) est appelée ***équation associée*** *à l'ERLH.*

On note $\Delta = a_2^2 + 4a_1$ le discriminant de l'équation associée.

Propriété 2.4 *Avec les notations précédentes :*

- *Si* $\Delta > 0$*, soient* α *et* β *les deux racines réelles de (E). Alors* $(\alpha^n)_{n\in\mathbb{N}}$ *et* $(\beta^n)_{n\in\mathbb{N}}$ *forment une base de* S_0
- *Si* $\Delta = 0$*, soit* α *la racine double de (E). Alors* $(\alpha^n)_{n\in\mathbb{N}}$ *et* $(n\alpha^n)_{n\in\mathbb{N}}$ *forment une base de* S_0
- *Si* $\Delta < 0$*, soient* $\alpha = \rho\ e^{i\theta}$ *et* $\overline{\alpha} = \rho\ e^{-i\theta}$ *les deux racines complexes conjuguées de (E). Alors* $(\rho^n\ \cos(n\theta))_{n\in\mathbb{N}}$ *et* $(\rho^n\ \sin(n\theta))_{n\in\mathbb{N}}$ *forment une base de* S_0

Démonstration 13 *On montre que, dans chacun des trois cas, les deux suites forment une famille libre. Comme* S_0 *est de dimension* 2*, si on montre que ce sont des éléments de* S_0*, c'est gagné.*

Premier cas : $\Delta > 0$
Avec les notations de la propriété, on a :

$$\alpha^{n+2} - a_2\alpha^{n+1} - a_1\alpha^n = \alpha^n(\alpha^2 - a_2\alpha - a_1) = 0$$

car α *est racine de (E). Le même raisonnement est valable pour* β*. On en déduit que* $(\alpha^n)_{n\in\mathbb{N}}$ *et* $(\beta^n)_{n\in\mathbb{N}}$ *forment une base de* S_0*.*

Deuxième cas : $\Delta = 0$
α *étant une racine double de (E), on a :* $\alpha^2 - a_2\alpha - a_1 = 0$ *et* $2\alpha - a_2 = 0$*. Le même raisonnement que précedemment montre que* $(\alpha^n)_{n\in\mathbb{N}}$ *est élément de* S_0*. D'autre part :*

$$(n+2)\alpha^{n+2} - a_2(n+1)\alpha^{n+1} - a_1 n\alpha^n = n\alpha^n(\alpha^2 - a_2\alpha - a_1) + \alpha^{n+1}(2\alpha - a_2) = 0$$

Donc $(n\alpha^n)_{n\in\mathbb{N}}$ *est également élément de* S_0.

Troisième cas : $\Delta < 0$
$\alpha = \rho\ e^{i\theta}$ *vérifie* $\alpha^2 - a_2\alpha - a_1 = 0$, *donc :*

$$\rho^{n+2}e^{i(n+2)\theta} - a_2\rho^{n+1}e^{i(n+1)\theta} - a_1\rho^n e^{in\theta} = \rho^n e^{in\theta}(\alpha^2 - a_2\alpha - a_1) = 0$$

En passant à la partie réelle, puis à la partie imaginaire, on en déduit que $(\rho^n\ \cos(n\theta))_{n\in\mathbb{N}}$ *et* $(\rho^n\ \sin(n\theta))_{n\in\mathbb{N}}$ *sont éléments de* S_0

Exemple 2.2

Soit à résoudre l'ERLH :

$$\forall n \in \mathbb{N}\ \ u_{n+2} + 2u_{n+1} - 3u_n = 0$$

L'équation associée est :

$$x^2 + 2x - 3 = 0$$

Le discriminant de cette équation vaut 16, on en déduit que les racines de l'équation sont 1 *et* -3.

Les suites $(1)_{n\in\mathbb{N}}$ *et* $(-3)_{n\in\mathbb{N}}$ *forment donc une base de* S_0, *et la solution générale de l'ERLH est la suite de terme général :*

$$u_n = \lambda + \beta\ (-3)^n$$

où λ *et* β *sont des réels.*

Exemple 2.3

Soit à résoudre l'ERLH :

$$\forall n \in \mathbb{N}\ \ u_{n+2} - 2u_{n+1} + 2u_n = 0$$

L'équation associée est :

$$x^2 - 2x + 2 = 0$$

Le discriminant de cette équation vaut -4, on en déduit que l'équation admet deux racines complexes conjuguées $1+i=\sqrt{2}e^{i\pi/4}$ et $1-i=\sqrt{2}e^{-i\pi/4}$.

Les suites $(\sqrt{2}^{\,n}\cos(n\pi/4))_{n\in\mathbb{N}}$ et $(\sqrt{2}^{\,n}\sin(n\pi/4))_{n\in\mathbb{N}}$ forment donc une base de S_0, et la solution générale de l'ERLH est la suite de terme général :

$$u_n = \lambda\ \sqrt{2}^{\,n}\ \cos(n\pi/4) + \beta\ \sqrt{2}^{\,n}\ \sin(n\pi/4)$$

où λ et β sont des réels.

2.2.3 ERLH d'ordre p

On va maintenant revenir à la forme générale d'une ERLH, et généraliser les résultats précédents.

Définition 2.2 *Soit l'équation de degré p*

$$x^p - a_1x^{p-1} - \ldots - a_p = 0 \quad (E)$$

(E) est appelée ***équation associée*** *à l'ERL (1).*

Propriété 2.5 *Soient $\beta_1,\ldots,\beta_q$ les racines réelles de (E), d'ordre respectif $i_1,\ldots,i_q$.*
Soient $\alpha_1 = \rho_1e^{i\theta_1}, \overline{\alpha_1} = \rho_1e^{-i\theta_1}, \ldots, \alpha_r = \rho_re^{i\theta_r}, \overline{\alpha_r} = \rho_re^{-i\theta_r}$ les racines complexes conjuguées de (E), d'ordre respectif $j_1,\ldots,j_r$.
(On a en particulier : $(i_1+\ldots+i_q)+2(j_1+\ldots+j_r)=p$)

Alors les suites :

- $(\beta_1^n)_{n\in\mathbb{N}}, \ldots, (n^{i_1-1}\beta_1^n)_{n\in\mathbb{N}}$
- $\ldots$
- $(\beta_q^n)_{n\in\mathbb{N}}, \ldots, (n^{i_q-1}\beta_q^n)_{n\in\mathbb{N}}$
- $(\rho_1^n\cos(n\theta_1))_{n\in\mathbb{N}}, (\rho_1^n\sin(n\theta_1))_{n\in\mathbb{N}}, \ldots, (n^{j_1-1}\rho_1^n\cos(n\theta_1))_{n\in\mathbb{N}}, (n^{j_1-1}\rho_1^n\sin(n\theta_1))_{n\in\mathbb{N}}$
- $\ldots$
- $(\rho_r^n\cos(n\theta_r))_{n\in\mathbb{N}}, (\rho_r^n\sin(n\theta_r))_{n\in\mathbb{N}}, \ldots, (n^{j_r-1}\rho_r^n\cos(n\theta_r))_{n\in\mathbb{N}}, (n^{j_r-1}\rho_r^n\sin(n\theta_r))_{n\in\mathbb{N}}$

forment une base de S_0.

Démonstration 14 *Bien qu'un peu plus calculatoire, le principe de la démonstration est le même que pour une ERL d'ordre 2.*

Exemple 2.4

Soit à résoudre l'ERLH :

$$\forall n \in \mathbb{N} \quad u_{n+5} - 3u_{n+4} + 4u_{n+3} - 4u_{n+2} + 3u_{n+1} - u_n = 0$$

On notera la factorisation suivante :

$$x^5 - 3x^4 + 4x^3 - 4x^2 + 3x - 1 = (x^2 + 1)(x - 1)^3$$

Compte-tenu de l'indication, l'équation associée à l'ERLH admet :

- *Une racine réelle triple :* 1
- *Deux racines complexes conjuguées simples :* $i = e^{i\pi/2}$ *et* $-i = e^{-i\pi/2}$

Les suites $(1)_{n\in\mathbb{N}}$, $(n)_{n\in\mathbb{N}}$, $(n^2)_{n\in\mathbb{N}}$, $(\cos(n\pi/2))_{n\in\mathbb{N}}$ *et* $(\sin(n\pi/2))_{n\in\mathbb{N}}$ *forment donc une base de* S_0*, et la solution générale de l'ERLH est la suite de terme général :*

$$\begin{array}{rcl} u_n & = & \lambda_1 + \lambda_2\ n + \lambda_3\ n^2 \\ & + & \lambda_4\ \cos(n\pi/2) + \lambda_5\ \sin(n\pi/2) \end{array}$$

où $\lambda_1, \ldots, \lambda_5$ *sont des réels.*

Exemple 2.5

Soit à résoudre l'ERLH :

$$\forall n \in \mathbb{N} \quad u_{n+5} - 6u_{n+4} + 20u_{n+3} - 40u_{n+2} + 48u_{n+1} - 32u_n = 0$$

On notera la factorisation suivante :

$$x^5 - 6x^4 + 20x^3 - 40x^2 + 48x - 32 = (x^2 - 2x + 4)^2(x - 2)$$

L'équation $x^2 - 2x + 4 = 0$ *a pour discriminant* -12*, et admet deux racines complexes conjuguées* $1 + i\sqrt{3} = 2\ e^{i\pi/3}$ *et* $1 - i\sqrt{3} = 2\ e^{-i\pi/3}$*.*

On en déduit que l'équation associée à l'ERLH admet :

– *Une racine réelle simple :* 2
– *Deux racines complexes conjuguées doubles :* $2\ e^{i\pi/3}$ *et* $2\ e^{-i\pi/3}$

Les suites $(2^n)_{n\in\mathbb{N}}$, $(2^n\ \cos(n\pi/3))_{n\in\mathbb{N}}$, $(2^n\ \sin(n\pi/3))_{n\in\mathbb{N}}$, $(2^n\ n\ \cos(n\pi/3))_{n\in\mathbb{N}}$ *et* $(2^n\ n\ \sin(n\pi/3))_{n\in\mathbb{N}}$ *forment donc une base de* S_0*, et la solution générale de l'ERLH est la suite de terme général :*

$$\begin{array}{rcc} u_n & = & \lambda_1\ 2^n \\ & + & \lambda_2\ 2^n\ \cos(n\pi/3) + \lambda_3\ 2^n\ \sin(n\pi/3) \\ & + & \lambda_4\ 2^n\ n\ \cos(n\pi/3) + \lambda_5\ 2^n\ n\ \sin(n\pi/3) \end{array}$$

où $\lambda_1, \ldots, \lambda_5$ *sont des réels.*

2.3 Recherche d'une solution particulière d'une ERL

Il n'est pas toujours simple de produire une solution particulière d'une ERL. Nous allons voir deux cas où l'on peut toujours en obtenir une :

– Celui où le second membre de l'équation est un polynôme en n
– Celui où le second membre de l'équation est une puissance en n

Notons tout de suite un point important : **Les conditions initiales ne vont pas intervenir dans la recherche de cette solution particulière**.

2.3.1 Cas où le second membre est un polynôme en n

On suppose que v_n est un polynôme en n, c'est à dire qu'il existe un polynôme Q à coefficients réels tel que

$$\forall n\ v_n = Q(n)$$

On note q le degré du polynôme Q.

Si 1 n'est pas racine de l'équation associée

On peut trouver une solution particulière de terme général $u_n = R(n)$, où R désigne un polynôme à coefficients réels de degré q. On cherche donc des coefficients $b_0, \ldots, b_q$ définissant une suite de terme général

$$u_n = b_0 + b_1 n + \ldots + b_q n^q$$

Exemple :

Soit à résoudre l'équation de récurrence linéaire

$$u_{n+3} - 2u_{n+2} - 2u_{n+1} - 3u_n = n + 2$$

L'équation associée est $x^3 - 2x^2 - 2x - 3 = 0$, dont 3, $e^{i\frac{2\pi}{3}}$ et $e^{-i\frac{2\pi}{3}}$ sont racines simples.
Le second membre est un polynôme en n de degré 1, et 1 n'est pas racine de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = dn + e$$

où d et e sont des constantes. Après un peu de calcul, on trouve $d = -\frac{1}{6}$ et $e = -\frac{1}{4}$.

Si 1 est racine simple de l'équation associée

On peut trouver une solution particulière de terme général $u_n = nR(n)$, où R désigne un polynôme à coefficients réels de degré q. On cherche donc des coefficients $b_0, \ldots, b_q$ définissant une suite de terme général

$$u_n = n(b_0 + b_1 n + \ldots + b_q n^q)$$

Exemple :

Soit à résoudre l'équation de récurrence linéaire

$$u_{n+3} - u_n = 2n$$

L'équation associée est $x^3 - 1 = 0$, dont 1, $e^{i\frac{2\pi}{3}}$ et $e^{-i\frac{2\pi}{3}}$ sont racines simples.
Le second membre est un polynôme en n de degré 3, et 1 est racine simple de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = n(b_0 + b_1 n) = b_0\ n + b_1\ n^2$$

où b_0 et b_1 sont des constantes à déterminer. On doit avoir :

$$\begin{array}{lcc} & u_{n+3} - u_n & = & 2n \\ \Rightarrow & (b_0(n+3) + b_1(n^2 + 6n + 9)) - (b_0\ n + b_1\ n^2) & = & 2n \\ \Rightarrow & (3b_0 + 9b_1) + b_1(6n) & = & 2n \\ \Rightarrow & b_1 & = & 1/3 \end{array}$$

On en déduit que $b_0 = -1$ et $b_1 = 1/3$.

Si 1 est racine multiple de l'équation associée

De façon plus générale, supposons que 1 est racine de l'équation associée d'ordre de multiplicité m.
On peut trouver une solution particulière de terme général $u_n = n^m R(n)$, où R désigne un polynôme à coefficients réels de degré q. On cherche donc des coefficients $b_0, \ldots, b_q$ définissant une suite de terme général

$$u_n = b_0 n^m + b_1 n^{m+1} + \ldots + b_q n^{m+q}$$

Exemple :

Soit à résoudre l'ERL :

$$u_{n+5} - 3u_{n+4} + 4u_{n+3} - 4u_{n+2} + 3u_{n+1} - u_n = n + 2$$

L'équation associée est :

$$x^5 - 3x^4 + 4x^3 - 4x^2 + 3x - 1 = 0$$

Cette équation peut s'écrire sous forme factorisée :

$$(x^2 + 1)(x - 1)^3 = 0$$

On en déduit donc que 1 est racine triple de cette équation, et que $i = e^{\frac{i\pi}{2}}$ et $-i = e^{\frac{-i\pi}{2}}$ en sont racines simples.
Le second membre est un polynôme en n de degré 1, et 1 est racine triple de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = n^3(b_0 + b_1 n)$$

où b_0 et b_1 sont des constantes. On doit avoir :

$$u_{n+5} - 3u_{n+4} + 4u_{n+3} - 4u_{n+2} + 3u_{n+1} - u_n = n + 2$$

ce qui donne après calcul et identification des coefficients $b_0 = -\frac{1}{54}$ et $b_1 = \frac{1}{54}$.

Bilan :

Si le second membre de l'ERL est un polynôme en n de degré q, on cherche une solution particulière de l'ERL de terme général $n^m\ R(n)$, où :

– R **est un polynôme réel (à déterminer) de degré** q
– m **désigne l'ordre de multiplicité de** 1 **comme racine de l'équation associée à l'ERL (**m **peut être nul)**

2.3.2 Cas où le second membre est une puissance en n

On suppose qu'il existe deux constantes α et b telles que

$$\forall n \; v_n = b\alpha^n$$

Si α n'est pas racine de l'équation associée

On peut trouver une solution particulière de terme général $u_n = b_0\alpha^n$, où b_0 est une constante.

Exemple :

Soit à résoudre l'équation de récurrence linéaire

$$u_{n+4} - 6u_{n+3} + 13u_{n+2} - 12u_{n+1} + 4u_n = 3^n$$

L'équation associée est

$$x^4 - 6x^3 + 13x^2 - 12x + 4 = 0$$

que l'on peut encore écrire sous forme factorisée

$$(x-1)^2(x-2)^2 = 0$$

1 et 2 sont toutes les 2 racines doubles de cette équation.

Le second membre est une puissance en n, et $\alpha = 3$ n'est pas racine de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = b_0 3^n$$

On a :

$$\begin{array}{rcc} & u_{n+4} - 6u_{n+3} + 13u_{n+2} - 12u_{n+1} + 4u_n & = & 3^n \\ \Rightarrow & b_0 3^{n+4} - 6b_0 3^{n+3} + 13b_0 3^{n+2} - 12b_0 3^{n+1} + 4b_0 3^n & = & 3^n \\ \Rightarrow & b_0\,(81 - 162 + 117 - 36 + 4) & = & 1 \\ \Rightarrow & b_0 & = & 1/4 \end{array}$$

Si α est racine simple de l'équation associée

On peut trouver une solution particulière de terme général $u_n = b_0\ n\ \alpha^n$, où b_0 est une constante.

Exemple :

Soit à résoudre l'équation de récurrence linéaire

$$u_{n+2} + u_{n+1} - 2u_n = 4(-2)^n$$

L'équation associée est

$$x^2 + x - 2 = 0$$

que l'on peut encore écrire sous forme factorisée

$$(x-1)(x+2) = 0$$

1 et -2 sont toutes les deux racines simples de cette équation.

Le second membre est une puissance en n, et $\alpha = -2$ est racine simple de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = b_0 n(-2)^n$$

Comme cette suite doit vérifier

$$u_{n+2} + u_{n+1} - 2u_n = 4(-2)^n$$

on obtient après identification $b_0 = \frac{2}{3}$.

Si α est racine multiple de l'équation associée

Soit m l'ordre de multiplicité de α comme racine de l'équation. On peut trouver une solution particulière de terme général $u_n = b_0\ n^m\ \alpha^n$, où b_0 est une constante.

Exemple :

Soit à résoudre l'équation de récurrence linéaire

$$u_{n+4} + 2u_{n+3} - 2u_{n+1} - u_n = 3(-1)^n$$

L'équation associée est

$$x^4 + 2x^3 - 2x - 1 = 0$$

que l'on peut encore écrire sous forme factorisée

$$(x-1)(x+1)^3 = 0$$

Donc 1 est racine simple de l'équation, et -1 en est racine triple.

Le second membre est une puissance en n, et $\alpha = -1$ est racine triple de l'équation associée. On cherche donc une solution particulière de terme général :

$$u_n = b_0 n^3 (-1)^n$$

Comme cette suite doit vérifier

$$u_{n+4} + 2u_{n+3} - 2u_{n+1} - u_n = 3(-1)^n$$

on obtient après identification $b_0 = \frac{1}{4}$.

Bilan :

Si le second membre de l'ERL est une puissance en n, i.e. une suite de terme général $b\ \alpha^n$, on cherche une solution particulière de l'ERL de terme général $b_0\ n^m\ \alpha^n$, où :

- b_0 **est une constante**
- m **désigne l'ordre de multiplicité de α comme racine de l'équation associée à l'ERL (m peut être nul)**

2.4 Résolution de l'ERL

Les deux paragraphes précédents nous permettent de donner la solution générale de l'ERL ; il reste à tenir compte des conditions initiales pour déterminer l'unique suite solution du problème.

Exemple 1 :

Soit à résoudre l'ERL :

$$u_{n+5} - 6u_{n+4} + 20u_{n+3} - 40u_{n+2} + 48u_{n+1} - 32u_n = -\frac{221}{2}\ 3^n$$

avec les conditions initiales

$$u_0 = 0\ u_1 = 0\ u_2 = 0\ u_3 = 0\ u_4 = 1$$

L'ERLH associée est :

$$u_{n+5} - 6u_{n+4} + 20u_{n+3} - 40u_{n+2} + 48u_{n+1} - 32u_n = 0$$

Et l'équation associée vaut :

$$x^5 - 6x^4 + 20x^3 - 40x^2 + 48x^1 - 32 = 0 \quad (E1)$$

On a vu précedemment que cette équation s'écrivait sous forme factorisée :

$$(x^2 - 2x + 4)^2(x - 2) = 0$$

d'où l'on a déduit que

$$\begin{aligned} S_0 \ = \ & \Big\{(u_n)_{n\in\mathbb{N}} = \big(\beta_1\ 2^n + \beta_2\ 2^n \cos\tfrac{n\pi}{3} + \beta_3\ 2^n \sin\tfrac{n\pi}{3} \\ & +\beta_4\ n\ 2^n \cos\tfrac{n\pi}{3} + \beta_5\ n\ 2^n \sin\tfrac{n\pi}{3}\big)_{n\in\mathbb{N}}\,; (\beta_1, \ldots, \beta_5) \in \mathbb{R}^5\Big\} \end{aligned}$$

On cherche maintenant une solution particulière de l'ERL ; d'après ce qui précède, on la cherche sous la forme d'une suite de terme général $a\ 3^n$ où a est une constante à déterminer. On doit avoir :

$$a(3^{n+5} - 6 \times 3^{n+4} + 20 \times 3^{n+3} - 40 \times 3^{n+2} + 48 \times 3^{n+1} - 32 \times 3^n) = -\frac{221}{2}\ 3^n$$

$$\Leftrightarrow a(243 - 486 + 270 - 360 + 144 - 32) = -\frac{221}{2}$$

$$\Leftrightarrow a = 2$$

Une solution particulière de l'ERL est donc la suite de termes général 2×3^n, où a est donné par l'équation précédente.

On en déduit l'ensemble des solutions :

$$S = \Big\{(u_n)_{n\in\mathbb{N}} = \big(2\times 3^n + \beta_1\ 2^n + \beta_2\ 2^n \cos\tfrac{n\pi}{3} + \beta_3\ 2^n \sin\tfrac{n\pi}{3}$$

$$+\beta_4\ n\ 2^n \cos\tfrac{n\pi}{3} + \beta_5\ n\ 2^n \sin\tfrac{n\pi}{3}\big)_{n\in\mathbb{N}}\,;(\beta_1,\ldots,\beta_5)\in\mathbb{R}^5\Big\}$$

Les coefficients β_i sont donnés par la résolution du système (obtenu à l'aide des conditions initiales) :

$$\begin{cases} 2 & +\beta_1 & +\beta_2 & & & & = 0 \\ 6 & +2\beta_1 & +\beta_2 & +\beta_3\ \sqrt{3} & +\beta_4 & +\beta_5\ \sqrt{3} & = 0 \\ 18 & +4\beta_1 & -2\beta_2 & +2\beta_3\ \sqrt{3} & -4\beta_4 & +4\beta_5\ \sqrt{3} & = 0 \\ 54 & +8\beta_1 & -8\beta_2 & & -24\beta_4 & +\beta_5\ \sqrt{3} & = 0 \\ 162 & +16\beta_1 & -8\beta_2 & +-8\beta_3\ \sqrt{3} & -32\beta_4 & -32\beta_5\ \sqrt{3} & = 0 \end{cases}$$

La résolution de ce système nous donne :

$$(\beta_1,\beta_2,\beta_3,\beta_4,\beta_5) \simeq (-6.12\ , 4.12\ , 0.89,\ -1.17,\ 1.01)$$

La solution de l'ERL est donc la suite de terme général :

$$u_n = 2\times 3^n - 6.12\times 2^n + 4.12\times 2^n \times \cos\frac{n\pi}{3}$$

$$+0.89\times 2^n \times \sin\frac{n\pi}{3} - 1.17n\times 2^n \times \cos\frac{n\pi}{3} + 1.01n\times 2^n \times \sin\frac{n\pi}{3}$$

Chapitre 3

Formes bilinéaires et formes quadratiques

3.1 Formes bilinéaires

3.1.1 Définition

Soit E un $\mathbb{K}$-ev de dimension n et ϕ une application de $E \times E$ dans $\mathbb{K}$. On dit que ϕ est une **forme bilinéaire sur** $\mathbb{K}$ si elle est linéaire par rapport à chacune de ses composantes.
Autrement dit, ϕ est bilinéaire si $\forall (x, y, z) \in E^3 \ \forall (\lambda, \beta) \in \mathbb{K}^2$:

$$\phi(\lambda x + \beta y, z) = \lambda\phi(x, z) + \beta\phi(y, z)$$
$$\phi(z, \lambda x + \beta y) = \lambda\phi(z, x) + \beta\phi(z, y)$$

3.1.2 Ecriture matricielle

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Comme dans le cas d'une application linéaire, on peut donner un "résumé" d'une forme bilinéaire ϕ dans la base $\mathcal{B}$ à l'aide d'une matrice.

Soient x et y deux vecteurs quelconques de E, $X = (x_1, \ldots, x_n)^t$ et $Y = (y_1, \ldots, y_n)^t$ leurs vecteurs coordonnées dans la base $\mathcal{B}$.

Alors :

$$\phi(x, y) = \phi(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \phi(e_i, e_j) = X^t A Y$$

où $A = (\phi(e_i, e_j))_{i,j=1,\dots,n}$. On dit que A est la **matrice de la forme bilinéaire ϕ dans $\mathcal{B}$**, ou plus simplement la matrice de ϕ dans $\mathcal{B}$. On notera :

$$M_{\mathcal{B}}(\phi) = A$$

Attention à ne pas confondre A avec la matrice d'une application linéaire !

Exemple 1

On se place dans $\mathbb{R}^3$ et on note $\mathcal{B}_3$ la base canonique. On note ϕ_1 l'application :

$$\phi_1 : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & x_1 y_1 + x_2 y_2 + x_3 y_3 \end{array}$$

avec $X = (x_1, \dots, x_n)^t$ et $Y = (y_1, \dots, y_n)^t$ la représentation de x et y dans la base canonique.

On peut vérifier que ϕ_1 est une forme bilinéaire. On a :

$$M_{\mathcal{B}_3}(\phi_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On peut étendre la définition de ϕ_1 à l'espace $\mathbb{R}^n$.

Exemple 2

Toujours dans $\mathbb{R}^3$ muni de sa base canonique $\mathcal{B}_3$, on note ϕ_2 l'application :

$$\phi_2 : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & x_1 y_1 - x_2 y_2 + x_3 y_3 + (x_1 y_2 + x_2 y_1) - 3(x_1 y_3 + x_3 y_1) \end{array}$$

avec les mêmes notations que précedemment.

On vérifie que ϕ_2 est une forme bilinéaire et on a :

$$M_{\mathcal{B}_3}(\phi_2) = \begin{pmatrix} 1 & 1 & -3 \\ 1 & -1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$

Exemple 3

On se place dans $\mathbb{R}^2$ et on note $\mathcal{B}_2$ sa base canonique. Soit ϕ_3 l'application :

$$\phi_3 : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & 2x_1y_1 - x_2y_1 + 3x_1y_2 + x_2y_2 \end{array}$$

avec $X = (x_1, x_2)^t$ et $Y = (y_1, y_2)^t$ la représentation de x et y dans la base canonique.

ϕ_3 est une forme bilinéaire et on a :

$$M_{\mathcal{B}_2}(\phi_3) = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}$$

3.1.3 Changement de base

On conserve les mêmes notations que dans le paragraphe précédent. On munit E d'une autre base $\mathcal{C} = (f_1, \ldots, f_n)$, dans laquelle x et y admettent pour vecteurs coordonnées $\tilde{X} = (\tilde{x_1}, \ldots, \tilde{x_n})^t$ et $\tilde{Y} = (\tilde{y_1}, \ldots, \tilde{y_n})^t$.

On a alors :

$$\begin{aligned} \phi(x, y) &= X^t A Y \text{ comme vu précedemment} \\ &= \tilde{X}^t B \tilde{Y} \text{ de façon analogue avec } \; B = (\phi(f_i, f_j))_{i,j=1,\ldots,n} \\ &= \tilde{X}^t P^t A P \tilde{Y} \text{ car } \; X = P\tilde{X} \text{ et } V = P\tilde{V} \text{ en notant } P = P_{\mathcal{BC}} \end{aligned}$$

On en déduit la formule de changement de base pour une forme bilinéaire :

$$M_{\mathcal{C}}(\phi) = P^t_{\mathcal{BC}} M_{\mathcal{B}}(\phi) P_{\mathcal{BC}}$$

Exemples

Considérons la forme bilinéaire ϕ_1, notons $\mathcal{B}_3 = (e_1, e_2, e_3)$ la base canonique et donnons nous une autre base de $\mathbb{R}^3$, par exemple $\mathcal{C}_3 = (e_1, e_1 - e_2, e_1 + e_2 + e_3)$.

On a $P_{\mathcal{B}_3\mathcal{C}_3} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ donc en vertu de ce qui précède :

$$M_{\mathcal{C}_3}(\phi_1) = (P_{\mathcal{B}_3\mathcal{C}_3})^t M_{\mathcal{B}_3}(\phi_1) P_{\mathcal{B}_3\mathcal{C}_3} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

De la même façon, on montre que :

$$M_{\mathcal{C}_3}(\phi_2) = (P_{\mathcal{B}_3\mathcal{C}_3})^t M_{\mathcal{B}_3}(\phi_2) P_{\mathcal{B}_3\mathcal{C}_3} = \begin{pmatrix} 2 & -3 & -1 \\ -3 & 0 & 0 \\ -1 & 0 & -7 \end{pmatrix}$$

En notant $\mathcal{B}_2 = (e_1, e_2)$, si l'on munit $\mathbb{R}^2$ de la base $\mathcal{C}_2 = (e_1, e_1 - e_2)$ on a :

$$M_{\mathcal{C}_2}(\phi_3) = (P_{\mathcal{B}_2\mathcal{C}_2})^t M_{\mathcal{B}_2}(\phi_3) P_{\mathcal{B}_2\mathcal{C}_2} = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$$

3.2 Formes bilinéaires symétriques et formes quadratiques

3.2.1 Définitions

Une forme bilinéaire $\phi : E \times E \to \mathbb{K}$ est dite **symétrique** si

$$\forall (x, y) \in E^2 \ \phi(x, y) = \phi(y, x)$$

On notera fbs en abrégé.

Soit ϕ une fbs sur $\mathbb{K}$. On appelle **forme quadratique associée à** ϕ l'application

$$Q : \begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ x & \mapsto & Q(x) = \phi(x, x) \end{array}$$

3.2.2 Identité de polarisation

On peut remarquer que l'on peut définir la forme quadratique associée à une forme bilinéaire non symétrique. Mais la symétrie nous permet d'avoir une correspondance bijective entre fbs et forme quadratique via l'identité de polarisation :

Propriété 3.1 *Soit ϕ une fbs sur $\mathbb{K}$ et Q la forme quadratique associée. Alors $\forall (x, y) \in E^2$:*

$$\phi(x, y) = \frac{1}{2} \left(Q(x + y) - Q(x) - Q(y) \right)$$

Démonstration 15 *Laissée en exercice*

3.2.3 Interprétation matricielle de la symétrie

Propriété 3.2 *Soit $\mathcal{B}$ une base de E et ϕ une forme bilinéaire sur $\mathbb{K}$. Alors ϕ est symétrique $\Leftrightarrow M_{\mathcal{B}}(\phi)$ est une matrice symétrique.*

Démonstration 16 *La condition nécessaire est immédiate. Etudions la réciproque. On note $\mathcal{B} = (e_1, \ldots, e_n)$.*

Soient x et y deux vecteurs quelconques de E, et $X = (x_1, \ldots, x_n)^t$ et $Y = (y_1, \ldots, y_n)^t$ leurs vecteurs coordonnées dans la base $\mathcal{B}$. Alors :

$$\phi(x, y) = X^t M_{\mathcal{B}(\phi)} Y$$

Si l'on transpose cette expression :

$$\phi(x, y)^t = Y^t M^t_{\mathcal{B}(\phi)} X = Y^t M_{\mathcal{B}(\phi)} X = \phi(y, x)$$

Comme $\phi(x, y)$ est un scalaire, il est égal à sa transposée. Par conséquent $\phi(x, y) = \phi(y, x)$ et ϕ est symétrique.

Exemples

A l'aide de la propriété précédente, il est facile de vérifier que ϕ_1 et ϕ_2 sont des formes bilinéaires symétriques et pas ϕ_3.

La forme quadratique associée à ϕ_1 est :

$$Q_1 : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ x & \mapsto & x_1^2 + x_2^2 + x_3^2 \end{array}$$

La forme quadratique associée à ϕ_2 est :

$$Q_2 : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ x & \mapsto & x_1^2 - x_2^2 + x_3^2 + 2x_1x_2 - 6x_1x_3 \end{array}$$

3.2.4 Ecriture matricielle de la forme quadratique

Soit Q une forme quadratique sur E, de fbs associée ϕ. Soit $\mathcal{B}$ une base de E et A la matrice de ϕ dans $\mathcal{B}$.
Alors si x est un vecteur quelconque de E, de vecteur coordonnées X dans $\mathcal{B}$, on a :

$$Q(x) = \phi(x, x) = X^t A X$$

Si $\mathcal{C}$ est une autre base de E, et que l'on note $\tilde{X}$ le vecteur des coordonnées de x dans $\mathcal{C}$, on peut réécrire $Q(x)$ à l'aide des formules de changement de base données précedemment :

$$Q(x) = \tilde{X}^t B \tilde{X} \text{ avec } B = P^t A P \text{ et } P = P_{\mathcal{BC}}$$

3.3 Formes bilinéaires symétriques positives et produits scalaires

A partir de maintenant, on s'intéresse uniquement au cas $\mathbb{K} = \mathbb{R}$, c'est à dire qu'on ne traite que des espaces vectoriels réels et des formes bilinéaires réelles (comme on l'a d'ailleurs fait dans tous les exemples précédents).
Dans la suite du paragraphe, E désigne donc un $\mathbb{R}$ espace vectoriel de dimension finie égale à n.

3.3.1 Définitions

Soit E un $\mathbb{R}$ espace vectoriel et ϕ une forme bilinéaire symétrique de E.
On dit que ϕ **est positive** si $\forall u \in E\ \phi(u,u) \geq 0$. Autrement dit, une fbs est positive si elle est à valeurs dans $\mathbb{R}^+$.
On dit que ϕ **est définie** si $\phi(u,u) = 0 \Rightarrow u = 0$.

Une forme bilinéaire symétrique de E définie positive est appelée **produit scalaire** de E.

3.3.2 Interprétation matricielle

Soit B une matrice symétrique de $M_n(\mathbb{R})$.
On dit que B **est positive** si $\forall U \in \mathbb{R}^n\ U^tAU \geq 0$.
On dit que B **est définie positive** si $\forall U \in \mathbb{R}^n\ U \neq 0 \Rightarrow U^tAU > 0$.

Il existe un lien direct entre la positivité au sens d'une fbs et la positivité au sens d'une matrice.

Propriété 3.3 *Soit ϕ une fbs sur E, $\mathcal{B}$ une base de E et $A = M_{\mathcal{B}}(\phi)$ la matrice de ϕ dans $\mathcal{B}$.*
ϕ est positive $\Leftrightarrow$ A est positive.
ϕ est définie positive $\Leftrightarrow$ A est définie positive.

Démonstration 17 *Elle découle de l'écriture matricielle de la quantité $\phi(u,u)$.*

Exemples

Il est clair que, au vu de l'écriture de Q_1, l'application ϕ_1 est positive et même définie positive. On l'appelle le **produit scalaire usuel de** $\mathbb{R}^3$.

Pour ϕ_2, c'est moins net et nous avons besoin d'un peu de calcul. Si $\vec{u}$ s'écrit $(u_1, u_2, u_3)^t$ dans la base canonique, on a :

$$\begin{aligned} Q_2(\vec{u}) &= u_1^2 - u_2^2 + u_3^2 + 2u_1u_2 - 6u_1u_3 \\ &= u_1^2 + 2u_1(u_2 - 3u_3) - u_2^2 + u_3^2 \\ &= (u_1 + u_2 - 3u_3)^2 - (u_2^2 + 9u_3^2 - 6u_2u_3) - u_2^2 + u_3^2 \\ &= (u_1 + u_2 - 3u_3)^2 - 2u_2^2 + 6u_2u_3 - 8u_3^2 \\ &= (u_1 + u_2 - 3u_3)^2 - 2(u_2 - \frac{3}{2}u_3)^2 + \frac{9}{3}u_3^2 - 8u_3^2 \\ &= (u_1 + u_2 - 3u_3)^2 - 2(u_2 - \frac{3}{2}u_3)^2 - \frac{7}{3}u_3^2 \end{aligned}$$

A l'aide de cette forme, on voit que ϕ_2 n'est pas positive (prendre par exemple $u_3 \neq 0$, $u_2 - \frac{3}{2}u_3 = 0$ et $u_1 + u_2 - 3u_3 = 0$).

La méthode appliquée précedemment est valable pour toute forme bilinéaire symétrique, en dimension quelconque, et s'appelle la **méthode de Gauss**. Le principe est simple : on commence par isoler tous les termes contenant u_1 et à les mettre dans un carré. On fait ensuite de même pour les termes contenant u_2, et ainsi de suite.

Toute forme quadratique associée à une forme bilinéaire symétrique peut ainsi s'écrire sous forme d'une somme de carrés pondérés par des scalaires ; **la forme quadratique sera positive si et seulement si tous ces scalaires sont positifs**.

Nous reviendrons plus loin sur le cas particulier des matrices symétriques ; nous donnons simplement ici un petit résultat qui nous servira dans la suite.

Propriété 3.4 *Soit ϕ une fbs sur E, $\mathcal{B}$ une base de E. Si ϕ est un produit scalaire, alors $M_{\mathcal{B}}(\phi)$ est inversible.*

Démonstration 18 *Notons $A = M_{\mathcal{B}}(\phi)$. Si A n'est pas inversible, $\exists X$ vecteur non nul de $\mathbb{R}^n$ tel que $AX = 0$. On a alors $X^tAX = 0$ avec $X \neq 0$, ce qui est impossible par définition d'un produit scalaire.*

On peut énoncer ce dernier résultat d'une façon équivalente : les matrices définies positives sont inversibles.

3.3.3 Deux célèbres inégalités

Propriété 3.5 *Inégalité de Cauchy-Schwarz*

Soit ϕ une fbs positive sur E. Alors

$$\forall(\vec{u},\vec{v}) \in E^2 \ |\phi(\vec{u},\vec{v})| \leq \sqrt{\phi(\vec{u},\vec{u})}\ \sqrt{\phi(\vec{v},\vec{v})}$$

Si ϕ est un produit scalaire (i.e. si ϕ est en plus définie), on a égalité $\Leftrightarrow$ $\vec{u}$ et $\vec{v}$ sont liés.

Démonstration 19 *Soit x un réel quelconque. On a $\phi(\vec{u}+x\ \vec{v},\vec{u}+x\ \vec{v}) \geq 0$ car ϕ est positive. D'autre part :*

$$\phi(\vec{u}+x\ \vec{v},\vec{u}+x\ \vec{v}) = x^2\phi(\vec{v},\vec{v}) + 2x\phi(\vec{u},\vec{v}) + \phi(\vec{u},\vec{u})$$

C'est un polynôme en x de degré 2. Comme il est toujours positif, son discriminant doit être négatif, donc :

$$4\phi(\vec{u},\vec{v})^2 - 4\phi(\vec{u},\vec{u})\phi(\vec{v},\vec{v}) \leq 0$$

$$\Rightarrow |\phi(\vec{u},\vec{v})| \leq \sqrt{\phi(\vec{u},\vec{u})}\ \sqrt{\phi(\vec{v},\vec{v})}$$

Supposons maintenant que ϕ est un produit scalaire. Alors :

$$|\phi(\vec{u},\vec{v})| = \sqrt{\phi(\vec{u},\vec{u})}\ \sqrt{\phi(\vec{v},\vec{v})}$$

$\Leftrightarrow$ *Le polynôme $x^2\phi(\vec{v},\vec{v}) + 2x\phi(\vec{u},\vec{v}) + \phi(\vec{u},\vec{u})$ admet une racine réelle*

$$\Leftrightarrow \exists x\ \phi(\vec{u}+x\ \vec{v},\vec{u}+x\ \vec{v}) = 0$$

$$\Leftrightarrow \exists x\ \vec{u}+x\ \vec{v} = 0$$

$$\Leftrightarrow \vec{u} \text{ et } \vec{v} \text{ sont liés}$$

Exemple

En appliquant ce résultat au produit scalaire usuel ϕ_1 de $\mathbb{R}^3$, on a :

$$\forall(u_1,u_2,u_3),(v_1,v_2,v_3) \in \mathbb{R}^3 \times \mathbb{R}^3$$

$$(u_1v_1+u_2v_2+u_3v_3)^2 \leq (u_1^2+u_2^2+u_3^2)(v_1^2+v_2^2+v_3^2)$$

Application

Soient $x_1, \ldots, x_n$ n réels strictement positifs. Pour p appartenant à $\mathbb{N}^*$, on pose $u_p = \frac{\sum_{i=1}^{n} x_i^p}{\sum_{i=1}^{n} x_i^{p+1}}$.
Montrer que la suite $(u_p)_{p\in\mathbb{N}^*}$ est décroissante.
En déduire qu'elle converge, et donner sa limite.

De l'inégalité de Cauchy-Schwarz, on tire une autre égalité moins fine mais qui nous sera utile lors de l'étude des espaces euclidiens.

Propriété 3.6 *Inégalité de Minkowsky*

Soit ϕ une fbs positive sur E. Alors $\forall(\vec{u},\vec{v}) \in E^2$:

$$\sqrt{\phi(\vec{u}+\vec{v},\vec{u}+\vec{v})} \leq \sqrt{\phi(\vec{u},\vec{u})} + \sqrt{\phi(\vec{v},\vec{v})}$$

Démonstration 20 *D'après l'identité de polarisation :*

$$\phi(\vec{u}+\vec{v},\vec{u}+\vec{v}) - \phi(\vec{u},\vec{u}) - \phi(\vec{v},\vec{v}) = 2\phi(\vec{u},\vec{v})$$

et d'après l'inégalité de Cauchy-Schwartz :

$$\phi(\vec{u},\vec{v}) \leq |\phi(\vec{u},\vec{v})| \leq \sqrt{\phi(\vec{u},\vec{u})}\ \sqrt{\phi(\vec{v},\vec{v})}$$

Donc :

$$\phi(\vec{u}+\vec{v},\vec{u}+\vec{v}) \leq \phi(\vec{u},\vec{u}) + \phi(\vec{v},\vec{v}) + 2\sqrt{\phi(\vec{u},\vec{u})}\ \sqrt{\phi(\vec{v},\vec{v})}$$

$$\Leftrightarrow \phi(\vec{u}+\vec{v},\vec{u}+\vec{v}) \leq (\sqrt{\phi(\vec{u},\vec{u})} + \sqrt{\phi(\vec{v},\vec{v})})^2$$

$$\Leftrightarrow \sqrt{\phi(\vec{u}+\vec{v},\vec{u}+\vec{v})} \leq \sqrt{\phi(\vec{u},\vec{u})} + \sqrt{\phi(\vec{v},\vec{v})}$$

Exemple

En appliquant ce résultat au produit scalaire usuel ϕ_1 de $\mathbb{R}^3$, on a :

$$\forall(u_1,u_2,u_3),(v_1,v_2,v_3) \in \mathbb{R}^3 \times \mathbb{R}^3$$

$$\sqrt{(u_1+v_1)^2+(u_2+v_2)^2+(u_3+v_3)^2} \leq \sqrt{u_1^2+u_2^2+u_3^2} + \sqrt{v_1^2+v_2^2+v_3^2}$$

Chapitre 4

Espaces euclidiens

4.1 Structure euclidienne

4.1.1 Définition et notations

On appelle **espace euclidien** un $\mathbb{R}$ espace vectoriel de dimension finie muni d'un produit scalaire $(.|.)$.

On suppose dans la suite de ce chapitre que E désigne un $\mathbb{R}$ espace vectoriel de dimension finie égale à n, auquel on adjoint un produit scalaire $(.|.)$.

4.1.2 Norme associée

Si Q est la forme quadratique associée à $(.|.)$, on note $\|.\|$ l'application

$$\|.\| : \begin{array}{ccc} E & \longrightarrow & \mathbb{R}^+ \\ u & \mapsto & \sqrt{Q(u)} = \sqrt{\phi(u,u)} \end{array}$$

Elle vérifie les propriétés suivantes :

Propriété 4.1
1. $\forall u \in E \ \|u\| = 0 \Rightarrow u = 0$
2. $\forall u \in E \ \forall \lambda \in \mathbb{R} \ \|\lambda \ u\| = |\lambda| \|u\|$
3. $\forall (u,v) \in E^2 \ \|u+v\| \leq \|u\| + \|v\|$

Les deux premières propriétés se vérifient simplement ; la troisième est l'inégalité de Minkowsky, que l'on a déjà rencontrée. Ces trois propriétés montrent que $\|.\|$ **est une norme**, ce qui justifie la notation.

On en déduit que tout espace euclidien est un espace vectoriel normé.

Exemple 1

On se place dans $\mathbb{R}^3$ muni de son produit scalaire usuel ϕ_1 défini dans le chapitre précédent. La norme associée à ϕ_1 est l'application

$$\|.\| : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ \vec{u} & \mapsto & \sqrt{u_1^2 + u_2^2 + u_3^2} \end{array}$$

où $U = (u_1, u_2, u_3)^t$ est le vecteur des coordonnées de $\vec{u}$ dans la base canonique $\mathcal{B}_3 = (\vec{e_1}, \vec{e_2}, \vec{e_3})$.

Il est important de noter que l'écriture de $\|\vec{u}\|$ dépend de la base de E choisie. Par exemple, dans la base $\mathcal{C}_3 = (\vec{e_1}, \vec{e_1} - \vec{e_2}, \vec{e_1} + \vec{e_2} + \vec{e_3})$ déjà considérée dans le chapitre précédent, on avait :

$$M_{\mathcal{C}_3}(\phi_1) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

Si $\tilde{U} = (\tilde{u_1}, \tilde{u_2}, \tilde{u_3})^t$ donne les coordonnées de $\vec{u}$ dans $\mathcal{C}_3$, alors $\phi_1(\vec{u}) = \tilde{U}^t M_{\mathcal{C}_3}(\phi_1)\tilde{U} = {\tilde{u_1}}^2 + 2{\tilde{u_2}}^2 + 3{\tilde{u_3}}^2 + 2\tilde{u_1}\tilde{u_2} + 2\tilde{u_1}\tilde{u_3}$ et :

$$\|\vec{u}\| = \sqrt{{\tilde{u_1}}^2 + 2{\tilde{u_2}}^2 + 3{\tilde{u_3}}^2 + 2\tilde{u_1}\tilde{u_2} + 2\tilde{u_1}\tilde{u_3}}$$

Exemple 2

Soit $E = M_n(\mathbb{R})$ l'ensemble des matrices réelles de taille n. On munit E de l'application

$$\phi : \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{R} \\ (A, B) & \mapsto & \sum_{i,j=1}^{n} a_{ij} b_{ij} \end{array}$$

On vérifie facilement que l'on définit ainsi un produit scalaire sur E. La norme associée est :

$$\|.\| : \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ A & \mapsto & \sqrt{\sum_{i,j=1}^{n} a_{ij} b_{ij}} \end{array}$$

4.1.3 Quelques formules

Propriété 4.2 *Formule d'Al Kashi*

$$\forall (u, v) \in E^2 \quad \|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2(u|v)$$

Propriété 4.3 *Identité du parallélogramme*

$$\forall (u,v) \in E^2 \quad \|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$$

La démonstration de ces deux résultats s'obtient en revenant à la définition de $\|.\|$ et en utilisant la bilinéarité de la fbs associée.

4.2 Orthogonalité

4.2.1 Définition

Soit E un espace euclidien et $(.|.)$ le produit scalaire associé. On dit que **deux vecteurs** u **et** v **de** E **sont orthogonaux** si $(u|v) = 0$. On note : $u \perp v$

Par exemple, plaçons nous dans $\mathbb{R}^3$ muni de son produit scalaire usuel. Les deux vecteurs $\vec{u}$ et $\vec{v}$, de coordonnées $(2,4,-7)^t$ et $(-3,5,2)^t$ dans la base canonique, sont orthogonaux pour le produit scalaire usuel car :

$$(\vec{u}|\vec{v}) = 2*(-3) + 4*5 + (-7)*2 = 0$$

Attention ! La notion d'orthogonalité dépend du produit scalaire considéré : deux éléments orthogonaux pour un produit scalaire ne le sont pas forcément pour un autre.

Théorème 4.4 *Théorème de Pythagore*

Soient u et v deux vecteurs de E. Alors u et v sont orthogonaux $\Leftrightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2$.

Démonstration 21 *C'est une conséquence directe de la formule d'Al Kashi.*

Soient A et B 2 parties de E. On dit que A **et** B **sont orthogonales** si

$$\forall u \in A \; \forall v \in B \; u \perp v$$

4.2.2 Familles orthogonales

Soit $\mathcal{F} = (\vec{u_1}, \ldots, \vec{u_p})$ une famille de vecteurs non nuls de E. On dit que $\mathcal{F}$ **est orthogonale** si ses éléments sont deux à deux orthogonaux.

On dit que la famille est orthonormale si elle est orthogonale et que tous ses éléments ont pour norme 1. Si on dispose d'une famille orthogonale, on l'orthonormalise facilement (il suffit de diviser chaque vecteur par sa norme). Cette remarque nous sera utile plus loin.

Propriété 4.5 *Une famille orthogonale est libre.*

Démonstration 22 *Notons $\mathcal{F} = (e_1, \ldots, e_p)$ famille orthogonale de vecteurs de E. Supposons que cette famille est liée. Alors il existe des réels $\lambda_1, \ldots, \lambda_p$ non tous nuls tels que :*

$$\lambda_1 e_1 + \ldots + \lambda_p e_p = 0$$

En prenant le produit scalaire du premier membre avec e_1, on trouve en vertu de l'orthogonalité des éléments de $\mathcal{F}$:

$$\lambda_1 (e_1|e_1) = 0$$

Comme $e_1 \neq 0$, $(e_1|e_1) > 0$ donc $\lambda_1 = 0$. On montre par un raisonnement analogue que tous les λ_i sont nuls, ce qui est contradictoire. La famille $\mathcal{F}$ est donc libre.

Théorème 4.6 *Théorème de Pythagore*
Soient $\mathcal{F} = (e_1, \ldots, e_p)$ une famille orthogonale. Alors :

$$\|e_1 + \ldots + e_p\|^2 = \|e_1\|^2 + \ldots + \|e_p\|^2$$

Démonstration 23 *e_1 est orthogonal à chacun des vecteurs $e_2, \ldots, e_p$ donc par linéarité il est orthogonal à leur somme. D'après le théorème de Pythagore pour deux éléments, on a donc :*

$$\|e_1 + \ldots + e_p\|^2 = \|e_1\|^2 + \|e_2 + \ldots + e_p\|^2$$

On conclut par récurrence.

4.2.3 Sous-espaces orthogonaux

Soit F un sev de E. On appelle orthogonal de F (que l'on note $F^\perp$) l'ensemble des éléments de E orthogonaux à tous les éléments de F. Autrement dit :

$$F^\perp = \{u \in E; \forall v \in F \ (u|v) = 0\}$$

Propriété 4.7 *Soit F un sev de E. Alors $F^\perp$ est un sev de E.*

Démonstration 24 *$\forall u \in F$ $(u|0) = 0$ donc $0 \in F^\perp$. D'autre part, soit $\lambda \in \mathbb{R}$ et $(u, v) \in (F^\perp)^2$. Soit w quelconque appartenant à F. Alors :*

$$(w|u + \lambda v) = (w|u) + \lambda(w|v) = 0 \text{ car } (w|u) = 0 \text{ et } (w|v) = 0$$

Donc, $u + \lambda v \in F^\perp$. On en déduit que $F^\perp$ est bien un sev de E.

Propriété 4.8 *Soit F un sev de E, p sa dimension et $\mathcal{F} = (e_1, \ldots, e_p)$ une base de F.*
Soit $u \in E$. Alors $u \in F^\perp \Leftrightarrow u$ est orthogonal à chaque élément de $\mathcal{F}$.

Démonstration 25 *La condition nécessaire est évidente. Réciproquement, supposons que u est orthogonal à chaque élément de $\mathcal{F}$. Soit $v \in F$; comme $\mathcal{F}$ est une base de F, v s'écrit comme une combinaison linéaire des éléments de $\mathcal{F}$:*

$$\exists (v_1, \ldots, v_p) \in \mathbb{R}^p \; v = v_1 e_1 + \ldots + v_p e_p$$

Alors :

$$\begin{aligned}(u|v) &= (u|v_1 e_1 + \ldots + v_p e_p) \\ &= v_1(u|e_1) + \ldots + v_p(u|e_p) = 0\end{aligned}$$

Définition-Propriété 4.1 *Soit F un sev de E. $F^\perp$ est un supplémentaire de F de E, appelé supplémentaire orthogonal de F dans E.*

Démonstration 26 *Soit $u \in F \cap F^\perp$. En tant qu'élément de $F^\perp$, u est orthogonal à tous les éléments de F, donc à lui-même. On a donc :*

$$(u|u) = \|u\|^2 = 0 \Rightarrow u = 0$$

Par suite, $F \cap F^\perp = \{0\}$ et F et $F^\perp$ sont en somme directe.

Notons p la dimension de F et $(e_1, \ldots, e_p)$ une base de F (pas nécessairement orthogonale). On la complète en une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E. Soit A la matrice du produit scalaire dans cette base.
Soit u un vecteur quelconque de E ; on note $U = (u_1, \ldots, u_n)^t$ le vecteur de ses coordonnées dans $\mathcal{B}$. Alors, en utilisant la propriété précédente :

$$\begin{aligned}&u \in F^\perp \\ &\Leftrightarrow \forall i = 1 \ldots p \; (u|e_i) = 0 \\ &\Leftrightarrow \forall i = 1 \ldots p \; u_1(e_i|e_1) + \ldots + u_n(e_i|e_n) = 0\end{aligned}$$

$$\Leftrightarrow A_p U = 0$$

où A_p est la matrice constituée des p premières lignes de A. Comme A est inversible, A_p est de rang p. On en déduit que $\{U\ ;\ A_p U = 0\}$ est de dimension $n - p$, donc que $F^\perp$ est de dimension $n - p$.
En résumé, on a donc montré que :

$$F \oplus F^\perp$$

$$dim(F) + dim(F^\perp) = n = dim(E)$$

Donc, F et $F^\perp$ sont supplémentaires dans E.

Exemple

Soit $E = \mathbb{R}^2$ et $\mathcal{B} = (\vec{e_1}, \vec{e_2})$ sa base canonique. On munit E du produit scalaire usuel ϕ_1, défini par

$$M_{\mathcal{B}}(\phi_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

et du produit scalaire ψ défini par

$$M_{\mathcal{B}}(\psi) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

$M_{\mathcal{B}}(\psi)$ est une matrice symétrique, donc ψ définit bien une fbs. A l'aide de la méthode de Gauss :

$$(x, y) M_{\mathcal{B}}(\psi) (x, y)^t$$
$$= x^2 - 2xy + 2y^2 = (x - y)^2 + y^2$$

Donc $(x, y) M_{\mathcal{B}}(\psi) (x, y)^t \geq 0$ avec égalité $\Leftrightarrow x = y = 0$, et ψ définit bien un produit scalaire.

Cherchons un vecteur $\vec{v}$ orthogonal à $\vec{e_1}$ pour le produit scalaire usuel. Si on note $(a, b)^t$ les coordonnées de ce vecteur dans $\mathcal{B}$, on doit avoir :

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0$$

$$\Leftrightarrow a = 0$$

On peut donc prendre par exemple $\vec{v} = e_2$.

Cherchons maintenant un vecteur $\vec{v}$ orthogonal à $\vec{e_1}$ pour le produit scalaire ψ. Si on note $(a,b)^t$ les coordonnées de ce vecteur dans $\mathcal{B}$, on doit avoir :

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0$$
$$\Leftrightarrow a - b = 0$$

On peut donc prendre par exemple $\vec{v} = (1,1)$ dans la base $\mathcal{B}$.

4.3 Existence d'une base orthonormée

4.3.1 Existence : le théorème de Gram-Schmidt

Le résultat qui suit est très important, aussi bien sur le plan théorique (nous verrons ensuite des propriétés qui en découlent) que pratique pour la construction d'une base orthonormée.

Théorème 4.9 *Tout espace euclidien E admet une base orthonormée.*

Démonstration 27 *E est un espace euclidien, donc de dimension finie n. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base quelconque de E. Nous allons commencer par définir une base orthogonale $\mathcal{C} = (f_1, \ldots, f_n)$ selon le procédé de Schmidt ; il est ensuite facile de la transformer en une base orthonormale.*
On commence par prendre $f_1 = e_1$.
On cherche ensuite f_2 sous la forme $f_2 = e_2 + \alpha f_1$. On doit avoir $(f_1|f_2) = 0$ ce qui impose :

$$\alpha = -\frac{(f_1|e_2)}{\|f_1\|^2}$$

On va poursuivre la construction. Supposons que l'on ait construits les $p+1$ premiers vecteurs $(f_1, f_2, \ldots, f_p)$ d'une base orthogonale. On cherche le $p+1$ème vecteur sous la forme :

$$f_{p+1} = e_{p+1} + \alpha_1 f_1 + \ldots + \alpha_p f_p$$

Comme les vecteurs $f_1 \ \ldots \ f_{p+1}$ sont mutuellement orthogonaux, on obtient facilement le système d'équations :

$$\begin{cases} 0 = (e_{p+1}|f_1) + \alpha_1 (f_1|f_1) \\ \qquad \ldots \\ 0 = (e_{p+1}|f_p) + \alpha_p (f_p|f_p) \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = -\frac{(e_{p+1}|f_1)}{(f_1|f_1)} \\ \qquad \ldots \\ \alpha_p = -\frac{(e_{p+1}|f_p)}{(f_p|f_p)} \end{cases}$$

On construit ainsi par récurrence une base orthogonale de E. Comme on l'a vu précedemment, on peut ensuite facilement la transformer en une base orthonormale.

4.3.2 Exemple

Soit E un espace euclidien de dimension 3 muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ quelconque. On note ψ le produit scalaire correspondant, défini dans la base $\mathcal{B}$ par

$$M_{\mathcal{B}}(\psi) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

On vérifie facilement que ψ est bien un produit scalaire ; on voit également que $\mathcal{B}$ n'est pas une base orthonormée pour ψ. On va en construire une en orthnormalisant $\mathcal{B}$ selon le procédé de Gram-Schmidt.

On commence par prendre $f_1 = e_1$.

On pose ensuite $f_2 = e_2 + \alpha f_1$. Comme on veut imposer $\psi(f_1, f_2) = 0$, on a :

$$0 = \psi(f_1, e_2) + \alpha\psi(f_1, f_1) = \psi(e_1, e_2) + \alpha\psi(e_1, e_1)$$

A l'aide de la forme matricielle de ψ, on a :

$$\psi(e_1, e_1) = (1, 0, 0) \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1$$

$$\psi(e_1, e_2) = (1, 0, 0) \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = -1$$

Donc, $0 = -1 + \alpha \Rightarrow \alpha = 1$ et on prend $f_2 = (1, 1, 0)^t$.

On pose enfin $f_3 = e_3 + \alpha f_1 + \beta f_2$. Comme on veut imposer $\psi(f_3, f_1) = 0$ et $\psi(f_3, f_2) = 0$, on a :

$$0 = \psi(f_1, e_3) + \alpha\psi(f_1, f_1) \text{ et } 0 = \psi(f_2, e_3) + \beta\psi(f_2, f_2)$$

$$\Rightarrow 0 = \psi(e_1, e_3) + \alpha\psi(e_1, e_1) \text{ et } 0 = \psi(f_2, e_3) + \beta\psi(f_2, f_2)$$

En utilisant là encore la forme matricielle de ψ, on trouve $\psi(e_1, e_3) = 0$, $\psi(f_2, e_3) = -1$ et $\psi(f_2, f_2) = 1$. On obtient donc :

$$0 = \alpha \text{ et } 0 = -1 + \beta$$

On prend donc $f_3 = (1, 1, 1)^t$.

En normalisant les vecteurs, on constate qu'ils le sont déjà. On peut donc prendre comme base orthonormée (f_1, f_2, f_3).

4.3.3 Ecriture d'un produit scalaire dans une base orthonormée

Soit E un espace euclidien muni d'une base orthonormée $\mathcal{B}$. Si on note ϕ le produit scalaire associé à E, on a de façon immédiate :

$$M_{\mathcal{B}}(\phi) = I$$

Autrement dit, la matrice d'un produit scalaire est la même quelle que soit la base orthonormée considérée. On voit donc l'un des avantages de travailler avec des bases orthonormées : on n'a pas à se soucier des formules de changement de base.

Soit $\mathcal{C}$ une autre base orthonormée ; on a donc $M_{\mathcal{C}}(\phi) = I$ également. Les formules de changement de base impliquent donc que $(P_{\mathcal{BC}})^t P_{\mathcal{BC}} = I$. On dit que $P_{\mathcal{BC}}$ est une matrice orthogonale.

Définition-Propriété 4.2 *Soit $A \in M_n(\mathbb{R})$. On dit que A est orthogonale si et seulement si*

$$A^t A = I$$

Soit E un espace euclidien. Les matrices de passage entre bases orthonormées sont orthogonales.

4.3.4 Quelques résultats supplémentaires

Propriété 4.10 *Soit F un sev de E. Alors F admet une base orthonormale.*

Démonstration 28 *F est lui-même un espace euclidien, donc on peut lui appliquer le théorême de Gram-Schmidt.*

Propriété 4.11 *Soit $\mathcal{F}$ une famille orthonormale d'éléments de E. Alors $\mathcal{F}$ peut être complétée en une base orthonormale de E.*

Démonstration 29 *On note F le sev de E engendré par $\mathcal{F}$. Son orthogonal admet une base orthonormale, d'après ce qui précède. En réunissant les deux familles, on obtient une base orthonormale de E.*

Chapitre 5

Projections

Dans ce chapitre, E désigne un espace euclidien de dimension finie égale à n.

5.1 Quelques rappels sur les projections

5.1.1 Définition et premières propriétés

Soient F et G deux sous-espaces supplémentaires dans E. On sait alors que pour tout élément $\vec{u}$ de E il existe une décomposition unique de $\vec{u}$ sous la forme

$$\vec{u} = \vec{v} + \vec{w} \text{ avec } \vec{v} \in F \text{ et } \vec{w} \in G$$

L'application qui à un vecteur $\vec{u}$ de E associe le vecteur $\vec{v}$ défini comme précedemment est appelée **projection (ou projecteur) sur F parallèlement à G**. Soit p cette application. C'est un endomorphisme de E, qui vérifie les propriétés suivantes :

- $p \circ p = p$: on dit que p est idempotente
- $Ker(p) = Im(Id - p) = G$
- $Im(p) = Ker(Id - p) = F$

Réciproquement, soit p un endomorphisme de E tel que $p \circ p = p$. Alors p est le projecteur sur $Im(p)$ parallèlement à $Ker(p)$. On a donc en particulier :

$$Ker(p) \oplus Im(p) = E$$

Notons que toute projection est diagonalisable. En effet, p annule le polynôme $X^2 - X$, scindé à racines simples.

5.1.2 Représentation matricielle

Soient F et G deux sous-espaces supplémentaires dans E, p le projecteur sur F parallèlement à G.

Compte tenu de ce qui précède, 0 et 1 sont les seules valeurs propres possibles pour p, et $Ker(p)$ et $Im(p)$ sont les sous-espaces propres qui leur sont respectivement associés.

Soit $(\overrightarrow{e_1}, \ldots, \overrightarrow{e_q})$ une base de $Ker(p)$ et $(\overrightarrow{e_{q+1}}, \ldots, \overrightarrow{e_n})$ une base de $Im(p)$.$\mathcal{B} = (\overrightarrow{e_1}, \ldots, \overrightarrow{e_n})$ est une base de E, et dans cette base :

$$M_{\mathcal{B}}(p) = \begin{pmatrix} 0 & \ldots & \ldots & 0 & 0 & \ldots & \ldots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0_q & 0_{n-q} \\ 0_{n-q} & I_{n-q} \end{pmatrix}$$

0_k désignant la matrice carrée nulle de taille k et I_k la matrice identité de taille k.

On peut remarquer que pour un projecteur, le rang est égal à la trace. En effet, dans une base de vecteurs propres, le nombre de termes non nuls de la diagonale est égal à la dimension de $Im(p)$, et chacun de ces termes vaut 1. Par conséquent $tr(p) = dim(Im(p)) \times 1 = rg(p)$.

Exemple

Soit $E = \mathbb{R}^3$. On pose $F = Vect((1,1,0)^t)$. On peut facilement trouver un supplémentaire de F, par exemple en prenant $G = Vect((1,0,0)^t, (0,1,1)^t)$.

On définit p_1 comme étant le projecteur sur F parallèlement à G. Si on note $f_1 = (1,1,0)^t$, $f_2 = (1,0,0)^t$ et $f_3 = (0,1,1)^t$, alors $\mathcal{C} = (\vec{f_1}, \vec{f_2}, \vec{f_3})$ est une base de E et dans cette base :

$$M_{\mathcal{C}}(p_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

On peut vérifier qu'en particulier, on a bien $(M_{\mathcal{C}}(p_1))^2 = M_{\mathcal{C}}(p_1)$.

Notons $\mathcal{B}$ la base canonique ; on a trivialement $P_{\mathcal{BC}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ et on peut montrer facilement que $P_{\mathcal{CB}} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. On en déduit la matrice de p dans la base canonique :

$$M_{\mathcal{B}}(p_1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

5.1.3 Le cas des symétries

Soient F et G deux sous-espaces supplémentaires dans E, p le projecteur sur F parallèlement à G.

Alors l'endomorphisme $s = 2p - Id$ est appelé symétrie par rapport à F parallèlement à G. s vérifie les propriétés suivantes :

- $s \circ s = Id$
- $Ker(Id - s) = Ker(Id - p) = F$
- $Ker(Id + s) = Ker(p) = G$

Réciproquement, soit s un endomorphisme de E tel que $s \circ s = Id$. Alors s est la symétrie sur $Ker(Id - s)$ parallèlement à $Ker(Id + s)$. On a donc en particulier :

$$Ker(Id - s) \oplus Ker(Id + s) = E$$

E est somme directe des sous-espaces propres de s, donc s est diagonalisable. C'est aussi une conséquence du théorême de Shreier (s annule le polynôme (X-1)(X+1)).

Nous n'en dirons pas plus sur les symétries ; pour les étudier, il suffit après tout de se ramener à un projecteur à l'aide de la définition.

5.2 Projections orthogonales

5.2.1 Définition

Soit F un sev de E. Alors le **projecteur sur F parallèlement à $F^\perp$** est appelé **projecteur orthogonal de E sur F**. Il est défini de façon unique

car le supplémentaire orthogonal de F est unique.

Le projecteur orthogonal a toutes les propriétés d'un projecteur. $Im(p)$ et $Ker(p)$ sont supplémentaires orthogonaux.

Réciproquement, soit p un endomorphisme de E tel que $p \circ p = p$ d'une part, $Im(p)$ et $Ker(p)$ sont orthogonaux d'autre part. Alors p est le projecteur orthogonal de E sur F.

5.2.2 Propriétés d'un projecteur orthogonal

Propriété 5.1 *Soit p un projecteur orthogonal de E. Alors :*

1. $\forall \vec{u} \in E \ (\vec{u}|p(\vec{u})) = \|p(\vec{u})\|^2$
2. $\forall (\vec{u}, \vec{v}) \in E^2 \ (p(\vec{u})|\vec{v}) = (\vec{u}|p(\vec{v}))$
3. $\forall \vec{u} \in E \ \|p(\vec{u})\| \leq \|\vec{u}\|$

Démonstration 30 *1 : Soit $\vec{u} \in E$. Comme p est un projecteur orthogonal, on a : $\vec{u} = p(\vec{u}) + (\vec{u} - p(\vec{u}))$ avec $\vec{u}$ et $\vec{u} - p(\vec{u})$ orthogonaux. Par conséquent :*

$$(\vec{u}|\vec{u} - p(\vec{u})) = 0 \Rightarrow (\vec{u}|\vec{u}) - (\vec{u}|p(\vec{u})) = 0$$

$$\Rightarrow (\vec{u}|p(\vec{u})) = \|p(\vec{u})\|^2$$

2 : Soient $\vec{u}$ et $\vec{v}$ dans E. Alors :

$$\begin{aligned}
(p(\vec{u})|\vec{v}) &= (p(\vec{u})|\vec{v} - p(\vec{v}) + p(\vec{v})) \\
&= (p(\vec{u})|\vec{v} - p(\vec{v})) + (p(\vec{u})|p(\vec{v})) \\
&= (p(\vec{u})|\vec{v} - p(\vec{v}) + p(\vec{v})) \\
&= (p(\vec{u})|p(\vec{v})) \ \textit{car } p(\vec{u}) \in Im(p) \ \textit{et } \vec{v} - p(\vec{v}) \in Im(p)^{\perp} = Ker(p) \\
&= (\vec{u} - (\vec{u} - p(\vec{u}))|p(\vec{v})) \\
&= (\vec{u}|p(\vec{v})) \ \textit{car } p(\vec{v}) \in Im(p) \ \textit{et } \vec{u} - p(\vec{u}) \in Im(p)^{\perp} = Ker(p)
\end{aligned}$$

3 : Soit $\vec{u} \in E$. On a vu que $p(\vec{u})$ et $\vec{u} - p(\vec{u})$ étaient orthogonaux, donc par le théorème de Pythagore :

$$\|p(\vec{u})\|^2 + \|\vec{u} - p(\vec{u})\|^2 = \|\vec{u}\|^2$$

$$\Rightarrow \|p(\vec{u})\|^2 \leq \|\vec{u}\|^2$$

La propriété suivante est le résultat fondamental utilisé pour la régression linéaire simple.

Théorème 5.2 *Soit $\vec{u} \in E$ et F un sev de E. On note p le projecteur orthogonal de E sur F. Alors $p(\vec{u})$ est le seul élément de F tel que :*

$$\|\vec{u} - p(\vec{u})\| = Min_{\vec{v}\in F}\|\vec{u} - \vec{v}\|$$

Il vérifie de plus :

$$\|\vec{u} - p(\vec{u})\|^2 + \|p(\vec{u})\|^2 = \|\vec{u}\|^2$$

Démonstration 31 *Soit $\vec{v}$ quelconque appartenant à F. Alors :*

$$\vec{u} - \vec{v} = \vec{u} - p(\vec{u}) + p(\vec{u}) - \vec{v}$$

$\vec{u} - p(\vec{u})$ appartient à $Im(Id - p) = Ker(p)$ et $p(\vec{u}) - \vec{v}$ appartient à $F = Im(p)$ car $p(\vec{u})$ et $\vec{v}$ appartiennent à F. $\vec{u} - p(\vec{u})$ et $p(\vec{u}) - \vec{v}$ sont donc orthogonaux et par le théorème de Pythagore :

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u} - p(\vec{u})\|^2 + \|p(\vec{u}) - \vec{v}\|^2$$

$$\Rightarrow \|\vec{u} - p(\vec{u})\|^2 \leq \|\vec{u} - \vec{v}\|^2$$

Par conséquent :

$$\|\vec{u} - p(\vec{u})\| = Min_{\vec{v}\in F}\|\vec{u} - \vec{v}\|$$

Supposons qu'il existe un autre $\vec{w} \in F$, différent de $p(\vec{u})$, tel que $\|\vec{u} - \vec{w}\| = Min_{\vec{v}\in F}\|\vec{u} - \vec{v}\| = \|\vec{u} - p(\vec{u})\|$. Alors comme $\vec{w}$ est un élément de F :

$$\|\vec{u} - \vec{w}\|^2 = \|\vec{u} - p(\vec{u})\|^2 + \|p(\vec{u}) - \vec{w}\|^2$$

$$\Rightarrow \|\vec{u} - \vec{w}\|^2 - \|\vec{u} - p(\vec{u})\|^2 = 0$$
$$= \|p(\vec{u}) - \vec{w}\|^2 > 0 \ car\ p(\vec{u}) \neq \vec{w}$$

On a donc contradiction ; d'où l'unicité.

Le dernier point découle du théorème de Pythagore.

Exemple

Reprenons l'exemple de la partie précédente, avec $E = \mathbb{R}^3$ muni de son produit scalaire usuel et $F = Vect((1, 1, 0)^t)$. On note $\vec{g_1} = (1, 1, 0)^t$.

On peut facilement produire le supplémentaire orthogonal de F, par exemple en orthogonalisant la base $\mathcal{C}$ selon le procédé de Gram-Schmidt. On prend ici comme base de $F^\perp$ les vecteurs $(0, 0, 1)^t$ et $(1, -1, 0)^t$, notés respectivement $\vec{g_2}$ et $\vec{g_3}$. Comme en particulier $F^\perp$ est un supplémentaire de F, $\mathcal{D} = (\vec{g_1}, \vec{g_2}, \vec{g_3})$

est une base orthogonale de E.

Soit p_2 le projecteur orthogonal sur F. Alors dans la base $\mathcal{D}$:

$$M_{\mathcal{D}}(p_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Nous allons vérifier les différentes propriétés d'un projecteur orthogonal sur cet exemple ; pour simplifier, nous allons revenir à la base canonique. On a de façon évidente : $P_{\mathcal{BD}} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ et on peut montrer facilement que $P_{\mathcal{CB}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$. On en déduit la matrice de p dans la base canonique :

$$M_{\mathcal{B}}(p_2) = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Soit $\vec{u}$ un vecteur quelconque de E, de coordonnées $(x, y, z)^t$ dans la base canonique. Alors $p_2(\vec{u})$ a pour coordonnées $\frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x+y \\ x+y \\ 0 \end{pmatrix}$.

On a :

- $\frac{1}{2}(p_2(\vec{u})|\vec{u}) = (x, y, z)^t \begin{pmatrix} x+y \\ x+y \\ 0 \end{pmatrix} = \frac{1}{2}(x^2 + xy) + \frac{1}{2}(y^2 + xy) = \frac{1}{2}(x+y)^2$ et $\|p_2(\vec{u})\|^2 = \frac{1}{2}(x+y)^2$, donc : $(p_2(\vec{u})|\vec{u}) = \|p_2(\vec{u})\|^2$

- $\|\vec{u}\|^2 = x^2 + y^2 + z^2 \geq x^2 + y^2 \geq \frac{1}{2}(x^2 + y^2) + xy$ car $x^2 + y^2 \geq 2xy$ (pourquoi ?). Donc, $\|\vec{u}\|^2 \geq \frac{1}{2}(x+y)^2 = \|p_2(\vec{u})\|^2$

- Soit $\vec{v}$ un vecteur quelconque de E, de coordonnées $(a, b, c)^t$ dans la base canonique. Alors $p_2(\vec{v})$ a pour coordonnées dans la base canonique $\frac{1}{2} \begin{pmatrix} a+b \\ a+b \\ 0 \end{pmatrix}$ d'après ce qui précède. On a $(p_2(\vec{u})|\vec{v}) = \frac{1}{2}(x+y, x+y, 0) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{1}{2}(x+y)(a+b)$ et par symétrie, on a également $(p_2(\vec{v})|\vec{u}) = \frac{1}{2}(x+y)(a+b)$. Donc, $(p_2(\vec{u})|\vec{v}) = (p_2(\vec{v})|\vec{u})$

Tous les points de la propriété 1.1 sont donc vérifiés.

Avec les mêmes notations que précédemment, on a d'autre part :

$$\|\vec{u} - p_2(\vec{u})\|^2 = \left\| \begin{pmatrix} \frac{1}{2}(x-y) \\ \frac{1}{2}(y-x) \\ z \end{pmatrix} \right\|^2 = \frac{1}{2}(x-y)^2 + z^2$$

Soit $\vec{v}$ un vecteur quelconque de F, de coordonnées $(a, a, 0)^t$ dans la base canonique. On a :

$$\|\vec{u} - \vec{v}\|^2 = (x-a)^2 + (y-a)^2 + z^2$$

Alors :

$$(x-y)^2 = (x-a+a-y)^2 = (x-a)^2 + (y-a)^2 - 2(x-a)(y-a)$$

$$\leq (x-a)^2 + (y-a)^2 + (x-a)^2 + (y-a)^2$$

$$\Rightarrow \frac{1}{2}(x-y)^2 \leq (x-a)^2 + (y-a)^2$$

On en déduit que pour tout vecteur $\vec{v}$ de F, $\|\vec{u} - p_2(\vec{u})\|^2 \leq \|\vec{u} - \vec{v}\|^2$ et donc que : $\|\vec{u} - p(\vec{u})\| = Min_{\vec{v} \in F} \|\vec{u} - \vec{v}\|$.

Remarquons enfin que $\|\vec{u} - p_2(\vec{u})\|^2 = \frac{1}{2}(x-y)^2 + z^2$ et $\|p_2(\vec{u})\|^2 = \frac{1}{2}(x+y)^2$, donc :

$$\|\vec{u} - p_2(\vec{u})\|^2 + \|p_2(\vec{u})\|^2 = x^2 + y^2 + z^2 = \|\vec{u}\|^2$$

et les vérifications sont terminées.

5.3 Application à la régression linéaire simple

La régression linéaire simple consiste, grosso modo, à expliquer une variable à l'aide d'autres variables ; plus exactement, à trouver la combinaison linéaire de variables la plus proche (dans un sens qui reste à préciser) d'une variable fixée.

Le cadre est généralement le suivant : on dispose d'un échantillon de n individus (au sens statistique du terme) sur lesquels on a relevé les valeurs de variables $y, x_1, \ldots, x_p$. On cherche approcher la variable y par la meilleure combinaison linéaire possible des variables $x_1, \ldots, x_p$.

Les valeurs relevés sont notées sous forme d'une matrice colonne $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ pour la variable à expliquer y et sous forme d'une matrice $X = \begin{pmatrix} x_1^1 & \dots & x_1^p \\ \vdots & \dots & \vdots \\ x_n^1 & \dots & x_n^p \end{pmatrix}$ pour les variables explicatives ; X_i, la ième colonne de X, donnant les valeurs prises par la variable x_i sur l'échantillon. Le problème est donc de trouver un vecteur $\vec{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}$ tel que $X\vec{\alpha}$ soit le plus proche possible de Y. La proximité est ici à prendre au sens de la norme associée au produit scalaire usuel de $\mathbb{R}^n$. On va donc chercher $\vec{\alpha}$ qui minimise :

$$\|Y - X\vec{\alpha}\|^2 = \sum_{i=1}^{n}(y_i - \sum_{j=1}^{p} \alpha_j x_i^j)^2$$

On peut résoudre ce problème numériquement, en dérivant l'expression précédente en les $\alpha_1, \dots, \alpha_p$ ce qui nous donnerait des équations aboutissant à une forme explicite pour $\vec{\alpha}$.
Nous allons plutôt adopter une approche géométrique. Le problème est de trouver des coefficients $\alpha_1, \dots, \alpha_p$ tels que $\|Y - \alpha_1 X_1 - \dots \alpha_p X_p\|^2$ soit minimal ; autrement dit, on cherche le projeté orthogonal de Y sur les colonnes de X. Soit $p(Y) = \hat{\vec{\alpha}}X$ ce projeté : $Y - p(Y)$ est orthogonal à l'espace sur lequel on projette, donc aux colonnes de X. On en déduit que :

$$X_1^t(Y - \hat{\vec{\alpha}}X) = 0$$

$$\dots$$

$$X_p^t(Y - \hat{\vec{\alpha}}X) = 0$$

Si on synthétise toutes ces relations, on a donc $X^t(Y - \hat{\vec{\alpha}}X) = 0$. Si on suppose que X est de plein rang, X^tX est inversible et on a finalement $\hat{\vec{\alpha}} = (X^tX)^{-1}X^tY$.

On peut également résoudre ce problème pour un produit scalaire ϕ quelconque. Notons M la matrice de ϕ dans la base canonique. Dans ce cas on cherche le projeté orthogonal de Y (au sens de ϕ) sur les colonnes de X, ce qui nous donne les équations :

$$X_1^t M(Y - \hat{\vec{\alpha}}X) = 0$$

$$\dots$$

$$X_p^t M(Y - \hat{\vec{\alpha}} X) = 0$$

que l'on peut synthétiser en $X^t M(Y - \hat{\vec{\alpha}} X) = 0$, ce qui donne finalement $\hat{\vec{\alpha}} = (X^t M X)^{-1} X^t M Y$.