ONZE LEÇONS D'ALGÈBRE LINÉAIRE POUR LA LICENCE DE MATHÉMATIQUES

Laurent Serlet

Version 2004-2005

Table des matières

Chapitre 1

Rappels sur les objets de base

Ce cours suppose connues les notions élémentaires de l'algèbre linéaire : notion d'espace vectoriel, famille libre et génératrice, base, sous-espace, somme de sous-espaces, sous-espaces supplémentaires, application linéaire, matrice, déterminant, système linéaire .... Il serait illusoire d'essayer d'aller plus loin s'il y a doute sur ces notions de base. Dans ce cas il faut se reporter aux ouvrages d'introduction. Néanmoins dans ce chapitre nous commençons par rappeler briévement certaines de ces notions élémentaires et, sans faire de preuve, certains résultats afin de fixer les notations et de redessiner succintement le cadre de notre étude.

La plus grande partie de ce cours se restreindra aux espaces de dimension finie, traitant ainsi l'algèbre matricielle. Toutefois certaines notions pourront être introduites dans le cas général. En particulier les produits scalaires sur des espaces de fonctions de dimension infinie sont d'une grande importance en analyse.

1.1 Matrices et applications linéaires

Les espaces avec lesquels nous allons travailler sont des espaces vectoriels sur un corps $\mathbb{K}$. Pour fixer les idées nous penserons que $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$ et nous appelerons $\mathbb{K}$ corps des scalaires. Les espaces vectoriels intervenant dans le présent chapitre seront tous supposés de dimension finie et il sera de même dans les chapitres suivants sauf indication contraire. Par choix d'une base il revient au même de travailler uniquement avec les espaces $\mathbb{K}^d$.

On notera $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à coefficients dans $\mathbb{K}$ ayant

n lignes et p colonnes et simplement $\mathcal{M}_n(\mathbb{K})$ pour l'ensemble des matrices carrées $\mathcal{M}_{n,n}(\mathbb{K})$. On notera $L(E,F)$ l'ensemble des applications linéaires de l'espace vectoriel E dans l'espace vectoriel F.

Une application linéaire u d'un espace vectoriel E dans un espace vectoriel F peut être représentée, une fois des bases (e_i) et (f_j) choisies respectivement dans E et F, par une matrice $A = (a_{ij})$ ayant $(\dim E)$ colonnes et $(\dim F)$ lignes et dont le terme général a_{ij} est fourni par :

$$\forall j \in \{1, \ldots, \dim E\}, \quad u(e_j) = \sum_{i=1}^{\dim F} a_{ij}\ f_i$$

On note alors $A = \mathrm{mat}_{(e_i),(f_j)}\ u$. On dispose des opérations addition et multiplication par un scalaire sur les matrices et les applications linéaires. Pour deux applications linéaires u et v la somme est simplement définie par $(u+v)(x) = u(x) + v(x)$ et pour λ scalaire, on pose $(\lambda\, u)(x) = \lambda\ u(x)$. Pour les matrices, la somme consiste à faire la somme des termes généraux ; pour le produit d'une matrice par un scalaire on multiplie le terme général par ce scalaire. L'ensemble des applications linéaires entre deux espaces fixés et identiquement l'ensemble des matrices de taille fixée, avec les opérations somme et produit par un scalaire décrites précédemment, forment deux espaces vectoriels. Les définitions d'opérations sur les applications linéaires d'une part et sur les matrices d'autre part sont cohérentes car, des bases étant fixées, la matrice de la somme de deux applications linéaires est la somme des matrices de ces deux applications linéaires et une propriété similaire vaut pour le produit avec un scalaire. En termes choisis, l'application de $L(E,F)$ dans $\mathcal{M}_{\dim F, \dim E}(\mathbb{K})$ qui, à une application linéaire associe sa matrice, dans des bases fixées, est un isomorphisme d'espaces vectoriels.

On dispose aussi de la composition des applications linéaires. Si E, F et G sont trois $\mathbb{K}$–espaces vectoriels, si u et v sont deux applications linéaires, respectivement de E dans F et de F dans G, alors on peut définir la composée $v \circ u$ par $(v \circ u)(x) = v(u(x))$ pour tout $x \in E$. Cette composée est une application linéaire. Pour deux matrices A et B, on peut définir le produit AB quand le nombre de colonnes de A est égal au nombre de lignes de B. Le coefficient général $(AB)_{ij}$ de la matrice AB s'obtient en fonction de ceux de A et B par

$$(AB)_{ij} = \sum_k A_{ik}\ B_{kj}$$

Produit de matrices et composé d'applications linéaires se correspondent de la facon suivante : si (e_i), (f_j), (g_k) sont des bases de E, F, G respectivement,

on a
$$\mathrm{mat}_{(e_i),(g_k)} v \circ u = \left(\mathrm{mat}_{(f_j),(g_k)} v\right) \left(\mathrm{mat}_{(e_i),(f_j)} u\right)$$

La *transposée* d'une matrice $A = (a_{ij})_{1\leq i\leq n,\ 1\leq j\leq p}$ à n lignes et p colonnes est la matrice à p lignes et n colonnes $B = (b_{ij})_{1\leq i\leq p,\ 1\leq j\leq n}$ donnée par $b_{ij} = a_{ji}$. Dans ce cours elle sera notée ${}^tA$. On trouve parfois la notation A' dans les cours de statistique.

Un cas particulier important d'application linéaire est celui des *endomorphismes* qui sont des applications linéaires d'un espace vectoriel dans lui-même. L'ensemble des endomorphismes de l'espace vectoriel E sera noté $L(E)$ plutôt que $L(E, E)$. La matrice d'un endomorphisme est une matrice carrée.

Parmi les matrices carrées, certaines sont inversibles ; ce sont les matrices A telles qu'il existe une matrice notée A^{-1} vérifiant $A\,A^{-1} = I$. Dans cette égalité I est la matrice unité–on dit parfois matrice identité– qui est précisement la matrice de l'endomorphisme identité. Dans le cours d'algèbre générale on dira que l'ensemble des matrices inversibles de taille $n \times n$ fixée à coefficients dans $\mathbb{K}$, muni de la multiplication matricielle, est un groupe (non commutatif) appelé *groupe linéaire*. On notera cet ensemble $\mathcal{GL}_n(\mathbb{K})$.

Les puissances de composition d'un endomorphisme sont bien définies : $u^k = u \circ u \circ \cdots \circ u$. La matrice de u^k dans une base donnée est la puissance k-ième de la matrice de u, dans la base considérée. Cette remarque sera prolongée plus loin pour définir la notion de polynome d'endomorphisme.

Supposons que, dans un espace E donné on dispose de deux bases (e_i) et (h_i). Un vecteur $x \in E$ se décompose, de manière unique, dans la base (e_i) et dans la base (h_i). Disons

$$x = \sum_i x_i\, e_i = \sum_i y_i\, h_i$$

Notons X la matrice colonne dont les composantes sont les x_i et Y la matrice colonne dont les composantes sont les y_i. Il y a entre X et Y une relation matricielle : $X = PY$ où P désigne la *matrice de passage* de la base (e_i) à la base (h_i). Les colonnes de cette matrice donnent les coordonnées des vecteurs de la "nouvelle base" (h_i) dans l' "ancienne base" (e_i). Ainsi, si p_{ij} est le terme général de la matrice P on a $h_j = \sum_i p_{ij} e_i$. On notera la matrice de passage $P = P_{(e_i)\to(h_i)}$. Cette matrice de passage est une matrice carrée inversible et son inverse est $P^{-1} = P_{(h_i)\to(e_i)}$. Si u est une application linéaire d'un espace E ayant deux bases (e_i) et (h_i) dans un espace F ayant deux bases (f_j) et (g_j) on a

$$\mathrm{mat}_{(e_i),(f_j)}\ u = P_{(f_j)\to(g_j)}\ \mathrm{mat}_{(h_i),(g_j)}\ u\ \ P_{(h_i)\to(e_i)}$$

Dans le cas d'un endomorphisme u sur un espace E ayant deux bases (e_i) et (h_i), la formule précédente se réécrit :

$$\mathrm{mat}_{(h_i)}\, u = P^{-1}_{(e_i)\to(h_i)}\, \mathrm{mat}_{(e_i)}\, u \;\; P_{(e_i)\to(h_i)}$$

Les matrices de u sur deux bases différentes sont donc deux matrices A et B qui vérifient une relation du type $A = P^{-1}\, B\, P$. Quand une telle relation a lieu entre les matrices A et B, on dit que les matrices A et B sont *semblables*. Ainsi on voit que deux matrices sont semblables si et seulement si ce sont les matrices d'un même endomorphisme sur deux bases différentes. Cette relation de similitude est une relation d'équivalence, c'est à dire qu'elle est réflexive, symétrique et transitive. Il ne faut pas confondre avec la notion de *matrices équivalentes*, qui induit aussi comme son nom l'indique une relation d'équivalence. Deux matrices A et B sont équivalentes si et seulement si il existe deux matrices P et Q inversibles telles que $A = PBQ$. En fait on prouve que deux matrices sont équivalentes si et seulement si elles ont même rang. Rappelons que le *rang* d'une matrice est égal à la dimension de l'espace engendré par ses vecteurs colonnes. En particulier deux matrices semblables ont même rang. On peut donc définir le rang d'une application linéaire $u \in L(E, F)$ comme le rang d'une de ses matrices. Ce rang est alors la dimension de l'*image* de u :

$$\mathrm{Im}(u) = \{u(x); x \in E\}$$

Rappelons la formule des dimensions appelée parfois formule de Grassman

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

On rappelle que $\ker u = \{x \in E, u(x) = 0\}$ appelé *noyau* de u.

Certaines identifications de notations peuvent se révéler commodes. Ainsi on peut convenir de noter de manière identique une matrice $d \times d$ et l'endomorphisme qui a cette matrice sur la base canonique de $\mathbb{K}^d$. On identifiera usuellement les vecteurs de $\mathbb{K}^d$ et les matrices colonnes de leurs composantes sur la base canonique.

1.2 Déterminants

Un outil utile en calcul matriciel est la notion de déterminant. Le déterminant d'une matrice carrée $A = (a_{ij})_{1\le i,j\le n}$ est un scalaire qui peut être défini par la formule

$$\det A = \sum_{\sigma\in\mathcal{S}_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Dans cette formule $\mathcal{S}_n$ désigne l'ensemble des bijections de $\{1, \dots, n\}$ sur lui-même, qui sont aussi appelées *permutations* de $\{1, \dots, n\}$. Pour $\sigma \in \mathcal{S}_n$, $\mathrm{sign}(\sigma)$ désigne la *signature* de la permutation σ. Cette signature vaut 1 ou -1 selon que le nombre d'inversion de σ défini comme le cardinal de $\{(i,j);\ 1 \leq i < j \leq n,\ \sigma(i) > \sigma(j)\}$ est pair ou impair. En fait cette formule ne peut être exploitée que pour les petites dimensions $n = 2$ ou $n = 3$ (règle de Sarrus). Remarquons que le déterminant d'une matrice A est égal à celui de sa transposée ${}^tA$.

Vu comme fonction des vecteurs colonnes de la matrice le déterminant est une application de $(\mathbb{K}^n)^n$ dans $\mathbb{K}$ qui, aux vecteurs $A_1, \dots, A_n$ de $\mathbb{K}^n$ associe le déterminant de la matrice A dont les colonnes sont successivement $A_1 \dots A_n$. Le déterminant est alors une forme *n-linéaire* –c'est à dire qu'il y a linéarité par rapport à chacune des n variables de $\mathbb{K}^n$– et cette forme n-linéaire est *alternée*–c'est à dire que le déterminant est nul dès que deux colonnes de la matrice sont identiques. Pour cette seconde propriété il est équivalent de dire que le déterminant est changé en son opposé si on échange deux colonnes. Toute forme n-linéaire alternée est proportionnelle au déterminant qui est donc la seule forme n-linéaire alternée qui donne la valeur 1 à la matrice identité.

Dans les calculs pratiques de déterminants de petite taille, il est commode d'utiliser le développement par rapport à une ligne ou une colonne. Pour la matrice $A = (a_{ij})_{1 \leq i,j \leq n}$ choisissons un indice de ligne $i_0 \in \{1, \dots, n\}$ et un indice de colonne $j_0 \in \{1, \dots, n\}$. On a

$$\det A = \sum_{j=1}^{n} a_{i_0 j}\, (-1)^{i_0 + j}\ \Delta_{i_0 j} = \sum_{i=1}^{n} a_{i j_0}\ (-1)^{i + j_0}\ \Delta_{i j_0}$$

Dans cette formule Δ_{ij} désigne le déterminant de la matrice $(n-1) \times (n-1)$ obtenue à partir de A en supprimant la i-ième ligne et la j-ième colonne. Grâce à cette formule on remplace le calcul d'un déterminant $n \times n$ par des déterminants $(n-1) \times (n-1)$. Cela est surtout intéressant si il y en a peu à calculer c'est à dire si beaucoup de coefficients de la ligne –ou de la colonne– selon laquelle on développe sont nuls. La quantité $(-1)^{i+j}\ \Delta_{ij}$ est appelé *cofacteurs* d'indices i, j. La matrice qui a ces quantités pour coefficients est logiquement appelé matrice des cofacteurs. Dans ce cours elle sera notée $\mathrm{cof}(A)$. Dans les livres on peut rencontrer le terme de comatrice qui désigne parfois la matrice des cofacteurs et parfois sa transposée. Avec nos notations on a l'égalité

$$A\ {}^t(\mathrm{cof}(A)) = {}^t\ (\mathrm{cof}(A))\ A = (\det A)\ I$$

En particulier on voit que pour une matrice A inversible, l'inverse est égal à la transposée de la matrice des cofacteurs multipliée par $1/\det A$.

Une propriété fondamentale du déterminant est que le déterminant du produit de deux matrices est égal au produit des déterminants de ces matrices. Il en résulte que deux matrices semblables ont même déterminant. On peut donc définir le déterminant d'un endomorphisme comme le déterminant de sa matrice dans une base quelconque, le choix de cette base n'ayant aucune importance.

1.3 Un rappel concernant les sommes

Une notion importante en théorie des espaces vectoriels est la notion de somme de sous-espaces vectoriels. Considérons $F_1, \dots, F_p$ sous-espaces vectoriels d'un espace vectoriel E. On définit la *somme* de ces sous-espaces par la formule

$$F_1 + \cdots + F_p = \sum_{i=1}^{p} F_i = \{x_1 + \cdots + x_p;\ \ x_1 \in F_1, \dots, x_p \in F_p\}$$

On dit que la somme est *directe* et on note alors cette somme $\bigoplus_{i=1}^{p} F_i$ ou $F_1 \oplus \cdots \oplus F_p$ si et seulement si l'une des conditions équivalentes suivantes est vérifiée

$$(i) \quad \forall k \in \{1, \dots, p\},\ \ F_k \bigcap \sum_{i \neq k} F_i = \{0\}$$

$$(ii) \quad \forall k \in \{2, \dots, p\},\ \ F_k \bigcap \sum_{i=1}^{k-1} F_i = \{0\}$$

$$(iii) \quad \forall y \in \sum_{i=1}^{p} F_i,\ \ \exists!(x_1, \dots, x_p) \in F_1 \times \cdots \times F_p,\ y = x_1 + \cdots + x_p$$

$$(iv) \quad (0 = x_1 + \cdots + x_p \text{ avec } x_1 \in F_1, \dots, x_p \in F_p) \Rightarrow x_1 = \cdots = x_p = 0$$

$$(v) \text{ (pour } E \text{ de dimension finie) } \dim\left(\sum_{i=1}^{p} F_i\right) = \sum_{i=1}^{p} \dim(F_i)$$

$$(vi) \forall \mathcal{B}_1, \dots, \mathcal{B}_p \text{ bases de } F_1, \dots, F_p,\ \ \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_p \text{ est une base de } \sum_{i=1}^{p} F_i$$

Dans cette dernière assertion on dit qu'une base de la forme $\mathcal{B}_1 \cup \ldots \cup \mathcal{B}_p$ est *adaptée* à la décomposition en somme directe.

Insistons. Il n'existe pas d'opération "somme directe" mais une opération somme qui est dans certains cas directe. La preuve du résultat qui précède est un bon exercice de révision.

1.4 Polynômes

Nous ne donnons ici que de très brefs rappels qui seront complétés au chapitre 3. On appelle polynôme à coefficient dans $\mathbb{K}$ une suite finie $p_0, \ldots, p_n$ d'éléments de $\mathbb{K}$ que l'on écrit sous la forme $p_0 + p_1X + \cdots + p_nX^n$ où X est appelé indéterminée. On dit alors que p_i est le coefficient du terme de degré i. Sur l'ensemble des polynômes noté $\mathbb{K}[X]$ on définit l'opération d'addition et de multiplication qui lui donnent la structure d'anneau commutatif. Avec les lois d'addition et de multiplication par un scalaire, $\mathbb{K}[X]$ a la structure d'espace vectoriel sur $\mathbb{K}$. La base dite canonique est la base $\{1, X, X^2, \ldots\}$.

Si $P(X) = p_0 + p_1X + \cdots + p_nX^n$ avec $p_n \neq 0$ on dit que P est de degré n. Par convention on dira que le degré du polynome nul est -1. L'ensemble des polynomes de degré inférieur ou égal à n, noté $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$ de dimension $n+1$ puisqu'on peut exhiber une base dite canonique en la famille $\{1, X, X^2, \ldots, X^n\}$.

On peut évaluer un polynôme $P(X) = p_0 + p_1X + \cdots + p_nX^n$ en $a \in \mathbb{K}$ en calculant la valeur de l'écriture précédente quand X est remplacé par a. Si $P(a) = p_0 + p_1a + \cdots + p_na^n = 0$ on dit que a est une racine de P. Cela équivaut au fait que P est multiple de $X - a$ ce qui signifie $P(X) = (X - a)\ Q(X)$ où $Q \in \mathbb{K}[X]$. On dira plus précisément que a est racine de P de multiplicité q $(\in \mathbb{N}^*)$ si P est multiple de $(X - a)^q$ mais pas de $(X - a)^{q+1}$. Cela équivaut au critère suivant portant sur les dérivées : $\forall j \in \{0, \ldots, q-1\}, \quad P^{(j)}(a) = 0$ et $P^{(q)}(a) \neq 0$. Quand P peut s'écrire comme produit de polynomes de degré 1 avec éventuelles répétitions on dit que P est scindé et si P peut s'écrire comme produit de polynômes de degré 1 distincts (en particulier n'a que des racines simples) on dit que P est simplement scindé.

1.5 Exercices

Exercice 1 Dans un espace vectoriel on considère trois sous-espaces vectoriels U, V, W et on suppose que $U \subset V \cup W$. Montrer que $U \subset V$ ou $U \subset W$.

Exercice 2 Dans un espace vectoriel réel de dimension finie n, une famille est dite positivement génératrice si tout vecteur de l'espace s'écrit comme combinaison linéaire à coefficients positifs d'éléments de cette famille. Quel est le cardinal minimum d'une telle famille ?

Exercice 3 Soit $A, B \in \mathcal{M}_d(\mathbb{R})$ On suppose que A et B sont semblables dans $\mathcal{M}_d(\mathbb{C})$ c'est à dire qu'il existe $P \in \mathcal{GL}_d(\mathbb{C})$ telle que $PB = AP$.

Montrer alors que A et B sont semblables dans $\mathcal{M}_d(\mathbb{R})$ c'est à dire qu'il existe $Q \in \mathcal{GL}_d(\mathbb{R})$ telle que $QB = AQ$.

Exercice 4 Montrer qu'un endomorphisme u d'un espace vectoriel E de dimension finie qui laisse stable toute droite vectorielle est une homothétie. Montrer qu'il en est de même si cet endomorphisme laisse stable tout hyperplan. Généralisation ?

Exercice 5 On note $M(a, b, c)$ la matrice réelle avec des a au dessous de la diagonale, des b sur la diagonale, des c au dessus de la diagonale.
1) Montrer que, pour toute matrice carrée (m_{ij}), par rapport à la variable réelle t, $\det[(m_{ij} + t)]$ est une fonction affine de t.
2) Pour $a \neq c$, calculer le déterminant de $M(a, b, c)$.
On suppose dorénavant $a = c$ et on note $M(a, b)$ pour $M(a, b, a)$.
3) Calculer le déterminant de $M(a, b)$

Exercice 6 (Extrait du partiel de novembre 2000) Pour $n \in \mathbb{N}^*$ et $\alpha_1, \ldots, \alpha_n$ nombres complexes donnés, on note $M(\alpha_1, \ldots, \alpha_n)$ la matrice $n \times n$ de terme général $(\alpha_i)^{j-1}$ où i désigne comme d'habitude l'indice de ligne et j l'indice de colonne.
1) Dans le cas $n = 3$, écrire explicitement $M(\alpha_1, \alpha_2, \alpha_3)$ et calculer son déterminant.
2) On revient au cas général $n \in \mathbb{N}^*$ quelconque.
2.a) Soit P un polynome unitaire (c'est à dire de coefficient dominant égal à 1) de degré $n - 1$. Montrer qu'on peut remplacer dans $\det M(\alpha_1, \ldots, \alpha_n)$ la dernière colonne dont le terme général est α_i^{n-1} par la colonne de terme général $P(\alpha_i)$.
2.b) Déterminer P pour qu'il s'annule en $\alpha_1, \ldots, \alpha_{n-1}$.
2.c) Avec ce choix de P en déduire une formule de récurrence sur $\det M(\alpha_1, \ldots, \alpha_n)$ c'est à dire exprimer $\det M(\alpha_1, \ldots, \alpha_n)$ en fonction de $\alpha_1, \ldots, \alpha_n$ et de

$\det M(\alpha_1, \dots, \alpha_{n-1})$.
2.d) En déduire une expression de $\det M(\alpha_1, \dots, \alpha_n)$.
2.e) Donner une condition nécessaire et suffisante pour que la matrice $M(\alpha_1, \dots, \alpha_n) \neq 0$ soit inversible.

Chapitre 2

Réduction : objectifs et premiers résultats

2.1 Objectifs de la réduction des matrices

Considérons un endomorphisme u d'un espace vectoriel E de dimension finie, qu'on aimerait pouvoir décrire de façon simple. On a ainsi l'espoir de trouver des sous espaces vectoriels de E sur lesquels u agit simplement. Remarquons que si on veut parler de restriction de u à un sous-espace vectoriel F au sens où cette restriction serait un endomorphisme de F il faut que l'image de F par u soit incluse dans F. Cela conduit au vocabulaire suivant. On dit qu'un sous-espace F est *stable* si $u(F) \subset F$.

Une situation agréable est celle où E se décompose comme somme directe de sous-espaces stables c'est à dire

$$E = F_1 \oplus F_2 \oplus \cdots \oplus F_k \text{ avec } u(F_i) \subset F_i$$

Dans ce cas on peut obtenir une base de E en réunissant k bases respectives de $F_1, F_2, \ldots, F_k$, une telle base étant dite adaptée à la décomposition en somme directe. La matrice de u dans une telle base est *diagonale par blocs* c'est à dire qu'elle s'écrit par blocs sous la forme

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix}$$

Dans cette écriture A_1 est une matrice carrée –un bloc– dont le nombre de lignes et colonnes est $\dim F_1$. De même, A_2 est un bloc dont le nombre de lignes et colonnes est dim F_2, etc. Le 0 situé en dessous de A_1 désigne un bloc de 0 –une matrice à coefficients tous nuls– ayant dim F_1 colonnes et dim F_2 lignes. Pour se convaincre d'une telle forme pour la matrice de u, prenons un des dim F_1 premiers vecteurs de la base. C'est donc un vecteur de F_1 et son image par u est dans F_1. Elle s'exprime donc comme combinaison linéaire des dim F_1 premiers vecteurs de la base. On justifie ainsi la présence des blocs de 0 sous A_1. Il en va de même pour les autres blocs de 0 apparaissant dans la matrice.

La décomposition évoquée ci-dessus est d'autant plus intéressante que l'action de u sur chacun des sous-espaces stables est simple. Le plus simple que l'on puisse imaginer est l'homothétie $x \mapsto \lambda\, x$ avec λ scalaire. Dans ce cas la matrice de la restriction de u à ce sous-espace est $\lambda\, I$ sur n'importe quelle base. On étudiera le cas où l'espace se décompose en somme directe de sous-espaces sur lesquels u agit comme une homothétie. Il s'agit du cas où u est *diagonalisable*. Ce cas est celui où la matrice de u sur une certaine base est diagonale. Malheureusement tout endomorphisme n'est pas diagonalisable. On peut alors chercher des bases sur lesquelles la matrice est d'une autre forme relativement simple. On pense à la forme triangulaire, par exemple triangulaire supérieure où tous les coefficients sous la diagonale sont nuls.

On ira encore plus loin dans un prochain chapitre avec la réduction sous forme de Jordan qui est une forme diagonale par bloc comme évoqué ci-dessus avec des blocs diagonaux très simple, en particulier triangulaires supérieurs.

2.2 Eléments propres

Rappelons que, disposant d'un endomorphisme u sur l'espace E de dimension finie d, le cas idéal serait de décomposer l'espace E en sous-espaces où u agit comme une homothétie. Cela nous conduit à introduire un certain vocabulaire. Si un vecteur non nul x est tel que $u(x) = \lambda\, x$, il est appelé *vecteur propre* associé à la *valeur propre* λ. Pour λ valeur propre, l'ensemble des vecteurs propres qui lui sont relatifs, auxquels on rajoute le vecteur nul, à savoir,

$$\{x \in E;\ u(x) = \lambda\, x\} = \ker(u - \lambda \mathrm{id})$$

est appelé *sous-espace propre* associé à la valeur propre λ. Un tel sous-espace est stable et la matrice de la restriction de u à ce sous-espace de dimension r est $\lambda\, I_r$ c'est à dire la matrice avec des λ sur la diagonale et des 0 ailleurs. On appelle *spectre* de u et on notera dans ce cours $\mathrm{sp}(u)$ l'ensemble des valeurs propres de u. Notons que les valeurs propres sont les racines du *polynôme* *caractéristique* de u noté dans ce cours C_u et défini par

$$C_u(X) = \det(X\,I - \mathrm{mat}\ u)$$

Dans cette définition mat u désigne la matrice de u sur une base qui n'a pas besoin d'être spécifiée car le déterminant de deux matrices semblables est le même. Avec cette définition C_u est un polynôme unitaire de degré $d = \dim E$. On trouve parfois dans les livres la définition $C_u = \det(\mathrm{mat}\ u - X\,I)$ qui diffère de la notre du facteur $(-1)^d$.

Concernant le polynôme caractéristique faisons maintenant une petite remarque qui interviendra bientôt. Si F est un sous-espace stable de E alors le polynôme caractéristique de la restriction de u à F divise C_u. En effet écrivons $E = F \oplus H$ où H est un supplémentaire quelconque de F et prenons une base adaptée à cette décomposition. Alors dans cette base

$$\mathrm{mat}\, u = \begin{pmatrix} \mathrm{mat}\, u|_F & A \\ 0 & B \end{pmatrix} \text{ d'où } X\,I - \mathrm{mat}\, u = \begin{pmatrix} X\,I - \mathrm{mat}\, u|_F & A \\ 0 & X\,I - B \end{pmatrix}$$

où A et B sont deux blocs sans importance. Alors

$$C_u(X) = \det(X\,I - \mathrm{mat}\, u|_F)\ \ \det(X\,I - B) = C_{u|_F}(X)\ \ \det(X\,I - B)$$

prouvant ainsi que C_u est multiple de $C_{u|_F}$.

2.3 Diagonalisation, trigonalisation

Revenons à la réduction de u. On dit que l'endomorphisme u de E est *diagonalisable* si

$$E = \bigoplus_{\lambda \in \mathrm{sp}(u)} \ker(u - \lambda\,\mathrm{id})$$

ce qui équivaut à dire qu'il existe une base de E telle que

$$\begin{pmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_k I \end{pmatrix}$$

en notant $\mathrm{sp}(u) = \{\lambda_1, \dots, \lambda_k\}$, ou plus simplement qu'il existe une base de E sur laquelle la matrice de u est diagonale. Cela revient à dire également qu'il existe une base de E formée de vecteurs propres de u.

On dira qu'une matrice A est diagonalisable si l'endomorphisme qui a A pour matrice dans une base quelconque est diagonalisable. Il est équivalent de dire que A est semblable à une matrice diagonale.

Selon la définition, pour savoir si un endomorphisme de E est diagonalisable, il faut calculer les valeurs propres puis les sous-espaces propres et voir si ces sous-espaces propres sont en somme directe et si cette somme est égale à E. En fait nous prouverons au prochain chapitre le résultat suivant que nous admettons pour l'instant.

Proposition 1 *La somme de sous-espaces propres (distincts) est toujours directe.*

Pour voir si u est diagonalisable il s'agit donc seulement de voir si la somme (directe) des sous-espaces propres est égale à E ou de manière équivalente si la somme des dimensions des sous-espaces propres est égale à E. A ce sujet il faut remarquer que pour toute valeur propre λ de l'endomorphisme u, c'est à dire pour toute racine du polynôme caractéristique C_u, la dimension du sous-espace propre $d_\lambda = \dim\ker(u - \lambda\,\mathrm{id})$ est inférieure ou égale à la multiplicité m_λ de λ comme racine de C_u. En effet on sait que le polynôme caractéristique de la restriction de u à $\ker(u-\lambda\,\mathrm{id})$ divise C_u. Or le polynôme caractéristique de cette restriction est $(X-\lambda)^{d_\lambda}$ d'où le résultat. En sommant les inégalités $\dim\ker(u - \lambda\,\mathrm{id}) \leq m_\lambda$, on a

$$\sum_{\lambda\in\mathrm{sp}(u)} \dim\ker(u - \lambda\,\mathrm{id}) \leq \sum_{\lambda;\ C_u(\lambda)=0} m_\lambda \leq d(C_u) = \dim(E)$$

On sait que u est diagonalisable si et seulement si les deux extremités de l'inégalite sont égales. Cela a lieu si et seulement si toutes les inégalités que l'on a sommé sont des égalités c'est à dire

$$\sum_{\lambda;\ C_u(\lambda)=0} m_\lambda = d(C_u) \text{ et } \forall\lambda \in \mathrm{sp}(u),\ \ \dim\ker(u - \lambda\,\mathrm{id}) = m_\lambda$$

Ainsi :

Proposition 2 *Un endomorphisme est diagonalisable si et seulement si son polynôme caractéristique est scindé et la dimension de tout sous-espace propre est égale à la multiplicité de cette valeur propre comme racine du polynôme caractéristique.*

La diagonalisation d'un endomorphisme n'est pas toujours possible, que le corps des scalaires soit $\mathbb{R}$ ou $\mathbb{C}$. Il faut envisager une autre réduction pour ces cas d'échec. On verra la réduction de Jordan au prochain chapitre. Voyons tout de suite un peu moins sophistiqué. On dit qu'un endomorphisme

u est *trigonalisable* si il existe une base de E sur laquelle sa matrice est triangulaire supérieure. Bien sûr on dira qu'une matrice carrée est trigonalisable si c'est la matrice d'un endomorphisme trigonalisable c'est à dire qu'elle est semblable à une matrice triangulaire supérieure. On dit parfois triangularisable voire trigonable au lieu de trigonalisable. On peut donner un critère simple.

Proposition 3 *Un endomorphisme est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

Comme tout polynôme à coefficients complexes est scindé, toute matrice à coefficients complexes est trigonalisable dans l'ensemble des matrices complexes $\mathcal{M}_n(\mathbb{C})$. On remarquera qu'une matrice à coefficients réels n'est qu'un cas particulier de matrice à coefficients complexes. Elle est donc trigonalisable mais dans $\mathcal{M}_n(\mathbb{C})$ c'est à dire avec matrice de passage et matrice triangulaire semblable complexes.

Nous n'allons pas prouver la proposition 3 maintenant puisque nous aurons beaucoup mieux au prochain chapitre avec le théorème de Jordan. Toutefois une preuve est proposée à l'exercice 9.

2.4 Exercices

Exercice 7 Cet exercice donne une localisation–assez grossière–des valeurs propres d'une matrice. Ce résultat s'appelle pompeusement théorème de Gerschgorin-Hadamard.

Dans $\mathbb{R}$ ou $\mathbb{C}$ on notera $\overline{B}(a,r)$ la boule fermée de centre a et de rayon r. Il s'agit donc simplement de l'intervalle $[a-r, a+r]$ dans le cas réel et le disque fermée de centre a et de rayon r dans le cas complexe. Soit λ une valeur propre de la matrice $A = (a_{ij})_{1\leq i,j\leq n}$. Montrer que

$$\lambda \in \bigcup_{i=1}^{n} \overline{B}\left(a_{ii}, L_i\right) \text{ où } L_i = \sum_{1\leq j\leq n; j\neq i} |a_{ij}|$$

En déduire une condition suffisante appelée condition de diagonale dominante qui porte sur les coefficients d'une matrice $A = (a_{ij})_{1\leq i,j\leq n}$ et qui assure que cette matrice est inversible.

Exercice 8 L'application linéaire de $\mathcal{M}_n(\mathbb{K})$ dans lui même qui, à une matrice associe sa transposée est elle diagonalisable ?

Exercice 9 Il s'agit de prouver qu'un endomorphisme d'un espace vectoriel de dimension finie est trigonalisable si et seulement si son polynôme caractéristique est scindé.

Prouver d'abord que si un endomorphisme est trigonalisable, son polynôme caractéristique est scindé.

Pour la réciproque, procéder par récurrence sur la dimension de l'espace. Pour se ramener de la dimension n à l'hypothèse de récurrence en dimension $n-1$, trouver un vecteur propre et appliquer l'hypothèse de récurrence à la composée de l'endomorphisme par la projection sur un supplémentaire de la droite propre.

Exercice 10 Donner un exemple de matrice 2×2 trigonalisable mais non diagonalisable. Donner un exemple de matrice 2×2 non trigonalisable (sur $\mathbb{R}$!).

Exercice 11 Soit A une matrice carrée à coefficients complexes et $\varepsilon > 0$. Montrer qu'il existe une matrice inversible P telle que $T = P^{-1}AP$ soit triangulaire supérieure et diagonale à ε près c'est à dire $|t_{ij}| \leq \varepsilon$ pour $j > i$.

Exercice 12 On reprend la matrice $M(a,b,c)$ étudiée à l'exercice 5 qui a des coefficients réels égaux à a en dessous de la diagonale, à b sur la diagonale et à c au dessus de la diagonale.

1) Calculer le polynôme caractéristique de $M(a,b,c)$.

2) Pour $a \neq c$, déterminer les valeurs propres complexes de $M(a,b,c)$.
Si $a = c$ et on note $M(a,b)$ au lieu de $M(a,b,a)$.
3) Déterminer le polynôme caractéristique de $M(a,b)$.
4) Déterminer les valeurs propres de $M(a,b)$ et leur multiplicité.

Exercice 13 (Extrait du partiel de novembre 2000) 1) Soit B est une matrice carrée réelle ou complexe, nilpotente (i.e. $\exists q \in \mathbb{N},\ \ B^q = 0$). Prouver que, pour tout $p \in \mathbb{N}^*$, la trace de B^p est nulle.
2) On suppose maintenant que A est une matrice carrée complexe telle que, pour tout $p \in \mathbb{N}^*$, la trace de A^p est nulle.
2.a) Justifier que la polynôme caractéristique s'écrit

$$C_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{\omega(i)}$$

où $k \in \mathbb{N}^*$, $\lambda_1, \dots, \lambda_k$ sont des nombres complexes distincts et $\omega(1), \dots, \omega(k)$ sont des entiers naturels non nuls.
2.b) Pour tout $p \in \mathbb{N}^*$, exprimer la trace de A^p en fonction de p, des $\lambda_1, \dots, \lambda_k$ et des $\omega(1), \dots, \omega(k)$.
2.c) Soit $M(\lambda_1, \dots, \lambda_k) = ((\lambda_j)^{i-1})_{1 \leq i,j \leq k}$ déja étudiée à l'exercice 6. Calculer $M(\lambda_1, \dots, \lambda_k)\, V$ où V est la matrice colonne $V = (\omega(i)\ \lambda_i)_{1 \leq i \leq k}$.
2.d) Montrer que A est nilpotente.

Chapitre 3

Polynômes et endomorphismes

3.1 Polynômes d'endomorphismes

Nous allons introduire une notion fondamentale pour développer la théorie de la réduction. Si A est une matrice carrée et si $P(X) = p_0 + p_1 X + \cdots + p_k X^k$ est un polynôme, on définit la matrice $P(A)$ par

$$P(A) = p_0 I + p_1 A + \cdots + p_k A^k$$

De même si u est un endomorphisme d'un espace vectoriel E, l'endomorphisme $P(u)$ est

$$P(u) = p_0 \,\mathrm{id} + p_1 u + \cdots + p_k u^k$$

où $u^k = u \circ u \circ \cdots \circ u$. Les deux définitions sont cohérentes dans la mesure où

$$\mathrm{mat}_{\mathcal{B}} P(u) = P(\mathrm{mat}_{\mathcal{B}} u)$$

Cette définition donne les propriétés algébriques souhaitables :

$$(P+Q)(A) = P(A) + Q(A),\ P(A)\,Q(A) = (PQ)(A) = Q(A)\,P(A)$$

pour tous polynômes P, Q et toute matrice carrée A. En particulier deux matrices qui sont polynômes d'une même matrice commutent. On peut aussi l'exprimer en termes d'endomorphismes :

$$(P+Q)(u) = P(u) + Q(u),\ P(u) \circ Q(u) = (PQ)(u) = Q(u) \circ P(u)$$

Un exemple très simple est le calcul d'un polynôme d'une matrice diagonale :

$$\text{si } A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ et si } P \in \mathbb{K}[X] \text{ alors } P(A) = \begin{pmatrix} P(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & P(\lambda_n) \end{pmatrix}$$

En effet cela est vrai si $P(X) = X^k,\ k \in \mathbb{N}$ et le cas général s'en déduit. Cette propriété sera prolongée à l'exercice 15. Notons aussi que :

$$\text{si } B = Q^{-1}AQ \text{ et si } P \in \mathbb{K}[X] \text{ alors } P(B) = Q^{-1}P(A)Q$$

Là encore on peut le déduire du cas $P(X) = X^k,\ k \in \mathbb{N}$ qui se prouve aisément pas récurrence.

L'intérêt des concepts définis dans ce paragraphe est de pouvoir utiliser les propriétés arithmétiques des polynômes pour travailler sur les endomorphismes associés. Dans le paragraphe qui suit nous allons rappeler quelques propriétés fondamentales des polynômes.

3.2 Arithmétique des polynômes

L'ensemble des polynômes muni des deux opérations addition et multiplication possède la structure algébrique d'anneau *euclidien*. Ce type de structure sera étudié en détail dans le cours d'algèbre générale. Pour nous le terme euclidien signifiera que l'on dispose d'une *division euclidienne* à savoir si A et B sont deux polynômes non nuls, il existe un unique couple de polynômes (Q, R) tel que $A = BQ + R$ et $d(R) < d(B)$. Dans cette dernière inégalité sur le degré, il faut rappeler que l'on convient que le degré du polynôme nul est -1. On dispose alors de la notion de divisibilité : le polynôme B divise le polynôme A si $A = BQ$. On note $B\,|A$. On dit aussi que A est multiple de B. Au niveau de la divisibilité deux polynômes qui diffèrent d'une constante multiplicative non nulle ont les mêmes propriétés. On dit qu'ils sont *associés*. Par exemple tout polynôme est divisible par lui-même et par tous les polynômes qui lui sont associés. Il est aussi divisible par les polynômes associés à 1 qui sont les constantes non nulles.

On dit qu'un polynôme est *irréductible* s'il est de degré au moins 1 et s'il n'est divisible que par 1 et lui-même et leurs associés. Par exemple dans l'anneau $\mathbb{C}[X]$ des polynômes à coefficients complexes les polynômes irréductibles sont ceux de degré 1 ; dans l'anneau $\mathbb{R}[X]$ des polynômes à coefficients réels les polynômes irréductibles sont ceux de degré 1 et ceux de degré deux sans racine réelle, c'est à dire de discriminant strictement négatif.

Dans l'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ et indéterminée X, il est utile de considérer certaines parties appelées *idéaux*. Un idéal de $\mathbb{K}[X]$ est une partie I non vide telle que $A, B \in I \Rightarrow A + B \in I$ et $A \in I \Rightarrow A\,P \in I$ pour tout polynôme P. Par exemple pour un polynôme H donné, la partie

$$\langle H \rangle = \{H\,P;\ P \in \mathbb{K}[X]\}$$

est un idéal appelé idéal engendré par le polynôme H. C'est l'ensemble des multiples du polynôme H. Il est facile de voir avec la division euclidienne qu'en fait tout idéal de l'anneau des polynômes est de ce type. Notons que deux polynômes engendrent le même idéal si et seulement si ils sont associés. Il est également facile de prouver que la somme d'idéaux ou l'intersection d'idéaux donne un idéal. Ainsi si $P_1, \dots, P_k$ sont des polynômes il en existe deux autres, D et M tels que

$$\langle P_1 \rangle + \cdots + \langle P_k \rangle = \langle D \rangle,\ \ \langle P_1 \rangle \cap \cdots \cap \langle P_k \rangle = \langle M \rangle$$

Les deux polynômes D et M sont uniques à association près. Ils deviennent donc uniques si on leur impose d'être unitaires. On note

$$D = \mathrm{PGCD}(P_1, \dots, P_k) \text{ et } M = \mathrm{PPCM}(P_1, \dots, P_k)$$

La dénomination PGCD –plus grand diviseur commun– vient de la caractérisation suivante de D : $D\,|P_1, \dots, D\,|P_k$ et si $H\,|P_1, \dots, H\,|P_k$ alors $H\,|D$. On pourra vérifier cela en exercice. Parallélement le PPCM M est le plus petit multiple commun à $P_1, \dots, P_k$ c'est à dire que $P_1\,|M, \dots, P_k\,|M$ et $P_1\,|H, \dots, P_k\,|H \Rightarrow M\,|H$.

Si $\mathrm{PGCD}(P_1, \dots, P_k) = 1$, les polynômes $P_1, \dots, P_k$ sont dits *premiers entre eux*. On notera qu'avec la définition précédente de PGCD, on obtient directement le résultat suivant.

Théorème 4 (Bezout) *Soit $P_1, \dots, P_k$ des polynômes et D leur PGCD alors il existe des polynômes $A_1, \dots, A_k$ tels que*

$$D = A_1\,P_1 + \cdots + A_k\,P_k$$

Inversement si il existe des polynômes $A_1, \dots, A_k$ tels que

$$1 = A_1\,P_1 + \cdots + A_k\,P_k$$

alors les polynômes $P_1, \dots, P_k$ sont premiers entre eux c'est à dire que leur PGCD est 1.

On notera que le PGCD et le PPCM de deux polynômes A et B sont tels que leur produit est associé à AB.

L'anneau des polynômes a comme propriété fondamentale qu'il est "*factoriel*" : tout polynôme se décompose de manière unique (à l'ordre près) comme produit de polynômes irréductibles. Si on a écrit les décompositions de deux polynômes comme produits d'irréductibles affectés de certaines puissances la décompostion en irréductibles du PGCD s'obtient en considérant les facteurs irréductibles qui apparaissent dans les deux décompositions et en les affectant de la plus petite des deux puissances apparaissant. Pour le PPCM on prend tout facteur irréductible qui apparait dans au moins l'une de deux décompositions et on l'affecte de la plus grande puissance apparaissant dans ces deux décompositions. On est parfois amené à utiliser les deux résultats suivants. Le *lemme d'Euclide* dit que si un polynôme irréductible H divise le produit $A\,B$ alors il divise A ou B. Le *théorème de Gauss* dit que si $H\,|A\,B$ et si H est premier avec A alors $H\,|B$.

Pour obtenir la décomposition d'un polynôme $P \in \mathbb{K}[X]$ en produit de polynômes irréductibles et pour bien d'autres usages, on commence par chercher les *racines*. On rappelle que ζ est racine de P si $P(\zeta) = 0$ ou de manière équivalente si $X - \zeta|P$. On dira plus précisément que ζ est racine de P de multiplicité q ($\in \mathbb{N}^*$) si $(X - \zeta)^q$ divise P mais $(X - \zeta)^{q+1}$ ne divise pas P. Cela équivaut au critère suivant portant sur les dérivées : $\forall j \in \{0, \ldots, q-1\},\ \ P^{(j)}(\zeta) = 0$ et $P^{(q)}(\zeta) \neq 0$. Ce critère vient essentiellement de la formule de Taylor pour les polynômes qui ressemble à la formule de Taylor de l'analyse mais qui est dans ce cadre algébrique exacte, c'est à dire sans reste. Elle s'écrit pour $P \in \mathbb{K}_d[X]$ et $a \in \mathbb{K}$,

$$P(X) = \sum_{k=0}^{d} \frac{P^{(k)}(a)}{k!}(X - a)^k$$

C'est en fait l'écriture de P dans la base $\{(X - a)^k, k \leq d\}$.

Quand $\mathbb{K} = \mathbb{C}$, le théorème de d'Alembert dit que tout polynôme non constant admet au moins une racine et donc que tout polynôme non constant est scindé. C'est cela qui entraine que dans $\mathbb{C}[X]$ les seuls polynômes irréductibles sont ceux de degré 1.

Quand $\mathbb{K} = \mathbb{R}$, on peut toujours considérer qu'un polynôme à coefficients réels est un cas particulier de polynôme à coefficients complexes et effectuer la factorisation dans $\mathbb{C}[X]$. Il est très facile de voir que si $a \in \mathbb{C}$ est racine de $P \in \mathbb{R}[X]$ alors le conjugué $\overline{a}$ est aussi racine de P et avec la même multiplicité que a. Cela entraine que dans la factorisation en irréductibles de P dans $\mathbb{C}[X]$ on peut regrouper $(X - a)^k$ et $(X - \overline{a})^k$ en $(X^2 - 2\mathrm{Re}(a)X -$

$|a|^2)^k$. Ainsi la factorisation de P s'écrit comme produit de polynômes du type précédent et de polynômes de degré 1 qui correspondent aux racines réelles. On voit donc bien que dans $\mathbb{R}[X]$ les polynômes irréductibles sont ceux de degré 1 et ceux de degré deux sans racine réelle.

3.3 Le lemme des noyaux

Nous allons d'abord revenir à l'utilisation des polynômes d'endomorphismes en énonçant un résultat qui nous sera utile dans les developpements suivants.

Théorème 5 (Lemme des noyaux) *Soit $P_1, \ldots, P_k$ des polynômes premiers entre eux deux à deux et u un endomorphisme alors*

$$\ker\left[(P_1 \cdots P_k)(u)\right] = \bigoplus_{i=1}^{k} \ker\left[P_i(u)\right]$$

Preuve. Montrons que les polynômes $\prod_{j\neq i} P_j$ pour $i \in \{1, \ldots, k\}$ sont premiers entre eux dans leurs ensemble, c'est à dire que le PGCD de ces k polynômes est 1. Supposons qu'ils sont tous divisibles par un polynôme irréductible H. Alors H divise le premier : $\prod_{j\neq 1} P_j$ donc il existe $j_1 \neq 1$ tel que $H\,|P_{j_1}$. Mais $H \mid \prod_{j\neq j_1} P_j$ donc il existe $j_2 \neq j_1$ tel que $H\,|P_{j_2}$. Cela contredit le fait que $\mathrm{PGCD}(P_{j_1}, P_{j_2}) = 1$.

Par le théorème de Bezout on en déduit l'existence de polynômes $A_1, \ldots, A_k$ tels que

$$1 = \sum_{i=1}^{k} A_i \prod_{j\neq i} P_j$$

d'où

$$\mathrm{id} = \sum_{i=1}^{k} A_i(u) \circ \prod_{j\neq i} P_j(u)$$

Soit alors $x \in \ker\left[\prod_{j=1}^{k} P_j(u)\right]$, Remarquons que

$$x_i = A_i(u) \circ \prod_{j\neq i} P_j(u)(x) \in \ker P_i(u)$$

et qu'on a $x = \sum_{i=1}^{k} x_i$. Donc le résultat de somme annoncé est vrai. Il reste à voir que cette somme est directe. Prenons donc $i \in \{1, \ldots, k\}$ et

$$y \in \ker\left[P_i(u)\right] \cap \left(\sum_{j\neq i} \ker\left[P_j(u)\right]\right)$$

Il s'agit de voir que $y = 0$. Nous écrivons $y = \sum_{j \neq i} y_j$ avec $y_j \in \ker [P_j(u)]$. On voit facilement par le lemme d'Euclide que P_i et $\prod_{l \neq i} P_l$ sont premiers entre eux. Par le théorème de Bezout, il existe A et B tels que

$$A\, P_i + B \prod_{l \neq i} P_l = 1$$

Donc en appliquant à u puis à $y = \sum_{j \neq i} y_j$ on obtient

$$y = A(u) \circ P_i(u)(y) + B(u) \circ \sum_{j \neq i} \prod_{l \neq i} P_l(u)(y_j)$$

En utilisant les hypothèses sur y et sur y_j cela conduit bien à $y = 0$.

Une conséquence de ce lemme des noyaux est que la somme de sous-espace propres distincts est directe c'est à dire la **preuve de la proposition 1**. En effet si $\lambda_1, \dots, \lambda_k$ sont des valeurs propres distinctes, les polynômes $X - \lambda_1, \dots, X - \lambda_k$ sont premiers entre eux deux à deux donc la somme $\sum_{i=1}^{k} \ker(u - \lambda_i \,\mathrm{id})$ est directe.

Le lemme des noyaux est particulièrement intéressant si les polynômes $P_1, \dots, P_k$ qui figurent dans son énoncé sont tels que $\prod_{i=1}^{k} P_i(u) = 0$ i.e. $P(u) = 0$ avec $P = \prod_{i=1}^{k} P_i$ car alors le noyau de ce dernier endomorphisme est E et on dispose d'une décomposition de E comme somme directe des sous-espaces $\ker P_i(u)$. Nous sommes donc naturellement amenés à étudier $\{P \in \mathbb{K}[X];\ P(u) = 0\}$

3.4 Idéal annulateur

Soit u un endomorphisme de l'espace de dimension finie E. Dans ce paragraphe on va noter $\mathrm{Ann}(u)$ l'ensemble des polynômes qui annulent u : $\mathrm{Ann}(u) = \{P;\ P(u) = 0\}$. Cet ensemble n'est pas réduit au polynôme nul. En effet la famille d'endomorphismes $\{\mathrm{id}, u, u^2, \dots, u^N\}$ où $N = (\dim E)^2$ est une famille qui possède $N + 1$ éléments dans l'espace vectoriel $L(E)$ des endomorphismes de E qui est de dimension $(\dim E)^2 = N$. La famille exhibée est donc liée et il existe des réels $\alpha_0, \dots, \alpha_N$ non tous nuls tels que

$$\alpha_0 \,\mathrm{id} + \alpha_1\, u + \alpha_2\, u^2 + \cdots + \alpha_N\, u^N = 0$$

ce qui veut exactement dire que $H(u) = 0$ où H est le polynôme non nul $H(X) = \sum_{i=0}^{N} \alpha_i\, X^i$.

Il est élémentaire de vérifier que $\mathrm{Ann}(u)$ est un idéal de l'ensemble des polynômes. On sait donc que cet idéal est engendré par un unique polynôme unitaire que l'on va noter M_u et que l'on appelle *polynôme minimal* de u. Il s'agit ainsi du polynôme de plus petit degré parmi tous les polynômes unitaires qui annulent u. L'étude du polynôme minimal est un des objectifs du prochain chapitre.

3.5 Exercices

Exercice 14 (Résultant) 1) Soit deux polynômes

$$P(X) = \sum_{0 \le k \le n} a_k \, X^k \text{ et } Q(X) = \sum_{0 \le k \le m} b_k \, X^k$$

de degrés respectifs n et m. Ecrire la matrice $S(P, Q)$ de la famille

$$\{P, XP, \ldots, X^{m-1}P, Q, XQ, \ldots, X^{n-1}Q\}$$

dans la base canonique.
2) Montrer que P et Q ont une racine commune dans $\mathbb{C}$ si et seulement si il existe deux polynômes non nuls U et V tels que $d(U) < d(Q)$ et $d(V) < d(P)$ et $UP + VQ = 0$.
3) Montrer que P et Q ont une racine commune dans $\mathbb{C}$ si et seulement si $\det[S(P, Q)] = 0$.

Exercice 15 Montrer que si A est une matrice triangulaire supérieure $n \times n$ de termes diagonaux $\lambda_1, \ldots, \lambda_n$ et si P est un polynôme alors $P(A)$ est une matrice triangulaire supérieure de termes diagonaux $P(\lambda_1), \ldots P(\lambda_n)$.

Que se passe-t-il si en plus $\lambda_1 = \cdots = \lambda_n = 0$?

Chapitre 4

Polynômes minimaux et caractéristiques

4.1 Polynôme minimal

Reprenons u un endomorphisme de l'espace de dimension finie E. On a vu au chapitre précédent que $\mathrm{Ann}(u) = \{P;\ P(u) = 0\}$ n'est pas réduit au polynôme nul et qu'il s'agit d'un idéal de l'ensemble des polynômes. On sait donc que cet idéal est engendré par un unique polynôme unitaire que l'on a noté M_u et que l'on appelle *polynôme minimal* de u. Il s'agit ainsi du polynôme de plus petit degré parmi tous les polynômes unitaires qui annulent u. L'égalité $\mathrm{Ann}(u) = \langle M_u \rangle$ nous dit de plus que tout polynôme qui annule u est un multiple de M_u.

On définit le polynôme minimal M_A d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ de manière similaire c'est à dire que M_A est l'unique générateur unitaire de l'idéal $\mathrm{Ann}(A) = \{P \in \mathbb{K}[X];\ \ P(A) = 0\}$. On remarque que si $P \in \mathbb{K}[X]$,et en notant $\mathrm{mat}u$ la matrice de l'endomorphisme u dans une certaine base $\mathcal{B}$ de E, on a

$$P(\mathrm{mat}_{\mathcal{B}} u) = \mathrm{mat}_{\mathcal{B}} P(u) \text{ d'où } \mathrm{Ann}(u) = \mathrm{Ann}(\mathrm{mat}_{\mathcal{B}} u)$$

Ainsi l'idéal annulateur d'un endomorphisme u est égal à l'idéal annulateur d'une de ses matrices. Le polynôme minimal d'un endomorphisme est donc égal au polynôme minimal (d'une) de ses matrices. Cela confirme que deux matrices semblables ont même polynôme minimal ce qui résulte trivialement de :

$$\forall P \in \mathbb{K}[X],\ \ \forall A \in \mathcal{M}_n(\mathbb{K}), \forall Q \in \mathcal{GL}_n(\mathbb{K}) P(Q^{-1} A Q) = Q^{-1}\, P(A)\, Q$$

Nous allons maintenant voir quelques exemples. Considérons d'abord une matrice diagonale :

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ alors } \forall P \in \mathbb{K}[X], \quad P(A) = \begin{pmatrix} P(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & P(\lambda_n) \end{pmatrix}$$

d'ou $P(A) = 0$ si et seulement si $P(\lambda_1) = \ldots = P(\lambda_n) = 0$. Ainsi on obtient déja le polynôme minimal de tout endomorphisme diagonalisable :

$$\text{si } u \text{ diagonalisable, } M_u(X) = \prod_{\lambda \in \mathrm{sp}(u)} (X - \lambda)$$

Le calcul du polynôme minimal est facile si on connait a priori un polynôme annulateur car le polynôme minimal en sera un diviseur. Par exemple une projection p vérifie $p^2 = p$ donc $M_p(X)$ est $X^2 - X$ ou $X - 1$ ou X les deux dernières possibilités étant à exclure sauf si $p = \mathrm{id}$ ou $p = 0$ respectivement. De même une symétrie s est involutive c'est à dire $s^2 = \mathrm{id}$ donc $M_s(X) = X^2 - 1$ pour une symétrie non dégénérée. Terminons par un dernier exemple. On considère dans un espace vectoriel de base $(b_1, \ldots, b_n)$ l'endomorphisme θ associé au décalage dans la base c'est à dire que θ est tel que $\theta(b_i) = b_{i+1(\mathrm{mod}\ n)}$. Le lecteur prouvera en exercice que le polynôme minimal est $M_\theta(X) = X^n$.

4.2 Application à la recherche de sous-espaces stables

Comme on l'a vu précédemment, la réduction des endomorphismes démarre avec la recherche de sous-espaces stables pour cet endomorphisme. Il est intéressant de remarquer que les polynômes d'endomorphismes nous donne de tels sous-espaces stables, par leurs noyaux. En effet si P est un polynôme et si $x \in E$ vérifie $P(u)(x) = 0$ alors $P(u)(u(x)) = u[P(u)(x)] = 0$ montrant ainsi que $\ker P(u)$ est stable. Nous pouvons prolonger cette remarque par le résultat suivant.

Proposition 6 *Si E est un espace vectoriel sur $\mathbb{C}$ de dimension finie et u un endomorphisme de E alors u admet un sous-espace stable de dimension 1, c'est à dire une droite vectorielle stable. Si E est un espace vectoriel sur $\mathbb{R}$ de dimension finie et u un endomorphisme de E alors u admet un sous-espace stable de dimension 1 ou 2, c'est à dire une droite ou un plan vectoriel stable.*

Preuve. Considérons un facteur irréductible unitaire H du polynôme minimal M_u de u. C'est à dire que $M_u = Q\,H$. On ne peut avoir $Q(u)(x) = 0$ pour tout $x \in E$ car sinon Q serait le polynôme minimal. Pour un certain x on a donc $y = Q(u)(x) \neq 0$. On a alors $H(u)(y) = H(u)[Q(u)(x)] = (HQ)(u)(x) = M_u(u)(x) = 0$. Ainsi le noyau $\ker H(u)$ est non trivial, c'est à dire non réduit à $\{0\}$.

Dans le cas complexe H s'écrit $H(X) = X - \lambda$ et $\ker H(u) = \ker(u - \lambda \mathrm{id})$ n'est autre que le sous-espace propre associé à la valeur propre λ. Il y a donc un vecteur propre associé à λ et celui-ci engendre une droite vectorielle qui est stable. Dans ce cas complexe nous aurions pu aussi raisonner sur le polynôme caractéristique qui admet forcément une racine.

Dans le cas réel H peut être de degré 1 et le raisonnement est identique au cas complexe. Mais H peut aussi être de degré 2 i.e. $H(X) = X^2 + \alpha\,X + \beta$. Prenons comme avant x tel que $H(u)(x) \neq 0$. Le sous-espace engendré par x et $u(x)$ est stable par u. Il suffit en effet de voir que $u(u(x))$ est combinaison linéaire de x et $u(x)$ ce qui se déduit facilement de $H(u)(x) = u(u(x)) + \alpha\,u(x) + \beta\,x = 0$. Le sous-espace engendré par x et $u(x)$ est évidemment de dimension au plus 2 et il est en fait de dimension 2. Car si il était de dimension 1 on aurait $u(x) = \lambda\,x$ et alors $H(u)(x) = H(\lambda)(x) = 0$ ce qui prouverait que H admet la racine réelle λ ce qui absurde.

4.3 Critère de diagonalisation

Théorème 7 *Un endomorphisme u est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

Preuve. En effet si $M_u(X) = (X - \lambda_1)\dots(X - \lambda_k)$ alors le lemme des noyaux donne $E = \ker 0 = \ker M_u(u) = \bigoplus_{i=1}^{k} \ker(u - \lambda_i\,\mathrm{id})$ ce qui prouve bien que u est diagonalisable.

Inversement si u est diagonalisable on a $E = \bigoplus_{\lambda \in \mathrm{sp}(u)} \ker(u - \lambda\,\mathrm{id})$. On pose alors $H(X) = \prod_{\lambda \in \mathrm{sp}(u)} (X - \lambda\,\mathrm{id})$. Pour tout $x \in E$ on utilise la décomposition $x = \sum_{\lambda \in \mathrm{sp}(u)} x_\lambda$ avec $x_\lambda \in \ker(u - \lambda\,\mathrm{id})$. Alors pour tout $\lambda \in \mathrm{sp}(u)$, $H(u)(x_\lambda) = \prod_{\lambda' \neq \lambda} (u - \lambda'\,\mathrm{id}) \circ (u - \lambda\,\mathrm{id})(x_\lambda) = 0$ d'où on déduit $H(u)(x) = 0$. On a donc bien $H(u) = 0$. Le polynôme minimal divise H donc est lui aussi scindé à racines simples. En fait on voit simplement $M_u = H$.

4.4 Le théorème de Cayley-Hamilton

Théorème 8 (Cayley-Hamilton) *Soit u un endomorphisme de l'espace de dimension finie E, C_u son polynôme caratéristique et M_u son polynôme* *minimal. Alors $C_u(u) = 0$ et donc $M_u \,|\, C_u$.*

Preuve. Soit $x \in E\backslash\{0\}$. Considérons l'ensemble des polynômes $\mathrm{Ann}(u, x) = \{P;\ P(u)(x) = 0\}$. C'est un idéal. Il est donc engendré par un polynôme $M_{u,x} = X^p + \alpha_{p-1}X^{p-1} + \cdots + \alpha_0$. Considèrons le sous-espace vectoriel F_x engendré par $x, u(x), \ldots, u^{p-1}(x)$. Il est facile de voir en utilisant la relation $u(u^{p-1}(x)) = u^p(x) = -\alpha_{p-1}u^{p-1}(x) - \cdots - \alpha_0\, x$ que F_x est stable par u. De plus par la minimalité du degré de $M_{u,x}$ dans $\mathrm{Ann}(u, x)$ on voit que $\{x, u(x), \ldots, u^{p-1}(x)\}$ est une famille libre et subséquemment une base de F_x. Ainsi on notera au passage que la dimension de F_x est égale à p qui est le degré de $M_{u,x}$. Complétons cette famille libre en une base $\mathcal{B}$ de E et écrivons la matrice de u dans cette base. Elle s'écrit par blocs sous la forme

$$\mathrm{mat}_{\mathcal{B}}\, u = \begin{pmatrix} A & ? \\ 0 & B \end{pmatrix}$$

Le bloc carré A a un nombre de lignes et de colonnes égal à $\dim F_x = p$. C'est la matrice de la restriction de u au sous-espace F_x écrite dans la base $\{x, u(x), \ldots, u^{p-1}(x)\}$. On voit que cette matrice A est

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & 0 & & 0 & -\alpha_1 \\ 0 & 1 & 0 & & 0 & -\alpha_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & 0 & -\alpha_{p-2} \\ 0 & \cdots & \cdots & 0 & 1 & -\alpha_{p-1} \end{pmatrix}$$

La forme par blocs de la matrice de u dans la base précisée et la définition du polynôme caractéristique montre que $C_u = C_A\, C_B$. Le polynôme caractéristique de A se calcule explicitement, par exemple en développant par rapport à la dernière colonne le déterminant de $X\,I - A$. On trouve

$$C_A(X) = X^p + \alpha_{p-1}\, X^{p-1} + \cdots + \alpha_0 = M_{u,x}$$

Ainsi $C_u(u)(x) = C_B(u) \circ M_{u,x}(u)(x) = 0$. Comme cela vaut pour tout $x \in E$ on a $C_u(u) = 0$ comme annoncé.

4.5 Décomposition en espaces caractéristiques

Le sous-espace stable

$$\mathrm{Car}(u, \lambda_i) = \ker\left[(u - \lambda_i\, \mathrm{id})^{\alpha_i}\right]$$

est appelé sous-espace *caractéristique* de u relatif à la valeur propre λ_i.

Supposons que le polynôme caractéristique C_u de l'endomorphisme u soit *scindé* c'est à dire $C_u(X) = \prod_{i=1}^{k}(X - \lambda_i)^{\alpha_i}$. Les $\lambda_1, \ldots, \lambda_k$ sont donc les valeurs propres de u. Alors le lemme des noyaux et le théorème de Cayley-Hamilton donnent

$$E = \bigoplus_{i=1}^{k} \ker\left[(u - \lambda_i \,\mathrm{id})^{\alpha_i}\right]$$

La restriction u_i de u à $\mathrm{Car}(u, \lambda_i)$ s'écrit donc $u_i = \lambda_i \,\mathrm{id} + n_i$ et $n_i^{\alpha_i} = 0$. On dit que n_i est *nilpotent*. On voit donc que quand le polynôme caractéristique est scindé la réduction d'un endomorphisme se ramène à la réduction d'un endomorphisme nilpotent.

4.6 Comparaison des polynômes caractéristique et minimal

Le théorème de Cayley-Hamilton nous dit déja que le polynôme caractéristique C_u d'un endomorphisme u de l'espace vectoriel de dimension finie E est un multiple du polynôme minimal M_u ce qui entraine que les diviseurs irréductibles de M_u sont des diviseurs irréductibles de C_u. Nous pouvons dire un peu plus

Proposition 9 *Soit u un endomorphisme d'un espace vectoriel de dimension finie E sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Le polynôme caractéristique C_u de u et son polynôme minimal M_u ont les mêmes diviseurs irréductibles. En particulier si $\mathbb{K} = \mathbb{C}$, ils s'écrivent*

$$C_u(X) = \prod_{\lambda \in sp(u)} (X - \lambda)^{m(\lambda)} \quad \textit{et} \quad M_u(X) = \prod_{\lambda \in sp(u)} (X - \lambda)^{r(\lambda)}$$
$$\textit{avec } 0 < r(\lambda) \leq m(\lambda) \quad \textit{et} \quad \sum_{\lambda \in sp(u)} m(\lambda) = \dim E$$

La preuve est l'objet de l'exercice 20.

4.7 Exercices

Exercice 16 On suppose qu'un espace vectoriel de dimension finie E se décompose comme somme directe de sous-espaces stables pour un endomorphisme u de E c'est à dire $E = F_1 \oplus F_2 \oplus \cdots \oplus F_k$ avec $u(F_i) \subset F_i$. Que valent le polynôme caractéristique et le polynôme minimal de u en fonction des polynômes minimaux et caractéristiques des restrictions de u aux sous-espaces $F_1, \ldots, F_k$?

Exercice 17 On considère la matrice $n \times n$

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

1) Calculer les puissances successives de M_0.
2) En déduire le polynôme minimal de M_0.
3) Quel est le polynôme minimal de M_λ ?
4) Montrer que M_λ est semblable à ${}^tM_\lambda$.

Exercice 18 Est il vrai que si A et B sont semblables alors $M_A = M_B$? Et la réciproque ? Si on a $M_A = M_B$ et $C_A = C_B$, les deux matrices A et B sont elles nécessairement semblables ?

Exercice 19 Soit $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$. Selon que l'on considère A comme une matrice coefficients réels ou complexes, le polynôme minimal est il le même ?

Exercice 20 Soit A une matrice carrée complexe.
1) Montrer que toute valeur propre est racine de tout polynôme annulateur de A.
2) Si le polynôme minimal de A s'écrit $M_A(X) = (X - \lambda)\, Q(X)$, montrer qu'il existe un vecteur v tel que $Q(A)\, v \neq 0$ et en déduire que λ est valeur propre de A.
3) En déduire que le polynôme minimal et le polynôme caractéristique ont les mêmes racines complexes .
4) En déduire que le polynôme minimal et le polynôme caractéristique ont les mêmes facteurs irréductibles.
5) Ce résultat persiste-t-il pour les matrices réelles ?

Exercice 21 Les affirmations suivantes sont elles vraies ou fausses :

$$a)\ \exists A \in \mathcal{M}_2(\mathbb{R}),\ M_A = X^2 + 1,$$

$$b)\ \exists B \in \mathcal{M}_3(\mathbb{C}),\ M_B = X^2 + 1,$$

$$c)\ \exists U \in \mathcal{M}_3(\mathbb{R}),\ M_U = X^2 + 1$$

Exercice 22 Soit $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{C})$. Parmi les assertions suivantes, quelles sont celles qui sont équivalentes à "A est nilpotente" :

$$(i)\ \mathrm{Sp}(A) = \{0\},\ \ (ii)\ C_A = X^n,\ \ (iii)\ A^n = 0,\ \ (iv)\ M_A | X^n$$

Exercice 23 On reprend la matrice $M(a,b,c)$ étudiée aux exercices 5 et 12 qui est la matrice avec des a en dessous de la diagonale, des b sur la diagonale, des c au dessus de la diagonale.
1) Calculer le polynôme minimal de $M(a,b)$.
2) Quand $M(a,b)$ est inversible, calculer son inverse.
3) Trouver, pour tout entier k, des constantes a_k et b_k telles que $M(a,b)^k = a_k\ M(a,b) + b_k\ I$

Exercice 24 1) Soit u un endomorphisme diagonalisable d'un espace vectoriel E de dimension finie et F un sous-espace stable pour u. Montrer que la restriction de u à F est diagonalisable.
2) Soit u et v deux endomorphismes d'un espace vectoriel qui commutent c'est à dire $u \circ v = v \circ u$. Montrer que tout sous-espace propre pour l'un est stable pour l'autre.
3) Montrer que si u et v sont deux endomorphismes diagonalisables de E et si $u \circ v = v \circ u$ alors il existe une base de E sur laquelle u et v diagonalisent.

Chapitre 5

Réduction de Jordan

5.1 Réduction en espaces cycliques pour un endomorphisme nilpotent

Commençons par définir la notion de sous-espace cyclique. Soit un endomorphisme u d'un $\mathbb{K}$-espace vectoriel de dimension finie E. Un vecteur v de E est *u-cyclique* de période r si $u^r(v) = 0$ avec r le plus petit entier naturel possible. On a déja dans la démonstration du théorème de Cayley-Hamilton étudié l'ensemble $\mathrm{Ann}(u, v)$ des polynomes P tels que $P(u)(v) = 0$. C'est un idéal de $\mathbb{K}[X]$ dont le générateur unitaire était noté $M_{u,v}$. On dit donc que v est u-cyclique de période r si $M_{u,v}(X) = X^r$. Alors il est immédiat de vérifier que la famille $\{v, u(v), \ldots, u^{r-1}(v)\}$ est libre. Le sous-espace

$$\langle\langle u, v\rangle\rangle = \mathrm{Vect}\{v, u(v), \ldots, u^{r-1}(v)\}$$

est un sous-espace vectoriel de E stable par u et de dimension r que l'on appellera *sous-espace cyclique*. Les matrices de la restriction de u à $\langle\langle u, v\rangle\rangle$ dans les bases $(v, u(v), \ldots, u^{r-1}(v))$ et $(u^{r-1}(v), \ldots, u(v), v)$ sont respectivement

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & 0 & & & \vdots \\ 0 & 1 & 0 & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ 0 & 0 & 0 & & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ \vdots & & & & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Cette seconde matrice $r \times r$ sera notée $M(r)$.

Théorème 10 *Soit u un endomorphisme de l'espace vectoriel E de dimension finie qui est nilpotent c'est à dire $u^r = 0$ pour un certain entier r. Alors*

$$E = F_1 \oplus \cdots \oplus F_k$$

avec

$$F_i = \langle\langle u, v_i \rangle\rangle = Vect(v_i, u(v_i), \ldots, u^{r_i-1}(v_i))$$

De plus la matrice de u sur la base

$$(u^{r_1-1}(v_1), u^{r_1-2}(v_1), \ldots, u(v_1), v_1, \ldots, u^{r_k-1}(v_k), \ldots, u(v_k), v_k)$$

est diagonale par blocs avec des blocs diagonaux qui sont respectivement $M(r_1), \ldots, M(r_k)$. Enfin, les entiers $k, r_1, \ldots, r_k$ sont déterminés par u.

Preuve. Nous allons procéder par récurrence sur la dimension de E. En dimension 0 (i.e. $E = \{0\}$) et 1, le résultat est vrai car le seul endomorphisme nilpotent est l'endomorphisme nul. Supposons le résultat acquis pour toutes les dimensions strictement inférieures à N et considérons un espace E de dimension N. Notons τ l'*indice de nilpotence* de u c'est à dire que l'entier τ vérifie $u^\tau = 0$ et qu'il est le plus petit qui convient. Notons que l'image de E par u, $\mathrm{Im}(u) = u(E)$ est un sous-espace vectoriel de E différent de E. En effet on a $u^{\tau-1} \neq 0$ donc il existe y tel que $u^{\tau-1}(y) \neq 0$. Cet y n'appartient pas à $u(E)$ car sinon on aurait $u^{\tau-1}(y) = 0$. Comme $u(E)$ est de dimension strictement plus petite que E on peut lui appliquer l'hypothèse de récurrence. Elle affirme que $u(E)$ est somme directe de sous-espaces cycliques :

$$u(E) = G_1 \oplus \cdots \oplus G_p$$

avec

$$G_i = \langle\langle u, w_i \rangle\rangle = \mathrm{Vect}(w_i, u(w_i), \ldots, u^{r_i-2}(w_i))$$

où $r_i \geq 2$. Autrement dit le vecteur w_i est supposé u-cyclique de période $r_i - 1$. On considère alors v_i tel que $u(v_i) = w_i$ et on pose

$$F_i = \langle\langle u, v_i \rangle\rangle = \mathrm{Vect}(v_i, u(v_i), \ldots, u^{r_i-1}(v_i))$$

Montrons dans un premier temps que la somme $\sum_{i=1}^{p} F_i$ est directe. Il suffit de montrer que 0 a une unique décomposition adaptée à cette somme. Supposons donc que

$$0 = P_1(u)(v_1) + \cdots + P_p(u)(v_p)$$

où P_i est un polynome. Car la forme générique d'un élément de F_i est $P_i(u)(v_i)$. Il s'agit de montrer que $P_1(u)(v_1) = \cdots = P_p(u)(v_p) = 0$. On fait agir u sur la décomposition de 0 ci-dessus ce qui donne $0 = P_1(u)(w_1) + \cdots + P_p(u)(w_p)$. Alors le caractère direct de la somme de G_i prouve que $0 = P_1(u)(w_1) = \cdots = P_p(u)(w_p)$. Cela implique que $X^{r_i-1} \mid P_i$ puisque w_i est u-cyclique de période $r_i - 1$. En particulier $X \mid P_i$ et on peut écrire $P_i = X\,Q_i$. On utilise cette forme des polynomes P_i dans la décompositon de 0 faisant intevenir les v_i. Cela donne $0 = Q_1(u)(w_1) + \cdots + Q_p(u)(w_p)$. On en déduit que $X^{r_i-1} \mid Q_i$ et donc que $X^{r_i} \mid P_i$ ce qui implique $P_i(u)(v_i) = 0$ comme désiré. On a donc établi que la somme des F_i est directe.

Posons $E' = F_1 \oplus \cdots \oplus F_p$. On a $u(E') = G_1 \oplus \cdots \oplus G_p = u(E)$. Considérons alors un vecteur y quelconque de E. Comme $u(E) = u(E')$ il existe $e' \in E'$ tel que $u(y) = u(e')$ d'où $y - e' \in \ker u$. Ainsi $E = E' + \ker(u)$. Le théorème de la base incomplète nous dit alors qu'on peut compléter une base de E' en une base de E en ajoutant des vecteurs $v_{p+1}, \ldots, v_k$ de $\ker(u)$. Un vecteur de $\ker u$ est cyclique de période 1. Les sous-espaces $F_{p+1} = \mathrm{Vect}(v_{p+1}), \ldots, F_k = \mathrm{Vect}(v_k)$ sont donc des sous-espaces cycliques et on a au total $E = F_1 \oplus \cdots \oplus F_p \oplus F_{p+1} \oplus \cdots \oplus F_k$ On a donc prouvé l'existence annoncée dans le théorème. Reste à voir que $k, r_1, \ldots, r_k$ sont uniquement déterminés par u. Pour cela il suffit de remarquer que pour tout p,

$$\dim u^{p+1}(E) - \dim u^p(E) = \mathrm{Card}\{i \in \{1, \ldots, k\}; r_i > p\}$$

5.2 Méthode pratique dans le cas nilpotent

Une première idée est d'utiliser la procédure mise en oeuvre dans la preuve par récurrence du paragraphe précédent. On a vu que $\dim u^p(E)$ décroit strictement jusqu'à ce que k atteigne l'indice de nilpotence τ de u. On peut donc appliquer la technique de passage de $u^p(E)$ à $u^{p-1}(E)$ vue ci-dessus pour des valeurs de p variant de τ à 1. Pour $p = \tau$, c'est facile ; il s'agit juste de trouver des vecteurs formant une base de $u^{\tau-1}(E)$. Ces vecteurs étant choisis, on leur cherche des antécédents dans $u^{\tau-2}(E)$ et on complète par les vecteurs de la restriction de $\ker u$ à $u^{\tau-2}(E)$. Et on continue d'itérer le procédé. Toutefois cette approche n'est pas très commode dans la pratique. Il est plus commode de travailler avec les noyaux plutôt que les images.

Pour les noyaux itérés, on a les inclusions

$$\{0\} \subset \ker u \subset \ker u^2 \subset \cdots \subset \ker u^{\tau-1} \subset \ker u^\tau = E$$

Prenons une base d'un supplémentaire de $\ker u^{\tau-1}$ dans E, que l'on note G_1. Ces vecteurs sont cycliques d'ordre τ. Leurs images engendrent $u(G_1)$ qui est inclus dans $\ker u^{\tau-1}$ mais qui vérifie $u(G_1) \cap \ker u^{\tau-2} = \{0\}$. On complète ces vecteurs de $u(G_1)$ pour former un supplémentaire G_2 de $\ker u^{\tau-2}$ dans $\ker u^{\tau-1}$. On forme une suite finie $G_1, \ldots, G_\tau$ de sous-espaces tels que $G_i \oplus \ker u^{\tau-i} = \ker u^{\tau-i+1}$ pour $i \in \{1, \ldots, \tau\}$ et $u(G_i) \subset G_{i+1}$. La base cherchée est obtenue en prenant les vecteurs utilisés à chaque étape pour compléter ainsi que tous les itérés non nuls de ces vecteurs.

5.3 Théorème de Jordan

Nous n'avons plus qu'à rassembler les résulats précédents pour obtenir un théorème puissant de réduction.

Théorème 11 (Jordan) *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie tel que son polynome caractéristique C_u est scindé. Alors il existe une base sur laquelle la matrice de u est diagonale par bloc avec des blocs diagonaux de Jordan c'est à dire de la forme*

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

où λ est une valeur propre de u. Cette forme est unique pour l'endomorphisme u, à permutation près des blocs de Jordan.

On notera qu'il peut y avoir plusieurs blocs de Jordan avec la même valeur propre λ.

Preuve. On décompose d'abord E en sous-espaces caractéristiques. Ensuite sur le sous-espace caractéristique associé à la valeur propre λ, u s'écrit comme $\lambda\,\mathrm{id}$ plus un endomorphisme nilpotent. On sait réduire un tel endomorphisme par la section précédente. Plus précisément on sait trouver une base sur-laquelle la matrice de l'endomorphisme nilpotent est diagonale par blocs avec des blocs de la forme

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

En réunissant les bases trouvées pour les sous-espaces caractéristiques on obtient une base de E qui convient. La méthode décrite qui fournit le résultat théorique est aussi celle que l'on doit mettre en oeuvre dans les calculs pratiques.

5.4 Calcul des puissances d'une matrice

Le théorème de Jordan est très utile pour calculer les puissances d'une matrice. Pour cela il faut faire trois remarques. La première remarque est que si M est diagonale par blocs alors les puissances sont encore diagonales par blocs :

$$\text{si } M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix} \text{ alors } M^n = \begin{pmatrix} A_1^n & 0 & \cdots & 0 \\ 0 & A_2^n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k^n \end{pmatrix}$$

La seconde remarque est qu'un bloc de Jordan s'écrit

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & & \vdots \\ 0 & 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda \end{pmatrix} = \lambda\, I + N \text{ avec } N = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

Comme I et N commutent on peut calculer les puissances par la formule du binome de Newton :

$$(\lambda\, I + N)^n = \sum_{j=0}^{n} \mathrm{C}_n^j \, \lambda^{n-j} \, N^j$$

Tout sera donc terminé si on sait calculer les puissances de N. Or $N^j = 0$ si j est plus grand que l'ordre de la matrice et N^2, N^3, ... sont obtenus "en déplacant l'alignement de 1 parallélement à la diagonale". Prenons l'exemple d'une matrice N d'ordre 6

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ; \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ;$$

$$N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ; \quad N^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ;$$

$$N^5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ; \quad N^j = 0 \text{ pour } j \geq 6$$

5.5 Exercices

Exercice 25 Soit a et b deux réels distincts et $A \in \mathcal{M}_6(\mathbb{R})$ de polynome minimal $M_A(X) = (X-a)^2\,(X-b)$ et de polynome caractéristique $C_A(X) = (X-a)^4\,(X-b)^2$. Quelles sont les réduites de Jordan possibles pour A ?

Exercice 26 effectuer la réduction de Jordan des matrices suivantes :

$$\begin{pmatrix} -15 & 12 & -9 & 1 \\ 29 & -24 & 19 & -2 \\ 60 & -49 & 38 & -4 \\ -6 & 5 & -4 & 1 \end{pmatrix} \begin{pmatrix} -1 & -2 & -3 & -3 & -3 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 3 \\ -1 & -1 & -1 & -1 & -2 \end{pmatrix}$$

Exercice 27 Montrer que si A et B sont deux matrices $n \times n$ avec $n < 4$ qui vérifient $M_A = M_B$ et $C_A = C_B$, alors les deux matrices A et B sont semblables.

Exercice 28 Montrer que toute matrice réelle ou complexe est semblable à sa transposée.

Exercice 29 (Partiel de novembre 2000, second des deux exercices) On note $W = (w_{ij})_{1\leq i,j\leq 3}$ la matrice 3×3 dont le terme général w_{ij} vaut 1 si $j = i+1$ et 0 sinon.
1.a) Ecrire explicitement J.
1.b) Trouver toutes les matrices M 3×3 telles que $M\,J = J\,M$.
1.c) L'ensemble des matrices M commutant avec J déterminé à la question précédente coincide-t-il avec l'ensemble des polynomes en J ?
2) Soit H une matrice 4×4 réelle dont le polynome minimal est $M_H(X) = (X-\alpha)^3\,(X-\beta)$ où α et β sont deux réels distincts. On note h l'endomorphisme de $\mathbb{R}^4$ qui a pour matrice H sur la base canonique de $\mathbb{R}^4$.
2.a) Quel est le polynome caractéristique de H ?
2.b) La matrice H est elle diagonalisable ?
2.c) La matrice H est elle trigonalisable ?
2.d) Justifier que $\ker[(h-\alpha\,\mathrm{id})^3]$ et $\ker[h-\beta\,\mathrm{id}]$ sont deux sous-espaces vectoriels stables par h supplémentaires dans $\mathbb{R}^4$ (c'est à dire que $\mathbb{R}^4$ est la somme directe de ces deux sous-espaces).
2.e) Quelle est la réduite de Jordan de h ?
3) On suppose que v est un endomorphisme de $\mathbb{R}^4$ tel que $v \circ h = h \circ v$.
3.a) Montrer que $\ker[(u-\alpha\,\mathrm{id})^3]$ et $\ker[u-\beta\,\mathrm{id}]$ sont stables par v.
3.b) Montrer que la matrice de la restriction de v à $\ker[(u-\alpha\,\mathrm{id})^3]$, dans une base bien choisie, commute avec J.
3.c) Montrer que $\mathcal{C} = \{v \in L(\mathbb{R}^4);\ \ v\circ h = h\circ v\}$ est un sous-espace vectoriel

de $L(\mathbb{R}^4)$.
3.d) Quelle est la dimension de $\mathcal{C}$?

Exercice 30 (Partiel de novembre 2002, second des deux exercices)
On considère dans un $\mathbb{K}$-espace vectoriel E de dimension 7 un endomorphisme u tel que, pour deux valeurs distinctes $\alpha, \beta \in \mathbb{K}$,

$$\dim \ker(u - \alpha \, \mathrm{id})^2 = 5, \quad \dim \ker(u - \alpha \, \mathrm{id}) = 3, \quad \dim \ker(u - \beta \, \mathrm{id}) = 2$$

On demande alors

1. de donner le polynôme minimal de u
2. de donner le polynôme caractéristique de u
3. de dire si u est diagonalisable
4. de dire si le théorème de Jordan s'applique à u et si oui de donner la réduite de Jordan
5. de décrire $\ker(u - \gamma \, \mathrm{id})^p$ pour $p \in \mathbb{N}^*$ et $\gamma \in \mathbb{K}$

On pourra répondre à ces questions dans l'ordre que l'on souhaite mais toute affirmation devra être argumentée (à l'aide des résultats du cours).

Existe-t-il un endomorphisme v de même polynôme minimal et caractéristique que u et qui n'est pas semblable à u ?

Chapitre 6

Formes quadratiques

6.1 Formes quadratiques

Nous travaillons sur un $\mathbb{R}$-espace vectoriel E. Une *forme bilinéaire symétrique* ϕ est une application de $E \times E$ dans $\mathbb{R}$ qui est linéaire par rapport à chacune de ses variables : pour $x, x', y, y' \in E$ et $\lambda \in \mathbb{R}$,

$$\phi(x + x', y) = \phi(x, y) + \phi(x', y),\ \phi(\lambda\, x, y) = \lambda\, \phi(x, y),$$

$$\phi(x, y + y') = \phi(x, y) + \phi(x, y'),\ \phi(x, \lambda\, y) = \lambda\, \phi(x, y)$$

et symétrique : $\phi(x, y) = \phi(y, x)$ pour tous $x, y \in E$. En fait, compte tenu de la symétrie, on peut se contenter d'énoncer la linéarité par rapport à une des deux variables. L'application q de E dans $\mathbb{R}$ définie par $q(x) = \phi(x, x)$ est appelée forme quadratique associée à ϕ. Les identités qui suivent, dites *identités de polarisation*, permettent alors de retrouver ϕ en fonction de q :

$$\begin{aligned} \phi(x, y) &= \frac{1}{2}(q(x + y) - q(x) - q(y)) \\ &= \frac{1}{4}(q(x + y) - q(x - y)) \end{aligned}$$

Supposons que le $\mathbb{R}$-espace vectoriel E est de dimension finie n et prenons une base $(b_1, \ldots, b_n)$. Si $x = \sum_{i=1}^n x_i\, b_i$ et $y = \sum_{i=1}^n y_i\, b_i$ alors on peut utiliser la bilinéarité pour obtenir une expression en fonction des coordonnées :

$$\phi(x, y) = \sum_{1 \le i,j \le n} x_i\, y_j\, \phi(b_i, b_j), \quad q(x) = \sum_{1 \le i,j \le n} x_i\, x_j\, \phi(b_i, b_j)$$

On peut écrire cela sous forme matricielle $\phi(x, y) = {}^t X\, Q\, Y$ où X est la matrice colonne des coordonnées de x, Y celle des coordonnées de y et Q est la matrice symétrique de terme général $\phi(b_i, b_j)$ appelée matrice de ϕ (ou de q) dans la base $(b_1, \ldots, b_n)$ et notée $\mathrm{mat}_{(b_i)}\phi$ ou $\mathrm{mat}_{(b_i)}q$. Bien sûr on a aussi $q(x) = {}^tX\,Q\,X$.

Nous pouvons alors nous poser la question du changement de base : si on considère une nouvelle base $(f_1, \ldots, f_n)$, quelle est la matrice $\tilde{Q}$ de ϕ sur cette base ? Notons P la matrice de passage de la base $(f_1, \ldots, f_n)$ à la base $(b_1, \ldots, b_n)$. Notons $\tilde{X}$ et $\tilde{Y}$ les matrices colonnes des composantes de x et y dans la base $(f_1, \ldots, f_n)$. On sait que l'on a $\tilde{X} = PX$, $\tilde{Y} = PY$ et $\phi(x, y) = {}^t\tilde{X}\tilde{Q}\tilde{Y} = {}^tX{}^tP\,\tilde{Q}\,PY = {}^tX\,Q\,Y$. Par identification, on a ainsi $Q = {}^tP\,\tilde{Q}\,P$. On dit parfois que les deux matrices $\tilde{Q}$ et Q sont *congruentes*. Redonnons la formule avec des notations plus explicites :

$$\mathrm{mat}_{(b_i)}\phi = {}^tP_{(f_i)\to(b_i)}\ \ \mathrm{mat}_{(f_i)}\phi\ \ P_{(f_i)\to(b_i)}$$

Notons que deux matrices congruentes sont équivalentes. Il en résulte en particulier qu'elles ont même rang. On peut donc définir le *rang* d'une forme quadratique comme le rang de sa matrice sur une base quelconque.

6.2 Réduction par la méthode de Gauss

Etant donnée une forme bilinéaire symétrique ϕ par exemple par sa matrice Q dans une certaine base, un des problèmes fondamentaux est de trouver une base où la matrice de ϕ est la plus simple possible, disons diagonale. Nous verrons que cela est possible, avec en plus une matrice de passage P d'un type particulier que nous appelerons orthogonale ce qui signifie ${}^tP = P^{-1}$. Il s'agit alors de trouver P telle que ${}^tPQP = P^{-1}QP$ soit diagonale. Le problème est donc ramené à la diagonalisation de la matrice symétrique Q et on retrouve un type de problème connu. Nous verrons cette méthode en détail dans un prochain chapitre.

L'algorithme de Gauss donne une autre méthode d'obtention d'une matrice P telle que ${}^tP\,Q\,P$ diagonale mais cette fois avec une matrice P qui est le plus souvent triangulaire. Nous résumons le résultat dans le théorème qui suit. La preuve par récurrence de ce théorème donne le principe de la méthode dite "de Gauss" qu'il faut appliquer itérativement dans les cas pratiques.

Théorème 12 (Sylvester) *Soit q forme quadratique sur le $\mathbb{R}$-espace vectoriel E de dimension n. Alors il existe des formes linéaires indépendantes $v_1, \ldots, v_n$ sur E telles que*

$$q(x) = \sum_{i=1}^{n} \alpha_i\ \ (v_i(x))^2\ \ \textit{avec}\ \alpha_i \in \{1, -1, 0\}$$

Le nombre de 1 et de -1 dépend de q et pas du choix des formes $v_1, \ldots, v_n$. Ces nombres de 1 et de -1 sont appelés signature de q.

Matriciellement, pour Q matrice symétrique, il existe une matrice P inversible telle que ${}^tP\,Q\,P$ soit diagonale avec des termes diagonaux égaux à 1,-1 ou 0. La preuve qui suit donne dans beaucoup de cas une matrice P triangulaire.

Preuve. Nous raisonnons par récurrence sur $n = \dim E$. Le cas $n = 1$ est trivial. Supposons donc le résultat acquis pour un espace de dimension inférieure ou égale à $n-1$ et considérons q forme quadratique sur E de dimension $n \geq 2$. Alors pour $x \in E$ se décomposant $x = \sum_i x_i\, b_i$ sur une certaine base fixée $(b_1, \ldots, b_n)$ de E et (q_{ij}) la matrice de q dans cette base, on a, en supposant d'abord $q_{11} \neq 0$,

$$\begin{aligned} q(x) &= \sum_{i,j} q_{ij}\, x_i\, x_j = q_{11}\, x_1^2 + 2\sum_{j>1} q_{1j}\, x_1\, x_j + \sum_{i,j\geq 2} q_{ij}\, x_i\, x_j \\ &= q_{11}\left[x_1 + \sum_{j>1} \frac{q_{1j}}{q_{11}} x_j\right]^2 + \sum_{i,j\geq 2} q_{ij}\, x_i x_j - q_{11}\left(\sum_{j>1} \frac{q_{1j}}{q_{11}} x_j\right)^2 \\ &= \text{signe}(q_{11})\; v_1^2(x) + \tilde{q}(x_2, \ldots, x_n) \end{aligned}$$

Dans cette dernière égalité

$$v_1(x) = \sqrt{|q_{11}|}\left[x_1 + \sum_{j>1} \frac{q_{1j}}{q_{11}} x_j\right]$$

est une forme linéaire sur E et

$$\tilde{q}(x_2, \ldots, x_n) = \sum_{i,j\geq 2} q_{ij}\, x_i\, x_j - q_{11}\left(\sum_{j>1} \frac{q_{1j}}{q_{11}} x_j\right)^2$$

est une forme quadratique sur $\mathbb{R}^{n-1}$ à laquelle on peut appliquer l'hypothèse de récurrence. Les formes linéaires $v_2, \ldots, v_n$ que l'on obtient alors sont indépendantes de v_1 car elles ne s'expriment qu'en fonction de $n-1$ dernières composantes alors que v_1 fait intervenir x_1.

Quitte à permuter les variables on voit que le procédé fonctionne dès qu'il existe j tel que $q_{jj} \neq 0$. Dans le cas contraire, on a $q_{jj} = 0$ pour tout j. Quitte à permuter encore les variables on peut supposer $q_{12} \neq 0$ et on écrit

$$\begin{aligned} q(x) &= 2q_{12}\, x_1 x_2 + 2\sum_{j\geq 3} q_{1j}\, x_1\, x_j + 2\sum_{j\geq 3} q_{2j}\, x_2\, x_j + \sum_{i,j\geq 3} q_{ij}\, x_i\, x_j \\ &= 2q_{12}\left(x_1 + \frac{1}{q_{12}}\sum_{j\geq 3} q_{2j} x_j\right)\left(x_2 + \frac{1}{q_{12}}\sum_{j\geq 3} q_{1j} x_j\right) + \tilde{q}(x_3, \ldots, x_n) \end{aligned}$$

où

$$\tilde{q}(x_3, \dots, x_n) = \sum_{i,j \geq 3} q_{ij}\, x_i\, x_j - \frac{2}{q_{12}} \left(\sum_{j \geq 3} q_{2j} x_j \right) \left(\sum_{j \geq 3} q_{1j} x_j \right)$$

On obtient alors

$$\begin{aligned} q(x) &= \frac{q_{12}}{2} \left(x_1 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{2j} x_j + x_2 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{1j} x_j \right)^2 \\ &- \frac{q_{12}}{2} \left(x_1 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{2j} x_j - x_2 - \frac{1}{q_{12}} \sum_{j \geq 3} q_{1j} x_j \right)^2 \\ &+ \tilde{q}(x_3, \dots, x_n) \end{aligned}$$

On peut appliquer l'hypothèse de récurrence à $\tilde{q}(x_3, \dots, x_n)$: il existe des formes linéaires indépendantes $w_3, \dots, w_n$ sur $\mathbb{R}^{n-2}$ telles que $\tilde{q}(x_3, \dots, x_n) = \sum_{i=3}^n \alpha_i \ (w_i(x_3, \dots, x_n))^2$ avec $\alpha_i \in \{1, -1, 0\}$. On pose

$$\begin{aligned} v_1(x) &= \sqrt{\frac{|q_{12}|}{2}} \left(x_1 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{2j} x_j + x_2 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{1j} x_j \right) \\ v_2(x) &= \sqrt{\frac{|q_{12}|}{2}} \left(x_1 + \frac{1}{q_{12}} \sum_{j \geq 3} q_{2j} x_j - x_2 - \frac{1}{q_{12}} \sum_{j \geq 3} q_{1j} x_j \right) \\ v_i(x) &= w_i(x_3, \dots, x_n) \end{aligned}$$

pour $i \geq 3$. Ces formes linéaires sont indépendantes et fournissent la représentation voulue :

$$q(x) = \operatorname{sign}(q_{12})\ v_1(x) + \operatorname{sign}(-q_{12})\ v_2(x) + \sum_{i=3}^{n} \alpha_i\ v_i(x)$$

Il reste à voir que le nombre de α_i égaux à 1 disons r_+ et le nombre de ceux égaux à -1 disons r_- ne dépend que de q. Cela va découler du fait que r_+ est le maximum des dimensions des sous-espaces vectoriels F de E tels que $\forall x \in F \setminus \{0\}, \ \ q(x) > 0$ et de même r_- est le maximum des dimensions des sous-espaces vectoriels F de E tels que $\forall x \in F \setminus \{0\}, \ \ q(x) < 0$. En effet considérons la base $(f_1, \dots, f_n)$ de E sur laquelle la matrice de q est diagonale avec r_+ 1 et r_- -1 puis des 0 sur la diagonale. On peut prendre comme telle base celle où les composantes d'un vecteur x sont donnés par $v_1(x), \dots, v_n(x)$. En termes sophistiqués, cette base est celle dont la base duale est $v_1, \ldots, v_n$. Supposons qu'il existe F tel que $\forall x \in F \setminus \{0\}$, $q(x) > 0$. Alors $\mathrm{Vect}(f_{r_+ +1}, \ldots, f_n)$ et F sont en somme directe. Cela se voit en évaluant q sur un vecteur de l'intersection. La dimension de leur somme (directe), est égale à la somme de leurs dimensions et elle est inférieure à la dimension de E : $(n - r_+) + \dim F \leq n$ c'est à dire $\dim F \leq r_+$. Dans cette inégalité on réalise l'égalité en prenant $F = \mathrm{Vect}(f_1, \ldots, f_{r_+})$. La caractérisation de r_+ est donc prouvée et celle de r_- est similaire.

6.3 Signatures particulières

On dit qu'une forme quadratique q et sa forme bilinéaire associée sont *positives* si $r_- = 0$ i.e. $q(x) \geq 0$ pour tout $x \in E$. On dit qu'une forme quadratique q et sa forme bilinéaire associée sont *définies positives* si signature est $(r_+, r_-) = (\dim E, 0)$ ou ce qui est équivalent si $q(x) > 0$ pour tout $x \in E \setminus \{0\}$. On dit qu'une matrice $n \times n$ symétrique A est positive si c'est la matrice d'une forme quadratique positive c'est à dire ${}^tX\,AX \geq 0$ pour toute matrice colonne X autrement dit $\sum_{i,j} a_{ij}\; x_i\, x_j \geq 0$ pour tous réels $x_1, \ldots, x_n$. On dit qu'une matrice $n \times n$ symétrique A est définie positive si c'est la matrice d'une forme quadratique définie positive c'est à dire ${}^tX\,AX > 0$ pour toute matrice colonne X non nulle autrement dit $\sum_{i,j} a_{ij}\; x_i\, x_j > 0$ pour tous réels $x_1, \ldots, x_n$ non tous nuls.

6.4 Cas des dimensions 2 et 3

Dans le cas des dimensions 2 et 3 la réduction de Gauss est particulièrement facile à mettre en oeuvre pour déterminer la signature d'une forme quadratique. Les lignes de niveau des formes quadratiques donnent des ensembles remarquables : les coniques en dimension 2 et certaines quadriques en dimension 3. Commençons par la dimension 2.

Les ensembles de points (x, y) du plan d'équation cartésienne $q((x, y)) = C$ où q est une forme quadratique sont appelés coniques du plan. Le type de conique dépend de la signature de q. Laissons tomber les cas dégénérés et intéressons nous seulement aux cas où le rang est 2. Alors il n'y a que 3 possibilités pour la signature : (2,0) cas défini positif, (1,1), (0,2) cas défini négatif. Dans le premier cas, ou le troisième par changement de signe, quitte à se placer dans une bonne base (dont on verra plus tard qu'elle peut être prise orthogonale) une ligne de niveau de q a typiquement l'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

qui est l'équation d'une ellipse. Le cas particulier $a = b$ est celui du cercle. Dans le cas de signature $(1, 1)$, on obtient l'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

qui est celle d'une hyperbole.

Dans le cas de la dimension 3, supposons encore le rang maximal sinon on obtient un cylindre (au sens mathématique) basé sur une conique : cylindre elliptique ou hyperbolique. Il a essentiellement trois cas après choix d'une bonne base, pour l'équation de lignes de niveau (non dégénérées)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

qui désignent respectivement l'ellipsoïde, l'hyperboloïde à une nappe et l'hyperboloïde à deux nappes.

En fait une quadrique est un ensemble de l'espace affine de dimension 3 d'équation $q(x, y, z) + L(x, y, z) = C$ où L est une forme linéaire. Dans le cas où q est de rang 3, la présence de L n'introduit pas de nouveau type. Il y a les 3 décrits précédemment plus les cas dégénérés : vide, point, cone. Dans le cas du rang 2 signalons deux nouveaux types d'équations typiques

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0,\ \frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$$

paraboloide elliptique et paraboloide hyperbolique. Pour le rang 1, on obtient le cylindre parabolique : $x^2/a^2 - 2p\,y = 0$. Nous avons laissé de côté pour les rangs 1 et 2 les cas dégénérés : vide, droite, deux plans, un plan.

La classification des quadriques dépassant le cadre de ce cours nous nous contenterons de cette petite introduction.

6.5 Exercices

Exercice 31 On considère la forme quadratique q sur $\mathbb{R}^3$ donnée par

$$q((x,y,z)) = x^2 + 6\,y^2 + 6\,z^2 + 4\,xy - 2\,xz$$

1) Quelle est la matrice de q dans la base canonique ?
2) Quelle est la forme bilinéaire symétrique associée ?
3) Appliquer l'algorithme de Gauss à la forme quadratique q.
4) Quelle est la signature de q ?
5) Donner l'écriture matricielle qui découle des questions précédentes.
6) Quelle est la nature de l'ensemble d'équation $q((x,y,z)) = 1$?

Exercice 32 Reprendre l'exercice précédent avec

$$q((x,y,z)) = xy + yz + xz$$

Exercice 33 Pour une matrice symétrique le fait d'avoir tous ses coefficients positifs est il une condition suffisante pour que la matrice soit positive ? Est ce une condition nécessaire ?

Exercice 34 Montrer que la matrice $n \times n$ de terme général $1/(i+j+1)$ est symétrique définie positive.

Exercice 35 On considère des réels $0 < t_1 < \cdots < t_n$. Montrer que la matrice de terme général $\inf(t_i, t_j)$ est définie positive.

Exercice 36 Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices symétriques positives. Montrer que la matrice de terme général $c_{ij} = a_{ij}\,b_{ij}$ est symétrique positive.

Chapitre 7

Structure euclidienne

7.1 Géométrie des espaces euclidiens

Un espace euclidien est un $\mathbb{R}$-espace vectoriel E muni d'un produit scalaire, c'est à dire d'une forme bilinéaire symétrique définie positive. Pour le moment nous n'avons pas besoin de le supposer de dimension finie. On notera $\langle \cdot, \cdot \rangle$ ce produit scalaire. On pose pour tout $x \in E$, $||x|| = \sqrt{\langle x, x \rangle}$. On verra très bientôt par l'inégalité de Minkovski que cette notation est cohérente avec la topologie car il s'agit effectivement d'une norme. Notons que pour tous $x, y \in E$, on a immédiatement $||x+y||^2 = ||x||^2 + ||y||^2 + 2\langle x, y \rangle$. Voyons une inégalité importante.

Proposition 13 (Inégalité de Cauchy-Schwarz)

$$\forall x, y \in E, \ \ |\langle x, y \rangle| \leq ||x||\,||y||$$

La preuve est très simple. Il suffit de remarquer que pour tout t réel, $0 \leq ||x + ty||^2 = t^2\ ||x||^2 + 2t\langle x, y \rangle + ||y||^2$. Mais pour qu'un trinome soit de signe constant, il faut que son discriminant soit négatif ou nul ce qui donne $(2\langle x, y \rangle)^2 - 4\,||x||^2\ ||y||^2 \leq 0$ comme désiré. Le cas d'égalité a lieu si et seulement si le discriminant est nul. Dans ce cas le trinome considéré a une racine réelle c'est à dire qu'il existe un réel t tel que $||x+ty|| = 0$ i.e. $x = -ty$.

Un corollaire est le résultat suivant qui prouve que la “norme” que nous avons défini vérifie l'inégalité triangulaire et qu'elle est donc une norme au sens du cours de topologie.

Proposition 14 (Inégalité de Minkovski)

$$\forall x, y \in E, \ \ ||x + y|| \leq ||x|| + ||y||$$

Preuve. On développe le carré de la norme puis on utilise l'inégalité de Cauchy-Schwarz.

$$\begin{aligned} ||x+y||^2 &= ||x||^2 + ||y||^2 + 2\langle x, y\rangle \\ &\leq ||x||^2 + ||y||^2 + 2\,||x||\;||y|| \\ &\leq (||x|| + ||y||)^2 \end{aligned}$$

Les espaces euclidiens bénéficient de la notion géométrique d'orthogonalité. Deux vecteurs sont *orthogonaux* si leur produit scalaire vaut 0. En particulier si x et y sont orthogonaux, on a simplement $||x+y||^2 = ||x||^2 + ||y||^2$ ce qui constitue le *théorème de Pythagore*. Pour A partie de E, l'orthogonal est l'ensemble des vecteurs orthogonaux à tous ceux de A.

$$A^\perp = \{y \in E,\ \ \forall a \in A,\ \langle a, y\rangle = 0\}$$

C'est un sous-espace vectoriel de E. Il est immédiat que si $A \subset B$ alors $B^\perp \subset A^\perp$. Deux sous-espaces vectoriels F et G sont orthogonaux si tout vecteur du premier est orthogonal à tout vecteur du second i.e. $F \subset G^\perp$ ou $G \subset F^\perp$. On dit qu'une famille $\{b_i,\ i \in I\}$ de vecteurs de l'espace euclidien E est *orthogonale* si $\langle b_i, b_j\rangle = 0$ pour tous $i \neq j$. On dit que la famille est *orthonormale* ou *orthonormée* si on a en plus $||b_i|| = 1$ pour tout i. On vérifie très facilement qu'un famille orthonormale est libre.

Proposition 15 *Soit F sous-espace vectoriel de dimension finie du $\mathbb{R}$-espace vectoriel euclidien E. On a alors*

$$E = F \oplus F^\perp$$

La projection p sur F parallélement à $F^\perp$ est appelée projection orthogonale *sur F. Pour $(b_1, \ldots, b_k)$ base orthonormale de F, la projection p est donnée, pour tout $x \in E$, par*

$$p(x) = \sum_{i=1}^{k} \langle b_i, x\rangle\ \ b_i$$

Preuve. Si $x \in F \cap F^\perp$ on a $\langle x, x\rangle = 0$ ce qui entraine $x = 0$ prouvant ainsi que la somme considérée est directe. Si on définit $p(x)$ comme le vecteur de F donné par la formule de l'énoncé, on a $\langle p(x), b_j\rangle = \langle b_j, x\rangle$ en utilisant l'orthonormalité de $(b_1, \ldots, b_k)$. On a donc pour tout j , $\langle x - p(x), b_j\rangle = 0$ ce qui prouve que $x - p(x)$ est orthogonal à F. L'écriture $x = p(x) + (x - p(x))$ montre que tout vecteur est somme d'un élément de F et d'un vecteur perpendiculaire à F, prouvant ainsi le résultat annoncé. Le point éclipsé dans cette preuve est l'existence d'une base orthonormale pour F. Il suffira d'attendre la proposition 16 du prochain paragraphe.

Dans le paragraphe que nous venons de traiter, la dimension de l'espace considéré peut être infinie. La plus grande partie du cours s'interesse néanmoins au cas de la dimension finie. On notera que dans certains livres le terme "euclidien" est réservé aux espaces de dimension finie.

Le cas générique d'espace euclidien de dimension finie est $\mathbb{R}^d$ muni de sa *structure euclidienne canonique* c'est à dire celle qui rend orthonormale la base canonique, autrement dit, le produit scalaire est donné par

$$\forall x = (x_1, \dots, x_d),\ \forall y = (y_1, \dots, y_d),\ \ \langle x, y\rangle = \sum_{i=1}^{d} x_i\, y_i$$

7.2 La méthode de Gram-Schmidt

Si $(u_1, u_2, u_3, \dots)$ est une famille libre, il existe une famille orthonormale $(b_1, b_2, b_3, \dots)$ qui vérifie $\mathrm{Vect}(u_1, \dots, u_p) = \mathrm{Vect}(b_1, \dots, b_p)$ pour tout entier $p \geq 1$. Il suffit de poser d'abord $b_1 = (1/||u_1||)\ u_1$. Ensuite supposons $b_1, \dots, b_{p-1}$ choisis ($p \geq 2$) de façon voulue. Le vecteur $w_p = u_p + \alpha_1\, b_1 + \cdots + \alpha_{p-1}\, b_{p-1}$ peut être rendu orthogonal à $b_1, \dots, b_{p-1}$. Il suffit de prendre $\alpha_j = -\langle u_p, b_j\rangle$. Ce choix étant fait, on remarque que w_p est non nul. En fait w_p est la différence de u_p et de sa projection orthogonale sur $\mathrm{Vect}(b_1, \dots, b_{p-1}) = \mathrm{Vect}(u_1, \dots, u_{p-1})$. Finalement on normalise en posant enfin $b_p = (1/||w_p||)\ w_p$.

Ce procédé s'appelle *procédé d'orthonormalisation de Gram-Schmidt*. Une conséquence est le résultat suivant.

Proposition 16 *Tout espace euclidien de dimension finie ou tout sous-espace de dimension finie d'un espace euclidien admet une base orthonormale.*

Preuve. Il suffit de considérer une base quelconque et de lui appliquer le procédé de Gram-Schmidt. En fait, sauf en dimension 1, il y a une infinité de telles bases orthonormales, puisqu'il y a une infinité de choix pour la direction du premier vecteur. Notons que le procédé de Gram-Schmidt peut également être appliqué à une suite de vecteurs donc le résultat est vrai pour un espace admettant une base dénombrable.

Voyons maintenant une conséquence matricielle.

Proposition 17 *Pour toute matrice symétrique définie positive A il existe une unique matrice R triangulaire supérieure à coefficients diagonaux strictement positifs vérifiant $A = {}^tR\,R$*

Preuve. Notons q la forme quadratique sur $\mathbb{R}^n$ de matrice A dans la base canonique. La matrice R vérifie $A = {}^tR\,R$ si et seulement si R^{-1} est la matrice de passage de la base canonique $(e_1, \dots, e_n)$ à une base $(f_1, \dots, f_n)$ où la matrice de q est I c'est à dire une base orthonormale. On demande que la matrice R soit triangulaire supérieure à coefficients diagonaux strictement positifs ce qui revient au même que de demander que R^{-1} soit triangulaire supérieure à coefficients diagonaux strictement positifs. La base orthonormale cherchée $(f_1, \dots, f_n)$ s'obtient donc à partir de la base canonique $(e_1, \dots, e_n)$ de façon que, pour tout p, $f_p = r_{p,p}\, e_p + r_{p-1,p}\, e_{p-1} + \cdots + r_{1,p}\, e_1$. Dans cette égalité $r_{p,p}, r_{p-1,p}, \dots, r_{1,p}$ sont les coefficients de la p-ième colonne de R^{-1} avec l'exigence que $r_{p,p} > 0$. On voit donc que $(f_1, \dots, f_n)$ s'obtient donc à partir de $(e_1, \dots, e_n)$ par le procédé de Gram-Schmidt et que c'est la seule façon. Il y a donc existence et unicité de la base $(f_1, \dots, f_n)$. La matrice R^{-1} et donc R existe et est unique ce qui achève la preuve.

7.3 Endomorphismes orthogonaux

Dans un espace euclidien on dispose d'une notion de distance donnée par la norme. Il paraît intéressant de se pencher sur les endomorphismes qui préservent cette norme.

On dira qu'un endomorphisme u d'un espace euclidien de dimension finie est *orthogonal* si et seulement si $||u(x)|| = ||x||$ pour tout $x \in E$.

Quelques remarques s'imposent. D'abord cette définition n'est pas vide de contenu car il y a au moins l'identité (et $-$id) qui est orthogonale. La définition montre qu'un endomorphisme orthogonal est injectif donc bijectif, que son inverse est orthogonal et que la composée de deux endomorphismes orthogonaux est orthogonal. Ainsi l'ensemble des endomorphismes orthogonaux est un *sous-groupe* du groupe linéaire de E. On le note $\mathcal{O}(E)$. Remarquons que grâce aux identités de polarisation, un endomorphisme orthogonal conserve le produit scalaire : $\langle u(x), u(y)\rangle = \langle x, y\rangle$ pour tous $x, y \in E$. En particulier l'image par un endomorphisme orthogonal de toute base orthonormale est une base orthonormale. Réciproquement, si une endomorphisme transforme une certaine base orthonormale (e_i) en une autre base orthonormale, c'est un endomorphisme orthogonal. En effet si $x = \sum_i x_i\, e_i$ est la décomposition d'un vecteur quelconque $x \in E$ sur la base orthonormale (e_i)

alors on a

$$||u(x)||^2 = \left\|\sum_i x_i\, u(e_i)\right\|^2 = \sum_i x_i^2 ||u(e_i)||^2 = \sum_i x_i^2 = \left\|\sum_i x_i\, e_i\right\|^2 = ||x||^2$$

en utilisant deux fois le théorème de Pythagore.

Passons maintenant à l'aspect matriciel. Une matrice carrée réelle M est dite *orthogonale* si elle vérifie ${}^tM\, M = I$. Par conséquent cette matrice a pour déterminant 1 ou -1 et en particulier est inversible. L'ensemble des matrices réelles $n \times n$ orthogonales est un sous-groupe du groupe des matrices inversibles et il est noté $\mathcal{O}(n)$. Dans $\mathcal{O}(n)$ l'ensemble des matrices de déterminant 1 forme un sous-groupe noté $\mathcal{O}^+(n)$.

Proposition 18 *Une matrice est orthogonale si et seulement si c'est, pour un certain produit scalaire, la matrice de passage d'une base orthonormale à une autre base orthonormale.*

Preuve. Supposons d'abord que dans un espace euclidien, on dispose d'une base orthonormale et d'une autre base telle que la matrice de passage de la première base à la seconde base est une matrice orthogonale P. Alors la matrice du produit scalaire dans la seconde est ${}^tP\, I\, P = I$ ce qui montre bien qu'elle est orthonormée. Inversement si P est la matrice de passage d'une base orthonormale à une autre base orthonormale dans un certain espace euclidien alors la matrice du produit scalaire dans chacune de ces bases est I et la relation de passage entre ces deux matrices donne ${}^tP\, I\, P = I$ prouvant ainsi que la matrice de passage est orthogonale.

Proposition 19 *Un endomorphisme est orthogonal si et seulement si sa matrice sur une base orthonormale est orthogonale.*

Preuve. La matrice d'un endomorphisme bijectif dans une base peut être vu comme la matrice de passage de cette base à la base image par cet automorphisme. L'endomorphisme est orthogonal si et seulement si il transforme une base orthonormale en une autre base orthonormale donc si et seulement si la matrice de passage– qui est la matrice de l'endomorphisme– est orthogonale.

7.4 Exemples d'endomorphismes orthogonaux

Quels exemples pouvons nous donner d'endomorphisme orthogonal ? Commençons par la dimension 2. Une condition nécessaire et suffisante est que sa matrice dans une base orthonormale (e_1, e_2) soit orthogonale ce qui conduit à l'équation :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = I \text{ soit } \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

L'égalité des coefficients diagonaux donne $a = \cos\alpha$, $c = \sin\alpha$ puis $b = \cos\beta$, $d = \sin\beta$. Alors l'égalité $ab+cd = 0$ s'écrit $\cos(\alpha-\beta) = 0$ donc $\beta = \alpha \pm (\pi/2)$ modulo 2π ce qui donne les matrices

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ ou } \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

Le premier type correspond à la rotation d'angle α. Ces matrices donnent d'ailleurs une paramétrisation de $\mathcal{O}^+(2)$. Le second type correspond à des matrices de déterminant -1. En particulier la matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est la matrice de la symétrie orthogonale selon $\mathrm{Vect}(e_1)$. Toute matrice du second type est produit de cette matrice par une matrice de rotation.

Donnons maintenant des exemples dans un espace E de dimension quelconque. Prenons un sous-espace F. Comme $E = F \oplus F^\perp$, tout $x \in E$ se décompose de manière unique $x = f + f^\perp$ selon cette décomposition. L'application qui à x associe f est la projection orthogonale sur F. L'application qui à x associe $f^\perp$ est la projection orthogonale sur $F^\perp$. Ces deux projections ne sont pas des endomorphismes orthogonaux. En revanche l'application qui à x associe $f - f^\perp$ est orthogonale ; c'est la symétrie orthogonale par rapport à F. Si on choisit une base orthonormale de E adaptée à la décompostion $E = F \oplus F^\perp$ ce qui signifie que les premiers vecteurs forment une base orthonormale de F et les suivants une base orthonormale de $F^\perp$ alors la matrice de la symétrie orthogonale par rapport à F est dans cette base

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

Dans le prochain chapitre nous ferons la réduction des automorphismes orthogonaux et nous verrons que la réduction ramène aux deux types que nous venons d'étudier : rotation vectorielle plane et symétrie orthogonale.

7.5 Décompositions et systèmes linéaires

Le problème majeur de l'analyse numérique matricielle est la résolution de systèmes linéaires. Il s'agit, une matrice $n \times n$ inversible A et un vecteur b étant donnés, de trouver le vecteur x tel que $A\,x = b$.

Notons que le calcul de l'inverse de A revient à la résolution de n tels systèmes car les colonnes $C_1\ C_2 \ldots C_n$ de A^{-1} sont solutions des systèmes $AC_i = e_i$ où e_i est le i-ième vecteur de la base canonique. Ainsi il est interessant de calculer l'inverse A^{-1} si il faut résoudre un nombre de systèmes grand devant la taille de la matrice intervenant. Pour un petit nombre de systèmes à résoudre le calcul de A^{-1} n'est pas nécessaire.

Il y a essentielement deux types de méthodes pour résoudre les systèmes linéaires. Les premières sont les méthodes exactes. Ces méthodes fourniraient le résultat exact si les ordinateurs menaient les calculs avec une précision infinie. C'est possible par exemple dans le cas de systèmes à coefficients rationnels ou avec un logiciel de calcul formel pour des systèmes de petite taille. Un exemple de telle méthode est la méthode du pivot de Gauss que l'on utilise pour résoudre à la main de petits systèmes. Le second type de méthodes de résolution des systèmes consiste en les méthodes itératives qui fournissent des approximations de la solution, approximations qui convergent vers la solution quand le nombre de pas d'itérations tend vers $+\infty$. Nous reparlerons de ce second type de méthodes dans un prochain chapitre et nous allons pour le moment décrire le principe de méthodes du premier type. Nous pouvons déja revenir sur la proposition 17. Nous la rappellons.

Proposition 20 (Décomposition de Cholesky) *Pour toute matrice symétrique définie positive A il existe une unique matrice R triangulaire supérieure à coefficients diagonaux strictement positifs vérifiant $A = {}^tR\,R$*

On voit son intérêt pour la résolution du système linéaire de matrice A : pour résoudre $A\,x = b$ on résout successivement ${}^tRy = b$ puis $R\,x = y$ qui sont deux systèmes triangulaires, le premier triangulaire inférieur et le second triangulaire supérieur. La question qui reste est : comment trouver la matrice R à partir de A ? Simplement par identification. L'équation $A = {}^tR\,R$ s'écrit sur les termes généraux $a_{ij} = \sum_{k \leq \inf(i,j)} r_{ki}\, r_{kj}$. On voit que les coefficients r_{ij}; $i \leq j$ se trouvent les uns après les autres en exprimant l'équation précédente pour a_{11} puis a_{12}, a_{22} etc ...

Cette méthode se généralise d'une certaine manière à presque toute matrice inversible à travers la décomposition dite L-U (de l'anglais lower-upper) appelée aussi parfois méthode de Crout.

Proposition 21 *Une matrice inversible $A(a_{ij})_{1 \leq i,j \leq n}$ s'écrit $A = L\,U$ où*

L est triangulaire inférieure à diagonale unité et U triangulaire supérieure si et seulement si tous les déterminants $\det[(a_{ij})_{1\leq i,j\leq r}]$ *avec $r \in \{1, \ldots, n\}$ appelés mineurs principaux sont non nuls.*

Preuve. Nous n'allons pas faire la démonstration compléte qui figure mieux à sa place dans tout cours d'analyse numérique matricielle. Fondamentalement ce résultat est une traduction de l'algorithme du pivot de Gauss bien connu. Car, pour trigonaliser un système par la méthode du pivot de Gauss, on procède à des opérations élémentaires sur les lignes qui peuvent en fait être vues comme des produits par des matrices élémentaires que l'on appelle dilatations (qui sont diagonales) et transvections (qui sont triangulaires). Parmi ces opérations il y a l'échange de deux lignes rendue nécessaire pour passer à un pivot non nul. C'est précisement ce qui n'arrive pas quand tous les mineurs principaux sont non nuls. Dans ce cas on voit en explicitant les matrices de transvection et de dilatations qui interviennent que l'on a au total multiplié A par un produit de ces matrices de transvections ou dilatations qui se trouve être une matrice triangulaire $\tilde{L}$ inférieure de diagonale unité. On a donc $\tilde{L}\,A = U$ triangulaire d'ou le résultat en inversant $\tilde{L}$.

La méthode de détermination pratique de L et U –qui serait d'ailleurs un bon moyen de prouver l'existence de la décomposition– est tout simplement d'identifier ce qui permet de déterminer de manière échelonnée les coefficients de L et U. Bien sûr l'intérêt de la décomposition L-U est de ramener la résolution du système linéaire de matrice A à deux systèmes triangulaires.

Voyons maintenant une nouvelle méthode basée, elle aussi, sur une décomposition matricielle.

Proposition 22 (Décomposition de Householder) *Pour toute matrice inversible M il existe un unique couple (O, T) constitué d'une matrice orthogonale O et d'une matrice T triangulaire supérieure à coefficients diagonaux strictement positifs et qui vérifie $M = O\,T$.*

Preuve. Si une telle décomposition a lieu, on a ${}^tM\,M = {}^tT\,T$. La matrice ${}^tM\,M$ étant symétrique définie positive, l'existence et l'unicité d'une matrice T vérifiant l'égalité précédente résulte donc de la proposition 17. Ensuite il suffit de poser $O = M\,T^{-1}$ et on vérifie qu'il s'agit bien d'une matrice orthogonale.

Pour la résolution des systèmes linéaires et en fait pour inverser une matrice M, on voit l'intérêt de la décompostion $M = O\,T$ car T est triangulaire donc facile à inverser et O est orthogonale donc l'inverse est simplement la transposée.

7.6 Exercices

Exercice 37 1) Montrer qu'un endomorphisme trigonalisable d'un espace euclidien est trigonalisable dans une base orthonormée.
2) Montrer que si A est une matrice dont le polynome caractéristique est scindé alors il existe une matrice P telle que ${}^tP\,A\,P$ est triangulaire supérieure et ${}^tP\,P = I$

Exercice 38 Soit S une matrice réelle $n \times n$ symétrique positive. Montrer que la matrice de ses cofacteurs est symétrique positive.

Exercice 39 (Inégalité d'Hadamard) Soit A une matrice réelle $n \times n$. L'espace $\mathcal{M}_{n,1}(\mathbb{R})$ des matrices colonnes réelles est muni de son produit scalaire canonique : pour $X = (x_{i1})_{1\leq i\leq n}$ et $Y = (y_{i1})_{1\leq i\leq n}$ matrices colonnes réelles, le produit scalaire vaut $\sum_i x_{i1}\, y_{i1}$. La norme associée est notée $||\cdot||$.
1) On suppose d'abord les colonnes $A_1, \ldots, A_n$ de A orthogonales au sens donné ci-dessus. Montrer que

$$|\det A| \leq ||A_1|| \ldots, ||A_n||$$

2) Montrer que le résultat vaut dans le cas général. On se raménera au cas des colonnes orthogonales en appliquant le procédé de Gram-Schmidt.

Exercice 40 1) Utiliser la proposition 17 pour montrer que si $A = (a_{ij})_{1\leq i,j\leq n}$ est une matrice symétrique définie positive alors

$$\det A \leq \prod_{i=1}^{n} a_{ii}$$

2) Retrouver l'inégalité d'Hadamard obtenue à l'exercice 39 en appliquant le résultat qui précède à $A = {}^tM\,M$ où M est une matrice inversible quelconque.

Exercice 41 Soit A une partie de l'espace vectoriel euclidien E de dimension finie.
1) Montrer que $A^\perp = \mathrm{Vect}(A)^\perp$ puis $A \cap A^\perp = \{0\}$
2) Montrer que $A^{\perp\perp} = \mathrm{Vect}(A)$.
3) Montrer que pour F et G sous-espaces vectoriels de E,

$$(F+G)^\perp = F^\perp \cap G^\perp, \;\; (F\cap G)^\perp = F^\perp + G^\perp$$

Exercice 42 1) Montrer que l'on définit un produit scalaire sur $\mathcal{M}_n(\mathbb{K})$ en posant $\langle A, B\rangle = \mathrm{Tr}({}^t\overline{A}\,B)$.
2) Pour $\mathbb{K} = \mathbb{R}$, quels sont les orthogonaux de $\mathcal{D}_n$, $\mathcal{S}_n$, $\mathcal{A}_n$ qui désignent respectivement l'ensemble des matrices diagonales, symétriques, antisymétriques ?

Exercice 43 (Matrices de Gram) *(Examen de janvier 2004)* **1)** On définit une matrice carrée M par blocs :

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

où A est une matrice carrée inversible.
1.a) Que vaut le produit de matrices suivant, les matrices étant écrites par blocs :

$$\begin{pmatrix} A & 0 \\ C & I \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

1.b) En déduire une expression de $\det M$.
2) Soit $X \in \mathcal{M}_{n,p}(\mathbb{R})$. Montrer que les matrices X et ${}^tX\ X \in \mathcal{M}_p(\mathbb{R})$ ont même rang. On pourra par exemple comparer les noyaux de X et ${}^tX\ X$.
3) On note E un $\mathbb{R}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.
3.a) Rappeler ce qu'est un produit scalaire.

Pour $x_1, \dots, x_p$ vecteurs de E, on appelle matrice de Gram de $(x_1, \dots, x_p)$ et on note $G(x_1, \dots, x_p)$ la matrice carrée $p \times p$ de terme général $\langle x_i, x_j \rangle$.
3.b) Si $p = n$ et $(x_1, \dots, x_n)$ base de E, quel est le nom habituel de $G(x_1, \dots, x_n)$?
3.c) Rappeler pourquoi il existe une base $\mathcal{B}$ de E orthonormale pour le produit scalaire $\langle \cdot, \cdot \rangle$. Cette base est-elle unique ?.

Dorénavant $\mathcal{B}$ désigne une telle base. On note X la matrice $n \times p$ dont les colonnes $X_1, \dots, X_p$ donnent les composantes respectives de $x_1, \dots, x_p$ sur la base $\mathcal{B}$.
3.d) Exprimer $G(x_1, \dots, x_p)$ en fonction de X et ${}^tX$.
4) On suppose dorénavant que $(x_1, \dots, x_p)$ est libre.
4.a) Justifier que $\det G(x_1, \dots, x_p) \neq 0$.
4.b) Soit $F = \mathrm{Vect}(x_1, \dots, x_p)$. Rappeler la définition de $F^\perp$, orthogonal de F et donner sa dimension.
4.c) Soit $(x_{p+1}, \dots, x_n)$ une base orthonormale de $F^\perp$. Que dire de $(x_1, \dots, x_n)$?
4.d) Soit $\tilde{X}$ la matrice carrée $n \times n$ dont les colonnes donnent les composantes de $x_1, \dots, x_p, x_{p+1}, \dots, x_n$ dans la base $\mathcal{B}$. Comment appelle-t-on habituellement cette matrice ?
4.e) Écrire $G(x_1, \dots, x_n)$ puis exprimer $\det G(x_1, \dots, x_p)$ en fonction de $\det \tilde{X}$.
4.f) Montrer que, pour tous $\lambda_1, \dots, \lambda_{p-1}$ réels,

$$\det G(x_1, \dots, x_p + \sum_{j<p} \lambda_j\ x_j) = \det G(x_1, \dots, x_p)$$

5) Pour $y \in E \setminus F$, on note $p_F(y)$ la projection orthogonale de y sur F.
5.a) Montrer que

$$||y - p_F(y)||^2 = \min_{z \in F} ||y - z||^2$$

Cette quantité sera désignée dans la suite par $d(y, F)$.
5.b) Justifier que

$$\det G(x_1, \dots, x_p, y) = \det G(x_1, \dots, x_p, y - p_F(y))$$

5.c) Donner une expression par blocs de $G((x_1, \dots, x_p, y - p_F(y))$.
5.d) En déduire que

$$d(y, F)^2 = \frac{\det G(x_1, \dots, x_p, y)}{\det G(x_1, \dots, x_p)}$$

6.a) Montrer que la matrice de p_F dans la base $\mathcal{B}$ est

$$X \, \left({}^t X \, X\right)^{-1} \, {}^t X$$

6.b) En utilisant la notation Y pour la matrice des composantes de y dans la base $\mathcal{B}$, écrire $G(x_1, \dots, x_p, y)$ par blocs en fonction de X et Y (et leurs transposées).
6.c) En utilisant 1.b), retrouver la formule obtenue en 5.d).

Chapitre 8

Structure hermitienne

La théorie que nous avons développée jusqu'à présent pour les espaces vectoriels réels s'adapte au cas d'un espace vectoriel dont le corps des scalaires est $\mathbb{C}$. Grosso modo il suffit d'ajouter des barres de conjugaison aux bons endroits.

8.1 Forme quadratique hermitienne

Nous travaillons donc maintenant sur un $\mathbb{C}$-espace vectoriel E. Au lieu de nous interesser comme avant aux formes bilinéaires, nous allons priviligier la notion de forme sesquilinéaire. Le préfixe "sesqui" signifie 3/2. Une forme sesquilinéaire ϕ est une application de $E \times E$ dans $\mathbb{C}$ qui est linéaire par rapport à la seconde variable

$$\begin{aligned} \forall x, y, y' \in E, \quad \phi(x, y + y') &= \phi(x, y) + \phi(x, y'), \\ \forall x, y \in E,\ \forall \lambda \in \mathbb{C}, \quad \phi(x, \lambda\, y) &= \lambda\, \phi(x, y) \end{aligned}$$

et semi-linéaire par rapport à la première variable

$$\begin{aligned} \forall x, x', y \in E, \quad \phi(x + x', y) &= \phi(x, y) + \phi(x', y), \\ \forall x, y \in E,\ \forall \lambda \in \mathbb{C}, \quad \phi(\lambda\, x, y) &= \overline{\lambda}\, \phi(x, y) \end{aligned}$$

Ainsi pour sortir une constante complexe en facteur de la première variable, il faut la conjuguer. C'est la seule différence avec le cas réel. Au lieu de favoriser comme avant les formes symétriques, nous regardons maintenant les formes à *symétrie hermitienne*. Cela signifie que

$$\forall x, y \in E, \quad \phi(x, y) = \overline{\phi(y, x)}$$

On notera en particulier que $\phi(x,x)$ est réel. En fait étant donné la symétrie hermitienne on peut se contenter d'énoncer la linéarité par rapport à la seconde variable. L'application q de E dans $\mathbb{R}$ définie par $q(x) = \phi(x,x)$ est appelée forme quadratique associée à ϕ. On notera que pour tout $\lambda \in \mathbb{C}$, $q(\lambda\, x) = |\lambda|^2\ q(x)$. Les identités de polarisation s'écrivent de façon légérement plus compliquée :

$$\begin{aligned} q(x+y) &= q(x)+q(y)+2\,\mathrm{Re}[\phi(x,y)] \\ q(x+iy) &= q(x)+q(y)-2\,\mathrm{Im}[\phi(x,y)] \end{aligned}$$

Si le $\mathbb{C}$-espace vectoriel E est de dimension finie n et admet pour base $(b_1,\ldots,b_n)$ alors en fonction des décompositions $x = \sum_{i=1}^n x_i\, b_i$ et $y = \sum_{i=1}^n y_i\, b_i$ on a

$$\phi(x,y) = \sum_{1\leq i,j\leq n} \overline{x_i}\, y_j\ \phi(b_i,b_j)$$

Sous forme matricielle $\phi(x,y) = {}^t\overline{X}\, Q\, Y$ où X est la matrice colonne des coordonnées de x, Y celle des coordonnées de y ; la matrice Q est la matrice de terme général $\phi(b_i,b_j)$ appelée matrice de ϕ (ou de q) dans la base $(b_1,\ldots,b_n)$ et notée $\mathrm{mat}_{(b_i)}\phi$ ou $\mathrm{mat}_{(b_i)}q$. On note usuellement $X^* = {}^t\overline{X}$. Bien sur on a alors $q(x) = X^*\, Q\, X$. Comme $\phi(b_i,b_j) = \overline{\phi(b_j,b_i)}$, la matrice Q est *hermitienne* c'est à dire que $Q^* = Q$.

Soit une nouvelle base $(f_1,\ldots,f_n)$ et P la matrice de passage de la base $(f_1,\ldots,f_n)$ à la base $b_1,\ldots,b_n)$. Notons $\tilde{X}$ et $\tilde{Y}$ les matrices colonnes des composantes de x et y dans la base $(f_1,\ldots,f_n)$. On sait que l'on a $\tilde{X} = PX$, $\tilde{Y} = PY$ et $\phi(x,y) = \tilde{X}^*\tilde{Q}\tilde{Y} = X^*P^*\tilde{Q}\,P\,Y = X^*Q\,Y$. Par identification, on a ainsi $\tilde{Q} = P^*\,Q\,P$.

8.2 Réduction de Gauss

Comme dans le cas réel, il est intéressant de trouver une base où la matrice de ϕ est diagonale. Nous verrons au prochain chapitre que cela est possible, avec en plus une matrice de passage P d'un type particulier que nous appelerons unitaire ce qui signifie $P^* = P^{-1}$. Il s'agit alors de trouver P telle que $P^*QP = P^{-1}QP$ soit diagonale. Le problème est donc ramené à la diagonalisation de la matrice hermitienne Q. Nous reparlerons de ce résultat.

Notons toutefois que, comme dans le cas réel, on peut utiliser l'algorithme de Gauss. Le théorème de Sylvester s'écrit dans le cas complexe de la manière suivante.

Théorème 23 (Sylvester complexe) *Soit q forme quadratique sur le $\mathbb{C}$-espace vectoriel E de dimension n. Alors il existe des formes linéaires indépendantes $v_1, \dots, v_n$ sur E telles que*

$$q(x) = \sum_{i=1}^{n} \alpha_i \ |v_i(x)|^2 \ \text{ avec } \alpha_i \in \{1, -1, 0\}$$

Le nombre de 1 et de -1 dépend de q et pas du choix des formes $v_1, \dots, v_n$. Ces nombres de 1 et de -1 sont appelés signature de q.

8.3 Espace hermitien

Nous allons nous concentrer maintenant sur le cas d'une forme quadratique q définie positive c'est à dire $q(x) > 0$ pour tout $x \in E \setminus \{0\}$. Dans le cas de la dimension finie cela veut dire que sa matrice A vérifie $X^* AX \geq 0$ pour toute matrice colonne X non nulle autrement dit $\sum_{i,j} a_{ij} \ \overline{x}_i x_j > 0$ pour tous complexes $x_1, \dots, x_n$ non tous nuls.

Quand un $\mathbb{C}$-espace vectoriel E est muni d'une forme sesquilinéaire à symétrie hermitienne définie positive, on appelle E espace *hermitien* et la forme sesquilinéaire est appelée produit scalaire. On notera comme avant $\langle \cdot, \cdot \rangle$ ce produit scalaire et on pose pour tout $x \in E$, $||x|| = \sqrt{\langle x, x \rangle}$. Notons que pour tous $x, y \in E$, on a $||x+y||^2 = ||x||^2 + ||y||^2 + 2\,\mathrm{Re}[\langle x, y \rangle]$. L'inégalité de *Cauchy-Schwarz* s'écrit comme dans le cas réel. Pour la prouver on note θ un argument du complexe $\langle x, y \rangle$ et on étudie $||x + t\, e^{-i\,\theta}\, y||^2 = ||y||^2\, t^2 + 2t\, |\langle x, y \rangle| + ||x||^2$. Comme dans le cas réel on exprime que le discriminant de ce trinome en t est négatif ou nul.

Un exemple d'espace hermitien de dimension finie est la *structure hermitienne canonique* sur $\mathbb{C}^d$ dont le produit scalaire est donné par

$$\forall x = (x_1, \dots, x_d) \in \mathbb{C}^d, \ \forall y = (y_1, \dots, y_d) \in \mathbb{C}^d, \ \ \langle x, y \rangle = \sum_{i=1}^{d} \overline{x}_i \, y_i$$

Dans un espace hermitien, la notion d'orthogonalité pour deux vecteurs ou deux sous-espaces se définit comme dans le cas réel. La proposition 15 reste vraie. De même on définit les notions de bases orthogonales et orthonormales. Le procédé de Gram-Schmidt reste valide avec ses corollaires.

8.4 Endomorphismes remarquables d'un espace hermitien

Dans un espace hermitien les endomorphismes qui préservent la norme sont dits *unitaires*. Ce sont ceux dont la matrice U dans une base orthonormale est unitaire ce qui signifie $U^* U = I$. Un endomorphisme est unitaire si et seulement si il transforme une base orthonormale en une base orthonormale. On effectuera la réduction des matrices unitaires dans le prochain chapitre.

Chapitre 9

Endomorphismes normaux

Dans ce chapitre nous allons traiter simultanément des espaces euclidiens et hermitiens. Nous allons nous interesser aux endomorphismes sur ces espaces que nous appelons endomorphismes normaux. Ce type est une classe assez large qui contient les types particuliers d'endomorphismes utiles dans la pratique. Nous ne différentierons le cas complexe du cas réel qu'en cas de nécessité.

On suppose donc que E est un $\mathbb{K}$-espace vectoriel *de dimension finie* qui est *euclidien* dans le cas $\mathbb{K} = \mathbb{R}$ et *hermitien* dans le cas $\mathbb{K} = \mathbb{C}$. Cela signifie que E est muni d'un produit scalaire que l'on notera $\langle \cdot, \cdot \rangle$.

Dans le cas réel $\mathbb{K} = \mathbb{R}$, le produit scalaire est une forme bilinéaire symétrique qui est définie positive ce qui veut dire que $\langle x, x \rangle > 0$ pour $x \neq 0$.

Dans le cas complexe $\mathbb{K} = \mathbb{C}$, le produit scalaire est une forme sesquilinéaire à symetrie hermitienne définie positive. Cela signifie que la forme est linéaire par rapport à la seconde variable et semi-linéaire par rapport à la première

$$\forall x, x', y \in E,\ \forall \lambda \in \mathbb{C},\ \langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle,\ \ \langle \lambda\, x, y \rangle = \overline{\lambda}\, \langle x, y \rangle$$

La symétrie hermitienne s'écrit $\langle x, y \rangle = \overline{\langle y, x \rangle}$. On notera en particulier que $\langle x, x \rangle$ est réel. Le caractère défini positif s'exprime comme précédemment : $\langle x, x \rangle > 0$ pour $x \neq 0$.

Les notions d'orthogonalité, de norme, etc ... et les résultats élémentaires ont été abordés dans le précédent chapitre.

9.1 Adjoint d'un endomorphisme

Si A est une matrice on pose $A^* = {}^t\overline{A}$. Dans le cas réel c'est donc la transposée et dans le cas complexe la trans-conjuguée.

Proposition 24 *Pour tout endomorphisme u de E, il existe un unique endomorphisme u^* de E appelé* adjoint *de u qui vérifie, pour tout x, y de E,*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle$$

Si $\mathcal{B}$ est une base orthonormale *de E, on a*

$$mat_{\mathcal{B}}(u^*) = (mat_{\mathcal{B}}(u))^*$$

Preuve. Voyons déja l'unicité en supposant que u^* existe. Le coefficient sur la i-ième ligne et la j-ième colonne de la matrice de u^* dans la base orthonormée $\mathcal{B} = (b_1, \ldots, b_n)$ est la composante suivant b_i de $u^*(b_j)$. Mais comme la base $\mathcal{B}$ est orthonormale cette composante est $\langle b_i, u^*(b_j)\rangle = \overline{\langle u(b_i), b_j\rangle}$. Or $\langle u(b_i), b_j\rangle$ est le coefficient d'indices j, i de A. La matrice de u^* dans $\mathcal{B}$ ne peut donc être que la trans-conjuguée de la matrice de u dans $\mathcal{B}$, prouvant ainsi l'unicité. Inversement si on définit de cette manière la matrice de u^*, on a pour tous i, j, $\langle u(b_i), b_j\rangle = \langle b_i, u^*(b_j)\rangle$. En sommant sur i, j avec des coefficients, on passe facilement à $\langle u(x), y\rangle = \langle x, u^*(y)\rangle$ pour tous x, y.

9.2 Endomorphismes normaux

Nous allons étudier une classe d'endomorphismes qui regroupe plusieurs cas intéressants. On dit qu'un endomorphisme u est <u>normal</u> si $u \circ u^* = u^* \circ u$ c'est à dire s'il commute avec son adjoint. Citons par exemple dans le cas réel

- les endomorphismes symétriques (ou auto-adjoint) : $u^* = u$
- les endomorphismes antisymétriques : $u^* = -u$
- les endomorphismes orthogonaux : $u^* = u^{-1}$

et dans le cas complexe

- les endomorphismes hermitiens (ou auto-adjoint) : $u^* = u$
- les endomorphismes antihermitiens : $u^* = -u$
- les endomorphismes unitaires : $u^* = u^{-1}$

On dit qu'une matrice A est *normale* si $A\,A^* = A^*\,A$. Ainsi un endomorphisme est normal si et seulement si sa matrice sur une base orthonormale est normale.

A titre d'exemple déterminons les endomorphismes normaux de $\mathbb{R}^2$ muni de sa structure euclidienne canonique. Une matrice réelle 2×2 est normale si et seulement si elle vérifie :

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

soit

$$\begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$$

La première possibilité est $b = c$ ce qui fournit les matrices symétriques. La seconde possibilité est $c = -b$ ce qui conduit ensuite à $b(a-d) = 0$ d'où $b = 0$ ou $a = d$. Au total les matrices réelles 2×2 normales sont celles de la forme

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} \text{ ou } \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Le second type est appelé matrice de similitude. Pour $a = 0$ on retrouve les matrices anti-symétriques. Pour $a^2 + b^2 = 1$ on retrouve les matrices orthogonales de déterminant 1 c'est à dire les matrices de rotation.

Nous allons procéder dans le prochain paragraphe à la réduction des endomorphismes normaux. Le point clé est le résultat suivant.

Proposition 25 *Soit u un endomorphisme normal de E espace euclidien ou hermitien. Si F est un sous-espace vectoriel u–stable, alors $F^\perp$ est u–stable et les restrictions de u à F et à $F^\perp$ sont normales.*

Choisissons des bases orthonormales de F et de $F^\perp$. La réunion de ces deux bases donne une base orthonormale $\mathcal{B}$ de E. La matrice de u sur $\mathcal{B}$ s'écrit par blocs sous la forme

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

Puisque $\mathcal{B}$ est orthonormée, la matrice de u^* sur $\mathcal{B}$ est

$$\begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}$$

Alors la relation $u \circ u^* = u^* \circ u$ s'écrit matriciellement

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} = \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

soit

$$\begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix} = \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix}$$

Cela entraine en particulier $BB^* = 0$. Or la trace de BB^* est $\sum_i \sum_k b_{ik}\overline{b_{ik}}$. La nullité de cette quantité prouve que la matrice $B = (b_{ij})$ est nulle donc $F^\perp$ est stable par u. L'égalité matricielle donne de plus $CC^* = C^*C$ ce qui montre que la restriction de u à $F^\perp$ est normale.

9.3 Réduction dans le cas complexe

Dans le cas **hermitien** tous les endomorphismes normaux sont diagonalisables.

Théorème 26 *Un endomorphisme d'un espace hermitien est normal si et seulement si il est diagonalisable dans une base orthonormale. Les valeurs propres sont réelles dans le cas auto-adjoint (hermitien), imaginaires pures dans le cas anti-hermitien et de module 1 dans le cas unitaire.*

Preuve. Nous procédons par récurrence sur la dimension de l'espace hermitien considéré. Le cas de la dimension 1 est trivial. Admettons le résultat dans tout espace de dimension $n-1$ avec $n \geq 2$ et considérons un endomorphisme normal u de l'espace hermitien E de dimension n. Par la proposition 6 on sait que u admet une droite stable F. Alors la proposition 25 prouvée juste avant nous dit que $F^\perp$ est stable par u et que la restriction de u à $F^\perp$ est normale. Nous pouvons appliquer l'hypothèse de récurrence à cette restriction car $F^\perp$ est de dimension $n-1$. On conclut alors en prenant comme base de diagonalisation la réunion d'un vecteur directeur de F et la base de diagonalisation de la restriction de u à $F^\perp$.

On remarquera qu'une conséquence évidente du théorème est que deux sous-espaces propres associés à des valeurs propres distinctes sont orthogonaux.

9.4 Réduction dans le cas réel

Dans le cas réel la réduction est différente.

Théorème 27 *Un endomorphisme d'un espace euclidien est normal si et seulement si il existe une base orthonormale $\mathcal{B}$ sur laquelle la matrice de u s'écrit sous la forme diagonale par blocs*

$$mat_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_p & & & \\ & & & S_1 & & \\ & 0 & & & \ddots & \\ & & & & & S_q \end{pmatrix}$$

où les $\lambda_1, \ldots, \lambda_p$ sont des réels et $S_1, \ldots, S_q$ sont des blocs 2×2 à coefficients réels qui sont des matrices de similitudes de la forme

$$S_l = \begin{pmatrix} a_l & -b_l \\ b_l & a_l \end{pmatrix}$$

En particulier, si u est symétrique, u est diagonalisable. Si u est anti-symétrique les coefficients λ_l et a_l sont nuls. Si u est orthogonal, les λ_l valent 1 ou -1 et les blocs 2×2 sont des matrices de rotation de la forme

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Preuve. Comme dans le cas complexe, nous procédons par récurrence sur la dimension de l'espace euclidien considéré. Le cas de la dimension 1 est encore trivial. Admettons le résultat dans tout espace de dimension inférieure ou égale à $n-1$ avec $n \geq 2$ et considérons un endomorphisme normal u de l'espace euclidien E de dimension n. Par la proposition 6 on sait que u admet une sous-espace stable F de dimension 1 ou 2. Nous savons que dans une base orthonormale de F quelconque la matrice de la restriction de u à F a la forme voulue car nous avons étudié les endomorphismes normaux réels en dimension 2 dans le paragraphe précédent. Par ailleurs la proposition 25 affirme que $F^\perp$ est stable par u et que la restriction de u à $F^\perp$ est normale. Nous pouvons appliquer l'hypothèse de récurrence à cette restriction car $F^\perp$ est de dimension $n-1$ ou $n-2$. On conclut en prenant comme base la réunion d'une base orthonormale de F et la base obtenue en appliquant l'hypothèse de récurrence à la restriction de u à $F^\perp$.

9.5 Formes quadratiques et endomorphismes

Les formes quadratiques et endomorphismes auto-adjoints sont intimement liés. En effet munissons un espace vectoriel de dimension finie E réel ou complexe d'une structure d'espace euclidien dans le cas réel et hermitien dans le cas complexe par le biais d'une forme de référence ϕ. Alors pour tout forme bilinéaire symétrique dans le cas réel ou toute forme sesqui-linéaire à symétrie hermitienne dans le cas complexe, que nous noterons $\tilde{\phi}$, il existe un unique endomorphisme u de E tel que

$$\forall x, y \in E, \ \ \tilde{\phi}(x, y) = \phi(x, u(y))$$

En effet, si on se donne une base $\mathcal{B}$ sur E, cette condition est équivalente à la forme matricielle

$$\mathrm{mat}_{\mathcal{B}}\tilde{\phi} = (\mathrm{mat}_{\mathcal{B}}\phi) \ \ (\mathrm{mat}_{\mathcal{B}}u)$$

Dans cette écriture figurent la notion de matrice d'une forme bilinéaire et celle de matrice d'un endomorphisme. En particulier si la base $\mathcal{B}$ est orthonormée pour ϕ, la matrice de $\tilde{\phi}$ et celle de u coincident. Il faut noter que l'endomorphisme u est auto-adjoint relativement au produit scalaire ϕ car sa matrice dans une base orthonormée a la symétrie voulue. Or il résulte des deux paragraphes précédents que les endomorphismes auto-adjoints sont diagonalisables dans une base orthonormale. Parce que ce résultat est important, réecrivons le sous forme matricielle, dans le cas réel tout d'abord.

Proposition 28 *Soit S une matrice symétrique réelle alors il existe une matice orthogonale P et une matrice diagonale réelle D telles que $S = P^{-1}\, D\, P = {}^tP\, D\, P$*

On voit là un moyen facile de déterminer la signature (r_+, r_-) de la forme bilinéaire symétrique associée à S, il suffit de compter sur D le nombre r_+ de termes diagonaux strictements positifs et le nombre r_- de termes diagonaux strictement négatifs.

Dans le cas complexe la formulation est similaire.

Proposition 29 *Soit H une matrice complexe hermitienne (${}^t\overline{H} = H$) alors il existe une matice unitaire U et une matrice diagonale réelle D telles que $H = U^{-1}\, D\, U = {}^t\overline{U}\, D\, U$*

Revenons aux formes bilinéaires pour obtenir un nouveau point de vue.

Proposition 30 (Théorème de réduction simultanée) *Soit E un espace vectoriel réel [resp. complexe] et ϕ une forme bilinéaire symétrique définie positive sur E [resp. une forme sesquilinéaire à symétrie hermitienne* *définie positive sur E]. Soit $\tilde{\phi}$ une autre forme bilinéaire symétrique sur E [resp. une autre forme sesquilinéaire à symétrie hermitienne sur E], il existe une base $(f_1, \ldots, f_n)$ de E telle que $\phi(f_i, f_j) = \delta_i^j$ et $\tilde{\phi}(f_i, f_j) = 0$ si $i \neq j$ autrement dit cette base est orthonormale pour ϕ et "orthogonale pour $\tilde{\phi}$".*

Voyons comment ce résultat se déduit de ce qui précède. Considérons une base $\mathcal{B}$ orthonormale pour ϕ. Nous avons vu qu'il existe un unique endomorphisme u de E tel que $\tilde{\phi}(x, y) = \phi(x, u(y))$ pour tous x, y et qu'il suffit de choisir pour matrice de u sur la base $\mathcal{B}$ la matrice de $\tilde{\phi}$ dans cette même base. Cet endomorphisme u est auto-adjoint dans (E, ϕ) car sa matrice dans une base ϕ-orthonormale est symétrique. On sait alors qu'il existe une base (f_i) qui est ϕ-orthonormale et qui est une base de vecteurs propres pour u disons $u(f_i) = \lambda_i\, f_i$. Alors $\tilde{\phi}(f_i, f_j) = \phi(f_i, u(f_j)) = \phi(f_i, \lambda_j\, f_j) = \lambda_j\, \delta_i^j$. On a bien trouvé en (f_i) une base convenable. Ecrivons maintenant une traduction matricielle.

Proposition 31 *Soit A une matrice symétrique réelle définie positive et B une matrice symétrique réelle. Alors il existe une matrice réelle inversible P telle que ${}^tP\,A\,P = I$ et ${}^tP\,B\,P$ diagonale.*

Soit H une matrice complexe hermitienne définie positive et K une matrice complexe hermitienne. Alors il existe une matrice complexe inversible P telle que ${}^t\overline{P}\,H\,P = I$ et ${}^t\overline{P}\,K\,P$ diagonale réelle.

Signalons un critère classique de stricte positivité d'une matrice symétrique réelle.

Proposition 32 *Une matrice symétrique réelle $A = (a_{ij})_{1 \leq i,j \leq n}$ est définie positive si et seulement si*

$$\forall r \in \{1, \ldots, n\},\ \det[(a_{ij})_{1 \leq i,j \leq r}] > 0$$

ces déterminants étant appelés mineurs principaux.

Preuve. Nous procédons par récurrence sur n. Le cas $n = 1$ est trivial. Supposons, pour $n \geq 2$, le résultat acquis pour toutes les dimensions strictement inférieures à n et considérons une matrice A symétrique réelle $n \times n$ dont tous les mineurs principaux sont strictement positifs. Notons q la forme quadratique sur $\mathbb{R}^n$ dont A est la matrice sur la base canonique $(c_1, \ldots, c_n)$ de $\mathbb{R}^n$. En appliquant l'hypothèse de récurrence à la restriction de q à $\mathrm{Vect}(c_1, \ldots, c_{n-1})$ on voit que cette restriction est définie positive. La signature de q est donc $(n-1, 1)$, $(n-1, 0)$ ou $(n, 0)$. Seul ce dernier cas est possible car le déterminant est strictement positif ce qui conclut la preuve.

9.6 Exercices

Exercice 44 Montrer que la matrice

$$\begin{pmatrix} 2 & i \\ i & 0 \end{pmatrix}$$

n'est pas diagonalisable. Cela est il en contradiction avec la théorie ?

Exercice 45 Soit $M = (m_{ij})_{1 \leq i,j \leq n}$ la matrice définie par $m_{ii} = a$, $m_{ij} = b$ si $|i-j| = 1$ et 0 pour les autres coefficients. Les constantes a et b sont réelles. Justifier que la matrice M est diagonalisable dans une base orthonormale. Procéder au calcul effectif. Vérifier l'orthogonalité des sous-espaces propres. Quelle est la signature de la forme quadratique associée à M ?

Exercice 46 (Décomposition polaire) 1) Montrer que si U est une matrice symétrique définie positive alors il existe une matrice symétrique définie positive S telle que $S^2 = U$
2) Montrer que si S est une matrice diagonalisable alors il existe un polynome H tel que $S = H(S^2)$.
3) Montrer que si deux matrices diagonalisables S et S' à valeurs propres positives vérifient $S^2 = S'^2$ alors elles sont égales.
4) En déduire que la matrice S obtenue en fonction de U à la question 1 est unique.
5) Soit M une matrice carrée inversible. On suppose que $M = O\,S$ où S est symétrique définie positive et O est orthogonale. Montrer alors que S est uniquement déterminée à partir de ${}^tM\,M$ par le procédé de la question 1.
6) Prouver que pour toute matrice inversible M, il existe effectivement un unique couple (S, O) où S est symétrique définie positive et O orthogonale tel que $M = O\,S$.

Exercice 47 Montrer qu'un endomorphisme u d'un espace hermitien (de dimension finie) est normal si et seulement si son adjoint u^* est un polynome en u.

Exercice 48 1) Trouver un endomorphisme f non nul d'un espace E euclidien qui vérifie $\langle f(x), x \rangle = 0$ pour tout $x \in E$.
2) Montrer qu'un endomorphisme f d'un espace E hermitien qui vérifie $\langle f(x), x \rangle = 0$ pour tout $x \in E$ est nul.
3) Montrer qu'un endomorphisme u d'un espace E hermitien est normal si et seulement si $||u(x)|| = ||u^*(x)||$ pour tout $x \in E$.

Exercice 49 On considère deux matrices carrées complexes H et K qui sont hermitiennes avec en plus H positive. Montrer que

$$|\det(H+iK)| \geq \det H$$

Exercice 50 Soit f un endomorphisme d'un espace hermitien (de dimension finie).
1) Montrer qu'il existe un unique couple (u,v) d'endomorphismes auto-adjoints tels que $f = u + i\,v$.
2) Avec les notations de la question précédente montrer que f est normal si et seulement si u et v commutent.

Exercice 51 Soit $A \in \mathcal{GL}_n(\mathbb{C})$. Montrer qu'il existe deux matrices unitaires U et V et une matrice diagonale D à coefficients réels strictement positifs telles que $A = U\,D\,V^*$.

Enoncer un résultat similaire dans le cas réel.

Chapitre 10

Topologie des espaces de matrices

Dans ce chapitre nous allons munir $\mathcal{M}_d(\mathbb{R})$ ensemble des matrices carrées $d \times d$ à coefficients réels ou $\mathcal{M}_d(\mathbb{C})$ ensemble des matrices carrées $d \times d$ à coefficients complexes d'une topologie. D'une part, nous souhaitons préciser les propriétés topologiques des ensembles remarquables de matrices étudiés précédemment. D'autre part, nous donnerons des exemples en cours ou en exercice de méthodes analytiques pour l'obtention de résultats algébriques.

Nous privilégions l'aspect matriciel pour sa commodité. Il va de soi que nous pouvons faire les traductions en termes d'ensembles d'endomorphismes sur des espaces vectoriels de dimension finie. Quand cela sera nécessaire nous effectuerons d'ailleurs les identifications habituelles entre l'aspect matriciel et l'aspect vectoriel.

La structure topologique dont nous munissons les espaces de matrices est la structure d'espace normé. Il s'agit donc d'un cas particulier de la théorie des espaces vectoriels normés et en plus, il s'agit ici d'espaces de dimension finie ce qui–on le sait–apporte une grande simplification. Pour l'autonomie de ce cours nous rappelons d'abord un certain nombre de définitions et résultats de cette théorie, bien sûr sans démonstration.

10.1 Topologie issue d'une norme

Sur $\mathcal{M}_d(\mathbb{R})$ ou $\mathcal{M}_d(\mathbb{C})$, comme sur des espaces vectoriels généraux, on appelle *norme* une application à valeurs réelles : $M \mapsto ||M||$ qui a les propriétés suivantes :

1. $||M|| = 0 \Leftrightarrow M = 0$

2. $||\lambda M|| = |\lambda| \, ||M||$
3. $||A + B|| \leq ||A|| + ||B||$

pour toutes matrices A, B, M et tout scalaire λ (c'est à dire un réel ou un complexe suivant le cas). Les espaces normés sont des cas particuliers d'espace métrique. On peut en effet définir une distance par la formule $\text{dist}(A, B) = ||B - A||$. Cette distance permet de définir les notions de boule ouverte $B(A, r) = \{M;\ ||M - a|| < r\}$ ($r \in \mathbb{R}_+^*$) ou fermée $\overline{B}(A, r) = \{M;\ ||M - A|| \leq r\}$. On appelle ouvert une partie qui, pour chacun de ses points, contient une boule ouverte centrée en ce point. L'ensemble des ouverts est appelé topologie. Il y a propriété de stabilité par réunion (quelconque) ou intersection finie. On dit qu'une partie est fermée si son complémentaire est ouvert.

Une notion fondamentale est la notion de limite. Rappelons en particulier qu'une suite (A_n) de matrices converge vers une matrice L pour la norme $|| \cdot ||$ si

$$\lim_{n \to +\infty} ||L - A_n|| = 0$$

Une partie est fermée si elle contient toute limite d'une suite d'éléments de cette partie. Une partie est dite compacte si de toute suite d'éléments de cette partie, on peut extraire une sous-suite convergente. Dans tout espace vectoriel normé de dimension finie, les compacts sont simplement les parties fermées et bornées.

Le détail de ces notions figure dans tout cours de topologie sur les espaces vectoriels normés.

Il y a bien sûr de nombreuses normes qui peuvent être définies sur l'espace des matrices. Nous verrons différents exemples plus loin mais il faut signaler que toutes conduisent à la même topologie. Autrement dit, les normes ne sont pas toutes égales mais elles donnent les mêmes notions d'ouvert, de fermé, de compact, de limite. En effet l'espace des matrices est de dimension finie. On sait alors que si $|| \cdot ||_a$ et $|| \cdot ||_b$ sont deux normes sur cet espace, il existe deux constantes α, $\beta > 0$ telles que, pour toute matrice M,

$$\alpha \, ||M||_a \leq ||M||_b \leq \beta \, ||M||_a$$

On dit que les deux normes $|| \cdot ||_a$ et $|| \cdot ||_b$ sont équivalentes.

Une autre conséquence de la finitude de la dimension pour l'espace des matrices est que celui-ci est complet. Cela signifie que toute suite de Cauchy est convergente. Une suite (A_n) de matrices est dite de Cauchy si, pour tout $\varepsilon > 0$ il existe un entier n_0 tel que, pour tous $q \geq p \geq n_0$, $||A_q - A_p|| \leq \varepsilon$. Comme la notion de limite, la notion de suite de Cauchy ne dépend pas, dans le cas de la dimension finie, de la norme choisie. Cela se vérifie facilement à partir de la relation d'équivalence des normes.

10.2 Exemples de normes de matrices

Notons le corps des scalaires $\mathbb{R}$ ou $\mathbb{C}$ par $\mathbb{K}$. Un bon moyen de construire une norme sur $\mathcal{M}_d(\mathbb{K})$ est de partir d'une norme $|\cdot|$ sur $\mathbb{K}^d$ et de poser

$$||M|| = \sup_{|X|\leq 1} |M\,X|$$

On définit ainsi la norme de M comme la "norme" (au sens évoqué en topologie) de l'endomorphisme de $\mathbb{K}^d$ –automatiquement continu– qui a pour matrice M dans la base canonique de $\mathbb{K}^d$. On dit alors que la norme $||\cdot||$ est <u>subordonnée</u> à la norme $|\cdot|$ sur $\mathbb{K}^d$. Il y a trois exemples de normes classiques sur $\mathbb{K}^d$: pour $x = (x_1, \dots, x_d) \in \mathbb{K}^d$,

$$|x|_1 = \sum_{i=1}^{d} |x_i|,\ |x|_2 = \left(\sum_{i=1}^{d} |x_i|^2\right)^{1/2} = (x^*\,x)^{1/2}\,,\ |x|_\infty = \sup_{1\leq i\leq d} |x_i|$$

On a noté x^* pour la matrice ligne qui est la transconjuguée de la matrice colonne x. C'est simplement la transposée dans le cas réel.

Les normes subordonnées correspondant aux normes ci-dessus peuvent pour deux d'entre elles s'exprimer à l'aide des coefficients de la matrice.

Proposition 33 *Pour toute matrice carrée* $M = (m_{ij})_{1\leq i,j\leq d}$*, on a*

$$(i)\ \ ||M||_1 = \sup_{|X|_1\leq 1} |M\,X|_1 = \sup_{j=1}^{d} \sum_{i=1}^{d} |m_{ij}|$$

$$(ii)\ \ ||M||_\infty = \sup_{|X|_\infty\leq 1} |M\,X|_\infty = \sup_{i=1}^{d} \sum_{j=1}^{d} |m_{ij}|$$

$$(iii)\ \ ||M||_2 = \sup_{|X|_2\leq 1} |M\,X|_2 = (\sup\{|\lambda|;\ \lambda \in \mathrm{Sp}(M^*M)\})^{1/2}$$

Preuve. (i) Pour $X = (x_1 \dots x_d)^T$ on écrit d'abord

$$\begin{aligned} |M\,X|_1 &= \sum_i \left|\sum_j m_{ij}\,x_j\right| \\ &\leq \sum_j |x_j| \sum_i |m_{ij}| \\ &\leq \left(\sup_j \sum_i |m_{ij}|\right) \sum_j |x_j| \end{aligned}$$

cq qui prouve déjà que $||M||_1 \leq \sup_{j=1}^d \sum_{i=1}^d |m_{ij}|$. Il reste à trouver X de norme égale à 1 tel que $|M\,X|_1$ soit égal au membre de droite ci dessus. Si on note j_0 qui réalise le sup dans le membre de droite, il suffit de prendre $x_j = \mathbf{1}(\{j = j_0\})$.

Pour **(ii)** la méthode est similaire. Un vecteur qui réalise la norme est cette fois donné par $x_j = \overline{m_{i_0j}}/|m_{i_0j}|$ où i_0 est l'indice qui réalise le sup dans l'expression $\sup_i \sum_j |m_{ij}|$.

(iii) Pour $||M||_2$, quelques remarques s'imposent. On appelle rayon spectral d'une matrice le plus grand module des valeurs propres de cette matrice considérée comme matrice à coefficients complexes. Nous reviendrons sur l'étude du rayon spectral dans un prochain paragraphe. Notons que la matrice M^*M est autoadjointe c'est à dire hermitienne dans le cas complexe et symétrique dans le cas réel. Elle est donc diagonalisable à valeurs propres réelles. Ces valeurs propres sont positives. En effet si x est vecteur propre associé à la valeur propre λ on a $M^*Mx = \lambda x$ et, en multipliant à gauche par x^*, on trouve $|Mx|_2^2 = \lambda\,|x|_2^2$. On sait même que la matrice auto-adjointe M^*M est diagonalisable dans une base orthonormale c'est à dire qu'il existe une matrice orthogonale dans le cas réel ou unitaire dans le cas complexe P telle que $P^*M^*MP = D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_d)$. Alors

$$\begin{aligned}
||M||_2^2 &= \sup_{X\neq 0} \frac{|M\,X|_2^2}{|X|_2^2} = \sup_{Y\neq 0} \frac{|M\,PY|_2^2}{|PY|_2^2} = \sup_{Y\neq 0} \frac{Y^*P^*M^*MPY}{|Y|_2^2} \\
&= \sup_{Y\neq 0} \frac{Y^*DY}{|Y|_2^2} = \sup_{Y\neq 0} \frac{\sum_{i=1}^d \lambda_i\, y_i^2}{\sum_{i=1}^d y_i^2} = \sup_{i=1}^d \lambda_i = \rho(M^*M)
\end{aligned}$$

On a effectué le changement $X = PY$ puis remarqué que $|PY|_2 = |Y|_2$. Cela achève la preuve de la proposition.

Les normes sur l'espace des matrices ne sont pas toutes subordonnées. Par exemple la formule

$$||M||_e = \left(\sum_{i,j} |m_{ij}|^2\right)^{1/2} = (\mathrm{Tr}(M^*\,M))^{1/2}$$

définit une norme comme on pourra le vérifier en exercice. c'est même une norme issue du produit scalaire $(A, B) \mapsto \mathrm{Tr}(B^*\,A)$. Mais elle n'est pas subordonnée car sinon la matrice unité aurait une norme égale à 1.

Parmi les normes sur l'espace des matrices on préfère celles qui sont *sous-multiplicatives* ce qui signifie que la propriété

$$||A\,B|| \leq ||A||\,||B||$$

est vraie pour toutes matrices A et B. Les normes subordonnées sont sous-multiplicatives. La norme $||\cdot||_e$ définie ci-dessus est sous-multiplicative (encore un bon exercice). Cela montre que les normes subordonnées ne sont pas les seules à être sous-multiplicatives. La norme définie par la formule

$$||M||_s = \sup_{i,j} |m_{ij}|$$

n'est pas sous-multiplicative. Il suffit de prendre pour A et B la matrice dont tout coefficient vaut 1.

10.3 Groupe linéaire et orthogonal

Comme d'habitude nous désignons par $\mathbb{K}$ le corps des scalaires $\mathbb{R}$ ou $\mathbb{C}$.

Proposition 34 *Le groupe linéaire $\mathcal{GL}_n(\mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$*

Preuve. $\mathcal{GL}_n(\mathbb{K})$ est un ouvert comme image réciproque de l'ouvert $\mathbb{K}\setminus\{0\}$ par l'application continue déterminant. Pour la densité, prenons M non inversible. Remarquons que l'application $\det(M - z\,I)$ est polynomiale en z dont s'annule pour un nombre fini de valeurs au plus. Donc il existe $\varepsilon > 0$ tel que $\det(M - z\,I) \neq 0$ pour tout $z \in]0, \varepsilon[$. Or $M - z\,I$ tend vers M si z tend vers 0, montrant ainsi que M est la limite de matrices inversibles. Donnons un exemple d'application.

Proposition 35 *Pour toutes matrices A et B, AB et BA ont même polynome caratéristique.*

Preuve. Il s'agit donc de prouver que pour tout $\lambda \in \mathbb{K}$, $\det(AB - \lambda\,I) = \det(BA - \lambda\,I)$. Si B est inversible, c'est facile à voir : $\det(AB - \lambda\,I) = \det[B^{-1}(BA - \lambda I)B] = \det[B^{-1}]\det[BA - \lambda I]\det[B] = \det[BA - \lambda I]$ Si B n'est pas inversible, elle est la limite quand $z \downarrow 0$ des matrices inversibles $B_z = B - z\,I$ pour lesquelles la relation est vraie : $\det(AB_z - \lambda\,I) = \det(B_zA - \lambda\,I)$. Il suffit donc de passer à la limite dans cette égalité en utilisant la continuité des deux membres par rapport à B_z pour obtenir le résultat souhaité.

Proposition 36 *L'application $A \mapsto A^{-1}$ de $\mathcal{GL}_n(\mathbb{K})$ dans lui-même est continue*

Preuve. Il suffit de se rappeler que $A^{-1} = (1/\det A)\ {}^t\mathrm{cof}\, A$ et que les applications déterminants intervenant dans cette formule sont continues.

Nous passons maintenant à l'étude du groupe orthogonal de $\mathcal{M}_n(\mathbb{R})$

Proposition 37 *Le groupe orthogonal $\mathcal{O}(n)$ est un compact de $\mathcal{M}_n(\mathbb{R})$.*

Preuve. Il suffit de remarquer qu'il est fermé et borné. Une petite application pour suivre.

Proposition 38 *Toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ peut s'écrire $M = O\,S$ où O est orthogonal et S symétrique*

Preuve. Le résultat est déjà connu si M est inversible. C'est la décomposition polaire prouvée à l'exercice 46 du chapitre consacré à la réduction des endomorphismes normaux. Dans ce cas la matrice S est même définie positive et il y a unicité de la décomposition. Si M n'est pas inversible, elle est néanmoins limite des matrices M_k inversibles. Pour ces matrices inversibles on peut écrire la décomposition polaire : $M_k = O_k\, S_k$. De la suite (O_k) on peut extraire une sous-suite $(O_{\phi(k)})$ convergeant vers la matrice orthogonale O. Alors $S_{\phi(k)} = O_{\phi(k)}^{-1}\, M_{\phi(k)}$ converge vers $O^{-1}\, M$ que l'on notera S. Cette matrice est symétrique comme limite de matrices symétriques. On a donc en définitive $M = O\,S$ comme souhaité. Notons qu'il n'y a pas unicité de la décomposition dans ce cas. Ce résultat est prolongé par l'exercice 54.

10.4 Matrices diagonalisables et trigonalisables

Comme toujours $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Proposition 39 *L'adhérence de l'ensemble $\mathcal{D}_n(\mathbb{K})$ des matrices diagonalisables de $\mathcal{M}_n(\mathbb{K})$ est égal à l'ensemble $\mathcal{T}_n(\mathbb{K})$ des matrices trigonalisables de $\mathcal{M}_n(\mathbb{K})$. En particulier $\mathcal{D}_n(\mathbb{C})$ est dense dans $\mathcal{T}_n(\mathbb{C}) = \mathcal{M}_n(\mathbb{C})$.*

L'intérieur de $\mathcal{D}_n(\mathbb{K})$ est égal à l'ensemble $\mathcal{V}_n(\mathbb{K})$ des matrices diagonalisables à valeurs propres simples.

Preuve. Montrons d'abord que toute matrice trigonalisable est limite de matrices diagonalisables et même de matrices diagonalisables à valeurs propres simples. Soit M une matrice trigonalisable i.e. $M = P^{-1}\,T\,P$ où T est triangulaire supérieure. Notons $t_{11}, \ldots, t_{nn}$ les termes diagonaux de T. Soit

$$\eta = \frac{1}{n}\inf\{|t_{pp} - t_{qq}|; 1 \leq p, q \leq n,\ t_{pp} \neq t_{qq}\}$$

Si tous les termes diagonaux $t_{11}, \ldots, t_{nn}$ sont égaux on pose simplement $\eta = 1$; la discussion qui suit vaut également dans ce cas plus facile que nous exclurons dorénavant. Pour $k \in \mathbb{N}^*$ considérons la matrice T_k de terme général $t_{ij}^k = t_{ij} + \mathbf{1}_{\{i=j\}}\eta\, i/k$. La matrice T_k est triangulaire supérieure avec des termes diagonaux tous distincts. En effet si $t_{ii}^k = t_{jj}^k$ avec $i \neq j$ alors $t_{ii} + \eta\, i/k = t_{jj} + \eta\, j/k$ ce qui entraine déja $t_{ii} \neq t_{jj}$ puis

$$|t_{jj} - t_{ii}| = \eta|j-i|/k \leq \left(\inf_{t_{pp}\neq t_{qq}} |t_{pp} - t_{qq}|\right)\frac{|j-i|}{k\,n} < \inf_{t_{pp}\neq t_{qq}} |t_{pp} - t_{qq}|$$

ce qui est absurde. La matrice T_k est diagonalisable à valeurs propres simples. Il en est de même des matrices $P^{-1}\,T_k\,P$ et ces matrices convergent vers $P^{-1}\,T\,P = M$.

Maintenant il s'agit de prouver que toute suite convergente (M_k) de matrices diagonalisables converge vers une limite M trigonalisable. Nous pouvons supposer $\mathbb{K} = \mathbb{R}$ car dans le cas complexe il n'y a rien à démontrer puisque toute matrice complexe est trigonalisable. Nous savons aussi par la première partie de la preuve qu'il existe des matrices diagonalisables B_k à valeurs propres simples $\lambda_1^k, \ldots, \lambda_n^k$ telles que $||M_k - B_k|| \leq 1/k$. Ainsi M est aussi la limite de (B_k). Nous travaillerons avec une norme de matrice $||\cdot||$ subordonée. Alors $|\lambda_i^k| \leq ||B_k||$. Donc les suites $(\lambda_i^k)_{k\geq 1}$ pour $1 \leq i \leq n$ sont bornées et on peut trouver une extraction qui les fait converger i.e. $\lambda_i^{\phi(k)} \to \lambda_i$ pour $1 \leq i \leq n$. Comme le polynome caractéristique dépend continument de la matrice sur laquelle il est calculé, on a

$$C_M(X) = \lim_{k\to+\infty} C_{B_{\phi(k)}}(X) = \lim_{k\to+\infty} \prod_{i=1}^{n}(X - \lambda_i^{\phi(k)}) = \prod_{i=1}^{n}(X - \lambda_i)$$

Dans cette dernière égalité nous avons simplement utilisé la continuité des coefficients d'un polynome par rapport à ses racines. Au total on trouve que C_M est scindé ce qui prouve bien que M est trigonalisable.

Montrons maintenant que toute matrice $V \in \mathcal{V}_n(\mathbb{K})$ est à l'intérieur de $\mathcal{D}_n(\mathbb{K})$. Sinon il existe une suite de matrices (W_k) qui ne sont pas dans $\mathcal{D}_n(\mathbb{K})$ et qui convergent vers V. Le polynome caractéristique de W_k admet des racines complexes $\lambda_1^k, \ldots, \lambda_n^k$ écrites avec éventuelles répétitions. Par un argument de compacité que nous avons déjà explicité ci-dessus, on peut trouver une sous-suite telle que $(\lambda_1^{\phi(k)}, \ldots, \lambda_n^{\phi(k)})$ converge vers un certain n-uple $(\lambda_1, \ldots, \lambda_n)$. Mais pour $i \in \{1, \ldots, n\}$ on a $C_{W_{\phi(k)}}(\lambda_i^{\phi(k)}) = 0$ d'où en passant à la limite $C_V(\lambda_i) = 0$. Par conséquent, $\lambda_1, \ldots, \lambda_n$ sont les racines du polynome caractéristique de V et ce sont donc n éléments distincts de $\mathbb{K}$. Nous pouvons trouver $\varepsilon > 0$ tel que les boules $B(\lambda_1, \varepsilon), \ldots, B(\lambda_n, \varepsilon)$, dans $\mathbb{C}$, soient disjointes. Pour k suffisamment grand, disons $k \geq k_0$, nous avons $\lambda_i^{\phi(k)} \in B(\lambda_i, \varepsilon)$. Cela entraine que les $\lambda_1^{\phi(k)}, \ldots, \lambda_n^{\phi(k)}$ sont distincts. Il s'agit, rappelons le, des racines complexes du polynome caractéristique de $W_{\phi(k)}$. Si $\mathbb{K} = \mathbb{C}$ cela entraine la diagonalisabilité de $W_{\phi(k)}$ ce qui est contraire à l'hypothèse. Dans le cas où $\mathbb{K} = \mathbb{R}$ pour $i \in \{1, \ldots, n\}$ et $k \geq k_0$, la racine $\lambda_i^{\phi(k)}$ du polynome caractéristique $C_{W_{\phi(k)}}$ à coefficients réels appartient à la boule centrée en le réel λ_i et de rayon ε. Le conjugué de cette racine est alors aussi racine et dans cette même boule. Il y a donc égalité entre la racine et la racine conjuguée car on ne peut avoir de racine supplémentaire pour $C_{W_{\phi(k)}}$

qui est de degré n et a déjà n racines. Donc en fait les $\lambda_1^{\phi(k)}, \dots, \lambda_n^{\phi(k)}$ sont des réels distincts ce qui contredit l'hypothèse que $W_{\phi(k)}$ est non diagonalisable.

Maintenant supposons que $H \in \mathcal{D}_n(\mathbb{K}) \backslash \mathcal{V}_n(\mathbb{K})$. La matrice H admet donc une valeur propre λ qui n'est pas simple. Elle s'écrit $P^{-1}\ D\ P$ où la matrice D est une matrice diagonale de diagonale $\lambda, \dots, \lambda, d_1, \dots, d_p$. Considérons la matrice D_ε obtenue en rajoutant à D le coefficient ε sur la première ligne en deuxième colonne. Alors quand $\varepsilon \downarrow 0$, $P^{-1}\ D_\varepsilon\ P$ converge vers $P^{-1}\ D\ P = H$. Mais $P^{-1}\ D_\varepsilon\ P$ n'est pas diagonalisable car son polynome minimal, égal à celui de D_ε est divisible par $(X - \lambda)^2$ et il n'est donc pas simplement scindé. Ainsi H est limite de matrices non diagonalisables ce qui prouve que H n'est pas dans l'intérieur de $\mathcal{D}_n(\mathbb{K})$.

10.5 Exercices

Exercice 52 Soit un entier p montrer que l'ensemble des matrices carrées $n \times n$ de rang supérieur à p est un ouvert.

Exercice 53 En utilisant le résultat de l'exercice 36, montrer que si $A = (a_{ij})_{1 \leq i,j \leq n}$ est une matrice symétrique positive alors la matrice de terme général $e^{a_{ij}}$ est aussi symétrique positive.

Exercice 54 On sait par l'exercice 46 que pour toute matrice inversible M, il existe un unique couple (S, O) où S est symétrique définie positive et O orthogonale tel que $M = O\,S$. Prouver que l'application qui à M associe le couple (S, O) est une homéomorphie, c'est à dire que c'est une bijection continue dont la réciproque est continue.

Exercice 55 (Preuve de la proposition 40) 1) Prouver d'abord que si $||\cdot||$ est une norme subordonnée à une norme vectorielle et λ une valeur propre de la matrice complexe A alors $|\lambda| \leq ||A||$.
2) Montrer que si P est une matrice inversible et $||\cdot||$ une norme subordonnée, alors l'application $M \mapsto ||P^{-1}\,M\,P||$ est une autre norme subordonnée.
3) Soit $\varepsilon > 0$ et T une matrice triangulaire supérieure et diagonale à ε près c'est à dire $|t_{ij}| \leq \varepsilon$ pour $j > i$. Trouver alors une norme subordonnée $||\cdot||$ telle que $||T|| \leq \max\{t_{ii},\ 1 \leq i \leq n\} + \varepsilon$.
4) En se rappelant l'exercice 11, montrer que pour $\varepsilon > 0$ et A matrice complexe, trouver une norme subordonnée $||\cdot||$ telle que $||A|| \leq \rho(A) + \varepsilon$.

Exercice 56 (Preuve de la proposition 42) 1) Prouver d'abord que si $||\cdot||$ est une norme sous-multiplicative et λ une valeur propre de la matrice complexe A alors $|\lambda| \leq ||A||$.
Indication : si V est un vecteur propre associé à λ considérer la matrice W dont toutes les colonnes sont V et calculer $A\,W$.
2) En déduire que $\rho(A) \leq \liminf_{k \to +\infty} ||A^k||^{1/k}$.
3) Soit $\varepsilon > 0$, calculer $\rho(B)$ où B est la matrice $(1/(\rho(A) + \varepsilon))\,A$.
4) En déduire que $(||B^k||)$ est bornée puis que $\limsup_{k \to +\infty} ||A^k||^{1/k} \leq \rho(A) + \varepsilon$.
5) Conclure.

Exercice 57 Prouver que $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable si et seulement si sa classe de similitude

$$\{P^{-1}\,A\,P;\ \ P \in \mathcal{GL}_n(\mathbb{C})\}$$

est fermée.

Chapitre 11

Quelques utilisations de l'analyse

11.1 Méthodes analytiques pour les systèmes

Nous commençons par introduire la notion de rayon spectral puis évoquons des applications.

Nous appelons *rayon spectral* $\rho(M)$ d'une matrice M réelle ou complexe le plus grand module des valeurs propres complexes de cette matrice. Notons déja que le rayon spectral peut être nul sans que la matrice soit nulle, comme dans le cas des matrices nilpotentes. Ce n'est donc pas une norme. Toutefois le rapport avec les normes est précisé dans le résultat suivant.

Proposition 40 *Le rayon spectral d'une matrice est égal à l'infimum des normes de cette matrice, pour toutes les normes subordonnées à une norme vectorielle.*

Preuve. Elle est l'objet de l'exercice 55. On en déduit très facilement le résultat suivant.

Proposition 41 *Pour $A \in \mathcal{M}_n(\mathbb{C})$ il y a équivalence entre les assertions suivantes :*

1. $\lim_{k \to +\infty} A^k = 0$
2. $\lim_{k \to +\infty} A^k x = 0$ *pour tout* $x \in \mathbb{C}^n$
3. $\rho(A) < 1$
4. *il existe une norme de matrice subordonnée* $||\cdot||$ *telle que* $||A|| < 1$.

Proposition 42 *Pour toute norme* $||\cdot||$ *sous-multiplicative sur* $\mathcal{M}_n(\mathbb{C})$ *et pour toute matrice* $A \in \mathcal{M}_n(\mathbb{C})$*, on a*

$$\rho(A) = \lim_{k\to+\infty} ||A^k||^{1/k}$$

La preuve de ce résultat est proposée à l'exercice 56.

Maintenant, considérons le système linéaire $Ax = b$ où A est une matrice $n \times n$ inversible. Supposons que A s'écrive $A = M - N$ où M est une matrice inversible facile à inverser. Alors le système $Ax = b$ équivaut à $x = M^{-1}N\,x + M^{-1}b$. La méthode des approximations successives est basée sur la suite (x_k) dont le terme initial est quelconque et qui vérifie la relation de récurrence

$$x_{k+1} = M^{-1}N\,x_k + M^{-1}b$$

Proposition 43 *La suite* (x_k) *converge vers la solution* x *du système si et seulement si* $\rho(M^{-1}N) < 1$*.*

Preuve. On a $x_k - x = (M^{-1}N)^k\,(x_0 - x)$ et il suffit alors d'utiliser la proposition 41.

La méthode de *Jacobi* consiste à prendre pour M la diagonale de A ce qui, bien sûr, ne peut se faire que si tous ces coefficients diagonaux sont non nuls. Dans la méthode de *Gauss-Seidel* on prend pour M la matrice triangulaire inférieure qui correspond aux termes sous la diagonale de A. Les méthodes de Jacobi et de Gauss-Seidel convergent par exemple quand A est à diagonale strictement dominante.On peut montrer aussi que la méthode de Gauss-Seidel converge pour A définie positive.

Dans le cas d'une matrice A définie positive, on peut aussi se ramener à une problème de minimisation.

Proposition 44 *Soit* A *une matrice* $n\times n$ *symétrique réelle définie positive et* $b \in \mathbb{R}^n$*. Alors l'application* φ *définie de* $\mathbb{R}^n$ *dans* $\mathbb{R}$ *par*

$$\varphi(x) = \frac{1}{2}\,{}^t xAx - {}^t b\,x$$

admet un unique extremum local qui est un minimum global au point u *solution du système* $Au = b$*.*

Preuve. En utlisant la relation $Au = b$, on peut écrire

$$\varphi(x) = \frac{1}{2}\,{}^t(x-u)A(x-u) - \frac{1}{2}\,{}^t uAu$$

Comme A est définie positive on a si $x-u\neq 0$, $\varphi(x) > -(1/2)^t uAu = \varphi(u)$, ce qui montre déjà que u réalise l'unique minimum global. Inversement remarquons que φ est différentiable de différentielle

$$\varphi'(x)\cdot h = {}^t x\, A\, h - {}^t b\, h = {}^t(Ax-b)\, h$$

Si φ admet un extremum local en un point u alors $\varphi'(u)$ est la forme linéaire nulle donc $Au-b=0$.

La résolution du système linéaire est donc ramenée à la recherche du minimum d'une fonctionnelle quadratique, problème pour lequel il existe des techniques standard. Décrivons brièvement la *méthode du gradient*. Cette méthode, qui fait partie des *méthodes de descente*, construit une suite de vecteurs (x_n) qui converge vers le minimum. Pour trouver x_{n+1} à partir de x_n, on choisit une direction de descente δ_n et on regarde sur la droite passant par x_n et de direction δ_n le point où le minimum de φ est atteint. Cela est facile car il s'agit d'une minimisation selon une variable réelle. Le point trouvé est x_{n+1}. Reste à préciser le choix de la direction de descente. Il y a une certaine liberté pour ce choix. Toutefois le bon sens nous dit que la direction qui parait la meilleure est celle selon laquelle la variation de φ est la plus forte c'est à dire le gradient.

Détaillons les calculs. Les valeurs de φ sur la droite passant par x_n et de direction δ_n sont données par

$$\varphi(x_n + t\,\delta_n) = \frac{1}{2}\,{}^t\delta_n A\delta_n\, t^2 + {}^t(Ax_n - b)\, t + \varphi(x_n)$$

Le minimum est atteint pour $t_0 = -{}^t(Ax_n-b)\delta_n/{}^t\delta_n A\delta_n$. Le point x_{n+1} est $x_{n+1} = x_n + t_0\,\delta_n$. Nous choisissons pour δ_n le gradient de φ en x_n qui est $Ax_n - b$. Avec ces choix (x_n) converge vers la solution u de $Au=b$ et en plus $|x_n - u|_2$ tend vers 0 à la vitesse de $(1-\mathrm{cond}(A))^{n/2}$. La quantité $\mathrm{cond}(A) = 1/(||A||_2||A^{-1}||_2) \in]0,1]$ est aussi égale à la racine du rapport de la plus grande et de la plus petite valeurs propres de A.

11.2 L'exponentielle de matrice

Considérons sur l'espace des matrices carrées $d\times d$ une norme sous-multiplicative $||\cdot||$ (subordonnée ou pas). Soit une matrice M. Alors pour tout entier n on a $||M^n|| \leq ||M||^n$ puisque la norme est sous-multiplicative. Il s'en suit que la série numérique $\sum_n \frac{||M^n||}{n!}$ est convergente. On en déduit par complétude que la série $\sum_n \frac{M^n}{n!}$ converge. On définit donc une matrice, dite *exponentielle* de la matrice M, en posant

$$\exp(M) = \sum_{k=0}^{+\infty} \frac{M^k}{k!}$$

Il est immédiat que si M est diagonale de coefficients diagonaux $\mu_1, \dots, \mu_d$ alors $\exp(M)$ est également diagonale de coefficients diagonaux $e^{\mu_1}, \dots, e^{\mu_d}$. Par exemple l'exponentielle de la matrice nulle est la matrice unité. On notera la propriété suivante qui se vérifie immédiatement :

$$P^{-1} \exp(M)\, P = \exp(P^{-1}\, M\, P)$$

valable pour toute matrice M et toute matrice inversible P. Il est également facile de voir que si M est triangulaire supérieure de coefficients diagonaux $\mu_1, \dots, \mu_d$ alors $\exp(M)$ est aussi triangulaire supérieure de coefficients diagonaux $e^{\mu_1}, \dots, e^{\mu_d}$. Si M est diagonale par blocs alors son exponentielle est diagonale par blocs avec des blocs diagonaux qui sont les exponentielles des blocs diagonaux de la matrice de départ. Par conséquent si on dispose de la réduction de Jordan d'une matrice, le calcul de l'exponentielle revient au calcul de l'exponentielle des blocs de Jordan. Nous y reviendrons bientôt. Prouvons pour le moment que l'exponentielle de matrice a la même propriété de morphisme que l'exponentielle réelle.

Proposition 45 *Pour toutes matrices carrées A et B qui commutent, on a*

$$\exp(A + B) = \exp(A)\, \exp(B)$$

Preuve. On a

$$\begin{aligned}
\sum_{k=0}^{N} \frac{(A+B)^k}{k!} &= \sum_{k=0}^{N} \frac{1}{k!} \sum_{i+j=k,\ i,j\geq 0} \mathrm{C}_k^i\, A^i\, B^j \\
&= \sum_{i,j\geq 0} \frac{A^i}{i!}\, \frac{B^j}{j!} \sum_{k=0}^{N} \mathbf{1}(i+j=k) \\
&= \sum_{i,j\geq 0} \frac{A^i}{i!}\, \frac{B^j}{j!}\, \mathbf{1}(0 \leq i+j \leq N)
\end{aligned}$$

On remplace N par $2\,N$. Alors

$$\sum_{k=0}^{2N} \frac{(A+B)^k}{k!} = \sum_{0\leq i,j\leq N} \frac{A^i}{i!}\, \frac{B^j}{j!} + R = \left(\sum_{0\leq i\leq N} \frac{A^i}{i!}\right) \left(\sum_{0\leq i\leq N} \frac{A^i}{i!}\right) + R$$

avec

$$R = \sum_{i,j} \frac{A^i}{i!} \frac{B^j}{j!} \mathbf{1}(i + j \leq 2N \text{ et } (i > N \text{ ou } j > N))$$

Par une majoration directe

$$\begin{aligned}
||R|| &\leq \sum_{i,j} \frac{||A||^i}{i!} \frac{||B||^j}{j!} (\mathbf{1}(i > N) + \mathbf{1}(j > N)) \\
&\leq \left(\sum_{i>N} \frac{||A||^i}{i!} \right) e^{||B||} + \left(\sum_{j>N} \frac{||B||^j}{j!} \right) e^{||A||}
\end{aligned}$$

On fait tendre N vers $+\infty$. La quantité R tend vers 0 et on obtient le résultat annoncé.

Une conséquence de la proposition qui précède est que l'exponentielle d'une matrice est inversible. Cela se déduit de $\exp(M) \exp(-M) = I$. On peut même voir que, pour toute matrice M,

$$\det \exp(M) = e^{\mathrm{Tr}(M)}$$

Pour cela il suffit de triangulariser la matrice M, ce qui est possible dans le corps des complexes.

Revenons au calcul de l'exponentielle d'un bloc de Jordan de taille r que nous écrirons $\lambda I + N$. Alors $\exp(\lambda I + N) = \exp(\lambda I) \exp(N) = e^\lambda \exp(N) = e^\lambda \left(\sum_{k=0}^{r-1} N^k / k!\right)$ puisque N est nilpotente d'indice r. Cela montre qu'en définitive, le calcul de la réduite de Jordan permet ensuite d'obtenir facilement l'exponentielle.

11.3 Exercices

Exercice 58 Calculer $\exp(M)$ dans les cas suivants :

$$M = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}, \quad M = \begin{pmatrix} 7 & -5\sqrt{3} \\ -5\sqrt{3} & -3 \end{pmatrix}$$

Exercice 59 Procéder à la réduction de Jordan de M puis calculer $\exp(tM)$ pour tout $t \in \mathbb{R}$ dans les cas suivants :

$$M = \begin{pmatrix} 3 & 8 & -7 \\ 0 & 7 & -4 \\ 0 & 0 & 3 \end{pmatrix}, \quad M = \begin{pmatrix} 2-i & 1 & 0 & 0 \\ 1 & 2+i & 0 & 0 \\ 2 & 6i & 0 & 4 \\ 1-i & 1+3i & -1 & 4 \end{pmatrix}$$

Exercice 60 Montrer que l'exponentielle réalise une surjection continue de l'ensemble des matrices anti-symétriques réelles sur l'ensemble des matrices orthogonales réelles de déterminant 1.

Exercice 61 Montrer que l'exponentielle réalise un homéomorphisme de l'ensemble des matrices symétriques réelles sur l'ensembles des matrices symétriques définies positives.

Exercice 62 Soit M une matrice carrée complexe. Montrer qu'il existe un polynome P tel que $\exp(M) = P(M)$.

Exercice 63 Quelles sont les matrices carrées complexes M telles que $\exp(M) = I$?

Chapitre 12

Auto-évaluation

12.1 Indications ou réponses pour les exercices

Exercice 1 Si $U \subset W$ c'est fini ; sinon il existe $u_0 \in U \setminus W$ donc $u_0 \in V \setminus W$. Montrer alors que tout $u \in U$ appartient à V. Pour cela procéder par l'absurde en utilisant les égalités : $u = u_0 + (u - u_0)$, $u_0 = u - (u - u_0)$.

Exercice 2 Soit $(b_1, \ldots, b_n)$ une base de E. Tout $x \in E$ s'écrit

$$x = x_1\, b_1 + \cdots + x_n\, b_n = (x_1 + y) b_1 + \cdots + (x_n + y)\, b_n + y\,(-b_1 - \cdots - b_n)$$

Pour un bon choix de y à préciser, cela permet de trouver une famille de $n+1$ vecteurs qui est positivement génératrice. Achever en prouvant qu'une famille de n vecteurs ne peut pas être positivement génératrice.

Exercice 3 Ecrire $P = U + iV$ où $U, V \in \mathcal{M}_d(\mathbb{R})$. Ecrire alors les deux relations fournies par l'égalité $PB = AP$. Pour $Q = U + tV$, $t \in \mathbb{R}$ montrer que $QB = AQ$. Comment trouver t tel que Q soit inversible ?

Exercice 4 Déduire de l'hypothèse que pour tout $x \in E$, il existe $c(x) \in \mathbb{K}$ tel que $u(x) = c(x)\, x$. Montrer que $c : E \to \mathbb{K}$ est constante i.e. $c(x) = c(y)$ pour tous x, y en séparant les deux cas x et y colinéaires et $\{x, y\}$ libre.
Pour la stabilisation des hyperplans, se ramener à la question précédente en remarquant qu'une droite est l'intersection des hyperplans qui la contiennent.

Exercice 5 1) Retrancher une ligne (ou une colonne) à toutes les autres puis développer selon cette ligne (ou colonne).
2) Prendre deux valeurs judicieuses pour t.
3) A partir du calcul précédent, passer à la limite $c \to a$.

Exercice 7 Ecrire la définition de λ valeur propre avec un vecteur propre $x = (x_1, \ldots, x_n)$. Exprimer la relation obtenue sur la i_0-ième coordonnée où i_0 est choisi pour maximiser $|x_i|$.

Exercice 8 Montrer que 1 et -1 sont des valeurs propres avec des espaces propres associés supplémentaires.

Exercice 11 Soit u l'endomorphisme canoniquement associé à A ; prendre une base de trigonalisation de u et multiplier chacun de ces vecteurs par un coefficient bien choisi.

Exercice 15 A quel type de polynôme P peut on se ramener ? La preuve se fait alors par un procédé bien connu. Avec les 0 sur la diagonale, voir qu'au fur et à mesure des itérations en puissance, les zéros remplissent la matrice vers le coin haut droit. On pourra se convaincre par un exemple. Ecrire comme propriété sur le terme général et prouver l'assertion par récurrence.

Exercice 16 Ecrire la matrice de u dans une base adaptée à la décomposition en somme directe. Voir que les polynômes de cette matrice diagonale par blocs s'obtiennent en prenant les polynômes des blocs diagonaux ; on pourra généraliser à partir des puissances. En déduire que $P(u) = 0$ si $P(u|_{F_i}) = 0$ pour toutes les restrictions de u à un sous-espace F_i. En déduire le polynôme minimal. Pour le polynôme caractéristique faire un calcul direct de déterminant à partir de la matrice diagonale par blocs.

Exercice 17 3) Comme $M_\lambda = M_0 + \lambda I$, voir que $P(X) \in \mathrm{Ann}(M_\lambda)$ si et ssi $P(X + \lambda) \in \mathrm{Ann}(M_0)$.
4) Se ramener à $\lambda = 0$. Soit u l'endomorphisme de matrice M_0 dans la base $(b_1, \ldots, b_n)$, écrire la matrice de u dans la base $(f_1, \ldots, f_n) = (b_n, \ldots, b_1)$.

Exercice 18 Donner un contre-exemple avec deux matrices 4×4 de polynôme carartéristique X^4 et de polynôme minimal X^2. L'une sera composée de deux blocs diagonaux 2×2 du type M_0 avec la notation de l'exercice 17.

Exercice 21 V,V,F

Exercice 22 Toutes les assertions proposées sont équivalentes à la nilpotence de A.

Exercice 24 a) Penser au critère avec le polynome minimal.

Exercice 26 Vérifier que la première matrice est nilpotente d'indice 4.

La seconde a pour polynôme caractéristique $(X-1)^3\,(X+1)^2$. En déterminant les dimensions des sous-espaces propres, ce qui n'est pas dur, on peut déja annoncer la forme de la réduite de Jordan.

Exercice 27 En examinant les différentes possibilités pour (C_A, M_A), montrer que dans ces petites dimensions la connaissance de (C_A, M_A) détermine la forme de la réduite de Jordan.

Exercice 28 Se placer d'abord dans le cas complexe. Faire la réduction de Jordan et utiliser l'exercice 17 4).

Exercice 34 Astuce : Se rappeler que $1/(k+1) = \int_0^1 t^k\, dt$.

Exercice 35 Introduire $r_k = t_k - t_{k-1}$ et écrire

$$\sum_{i,j} \inf(t_i, t_j)\, x_i\, x_j = \sum_{i,j} [\sum_k r_k \mathbf{1}_{k \leq \inf(i,j)}]\, x_i\, x_j$$

Puis permuter les sommations.

Exercice 36 Par exemple réduire la matrice B avec une matrice de passage (inconnue) $P = (p_{ij})$; écrire ce que cela entraine sur l'expression du terme général b_{ij} à l'aide des p_{ij} ; utiliser cette écriture pour vérifier, selon la définition, le caractère positif.

Exercice 37 1) Faire une orthonormalisation de Gram-Schmidt à une base de trigonalisation.
2) Traduire 1) en termes matriciels.

Exercice 38 On pourra discuter les cas selon que le rang de S est n, $n-1$ ou inférieur ou égal à $n-2$. Dans le cas de rang n, la matrice des cofacteurs s'exprime simplement à l'aide de l'inverse. Dans le cas de rang $\leq n-2$, la matrice des cofacteurs est très simple !

Exercice 39 On pourra se ramener à une matrice orthogonale (pour laquelle le déterminant a une valeur absolue bien connue).

Exercice 46 1) Faire la réduction de U.
2) En faisant une réduction, cela ramène au cas diagonal. Utiliser alors un polynôme d'interpolation c'est à dire un polynôme qui prend des valeurs données en des points donnés.
3) Par 2) montrer que S et S' commutent. Utiliser alors l'exercice 24.

Exercice 47 Si u est normal, le résultat de réduction vu dans ce chapitre permet de se ramener à une matrice diagonale ; on pourra alors utiliser un polynome d'interpolation.

Exercice 48 1) Penser à une rotation.
2) On pourra remplacer x par $x+iy$ et $x+y$ pour x, y quelconques.

Exercice 49 Faire une réduction simultanée de H et K.

Exercice 51 Procéder par analyse et synthèse.

12.2 QCM

QUESTION 1. Quelle partie de $\mathbb{R}[X]$ n'est pas un sous-espace vectoriel de $\mathbb{R}[X]$?
(a) $\{P\,;\ \ P''=0\}$
(b) $\{P\,;\ \ P'(1)\,(1+P(-1)^2)=0\}$
(c) $\{P\,;\ \ P'(1)\,.\,P(1)=0\}$

QUESTION 2. Quelle partie de $\mathbb{R}[X]$ n'est pas un idéal de $\mathbb{R}[X]$?
(a) $\{P\,;\ \ P(0)=P'(0)=0\}$
(b) $\{P\,;\ \ \int_0^1 P(t)^2\,dt=0\}$
(c) $\{P\,;\ \ P''=0\}$

QUESTION 3. Soit u un endomorphisme dont la matrice sur une base $\mathcal{B}$ est A. Soit $\mathcal{H}$ une base telle que la matrice de passage de $\mathcal{B}$ à $\mathcal{H}$ est notée P. Alors la matrice de u sur la base $\mathcal{H}$ est
(a) $P^{-1}\ A\ P$
(b) $P\ A\ P^{-1}$
(c) ${}^tP\ A\ P$

QUESTION 4. Le déterminant

$$\begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix}$$

(a) est le célèbre déterminant de Vanderplouc
(b) s'annule dès que a_1, a_2, a_3 ne sont pas distincts
(c) vaut $\prod_{1\leq i,j\leq 3}(a_j - a_i)$

QUESTION 5. La matrice

$$\begin{pmatrix} 4 & 0 & 1 & -2 \\ 1 & 7 & 2 & -3 \\ -4 & 0 & -7 & 1 \\ 2 & 3 & 0 & 8 \end{pmatrix}$$

(a) est inversible
(b) est de rang 2
(c) est nilpotente

QUESTION 6. Soit u un endomorphisme dont la matrice dans une base (b_1, b_2) est

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

Dans quelle base a-t-il pour matrice

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

(a) $(\frac{1}{2}b_1, \frac{1}{2}b_2)$
(b) $(\frac{1}{2}b_1, 2\, b_2)$
(c) $(b_1, 2\, b_2)$

QUESTION 7. Soit la matrice triangulaire supérieure $T = (t_{ij})_{1\leq i,j\leq n}$. Alors
(a) pour tout polynôme P la matrice $P(T)$ est triangulaire supérieure de terme général égal à $P(t_{ij})$ pour $j \geq i$
(b) T est nilpotente si et seulement si $t_{ii} = 0$ pour tout $i \in \{1, \ldots, n\}$
(c) T est nilpotente si et seulement si $(\mathrm{Tr}(T))^k = 0$ pour tout $k \in \mathbb{N}^*$

QUESTION 8. Pour que $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifient $A\,B = B\,A$,
(a) il faut que B soit un polynôme en A
(b) il suffit que A et B soient toutes deux des polynômes de la même matrice
(c) il suffit que $\det(AB - BA) = 0$

QUESTION 9. La matrice

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

(a) est diagonalisable si et seulement si $\alpha = 0$
(b) est diagonalisable pour tout $\alpha \in \mathbb{R}$
(c) n'est diagonalisable dans $\mathcal{M}_n(\mathbb{R})$ pour aucune valeur du réel α.

QUESTION 10. Parmi les 3 matrices suivantes déterminer celle qui n'est pas trigonalisable dans $\mathcal{M}_n(\mathbb{R})$: (a) $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

QUESTION 11. Parmi les 3 polynômes ci-dessous déterminer celui qui peut être le polynôme minimal d'une matrice de $\mathcal{M}_n(\mathbb{R})$ diagonalisable dans

$\mathcal{M}_n(\mathbb{R})$:
(a) $X^3 - X$
(b) $X^4 - X^2$
(c) $X^2 - 2X + 1$

QUESTION 12. Parmi les matrices suivantes déterminer la matrice qui n'est semblable à aucune des deux autres
(a) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

QUESTION 13. Si une matrice carrée A s'écrit par blocs carrés sous la forme

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

Alors
(a) le polynôme caractéristique de A est donné en fonction de ceux de U et V par $C_A(X) = \mathrm{PPCM}(C_U(X), C_V(X))$
(b) le polynôme minimal de A est donné en fonction de ceux de U et V par $M_A(X) = \mathrm{PGCD}(M_U(X), M_V(X))$
(c) le polynôme minimal de A est donné en fonction de ceux de U et V par $M_A(X) = \mathrm{PPCM}(M_U(X), M_V(X))$

QUESTION 14. Dans la situation de la question précédente,
(a) A peut être diagonalisable sans que U et V le soient toutes deux
(b) A peut être trigonalisable sans que U et V le soient toutes deux
(c) A est diagonalisable [respectivement trigonalisable] si et seulement si U et V sont diagonalisables [respectivement trigonalisables]

QUESTION 15. Quelle matrice a un polynôme minimal différent de celui des deux autres matrices ?

(a) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

QUESTION 16. Parmi les trois assertions suivantes, une n'est pas équivalente aux deux autres. Laquelle ?
(a) p est un endomorphisme de l'espace vectoriel E de dimension finie qui vérifie $p \circ p = p$
(b) Dans une certaine base de E la matrice de p s'écrit par blocs

$$\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

où I est une matrice identité et 0 désigne un bloc carré de zéros.
(c) Le polynôme caractéristique de p s'écrit $C_p(X) = X^k\ (X-1)^{n-k}$ où n est la dimension de E et $k \in \{1, \dots, n-1\}$.

QUESTION 17. Parmi les trois polynômes ci-dessous, déterminer celui qui peut être le polynôme minimal d'une matrice réelle trigonalisable dans $\mathcal{M}_n(\mathbb{R})$ mais pas diagonalisable dans $\mathcal{M}_n(\mathbb{R})$:
(a) $X^2 - X$
(b) $X^3 - X^2$
(c) $X^3 + X$

QUESTION 18. On considère l'endomorphisme u de $\mathbb{R}^n$ ($n \geq 3$) dont la matrice sur la base canonique est

$$\begin{pmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & & \vdots \\ 1 & 1 & 2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & 2 \end{pmatrix}$$

Alors
(a) $\ker(u - id)$ est une droite vectorielle
(b) $\ker(u - id)$ est un hyperplan c'est à dire un sous-espace de dimension $n - 1$
(c) 1 n'est pas valeur propre de u

QUESTION 19. Dans la situation de la question précédente,
(a) le polynôme carac téristique C_u de u n'est pas scindé.
(b) C_u est scindé avec pour seule racine 1 et u est trigonalisable sans être diagonalisable.
(c) u est diagonalisable.

QUESTION 20. Dans un seul des trois cas ci-dessous la donnée du polynôme minimal M_A et du polynôme caractéristique C_A de la matrice A permet de déterminer la réduite de Jordan de A. Lequel ?
(a) $C_A(X) = (X-1)^3\,(X+1)^2\,,\quad M_A(X) = (X-1)^2\,(X+1)$
(b) $C_A(X) = (X-1)^4\,(X+1)\,,\quad M_A(X) = (X-1)^2\,(X+1)$
(c) $C_A(X) = (X+1)^5\,,\quad M_A(X) = (X+1)^3$

QUESTION 21. La matrice

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

(a) est inversible
(b) admet une valeur propre réelle non nulle
(c) n'est pas diagonalisable dans $\mathcal{M}_3(\mathbb{R})$

QUESTION 22. Quelle matrice a un polynôme minimal différent de celui des deux autres matrices ?

(a) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

QUESTION 23. Dans un seul des trois cas ci-dessous la donnée du polynôme minimal M_A et du polynôme caractéristique C_A de la matrice A permet de déterminer la réduite de Jordan de A. Lequel ?
(a) $C_A(X) = (X-1)^4\,(X+1)\,,\ \ M_A(X) = (X-1)^2\,(X+1)$
(b) $C_A(X) = (X-1)^3\,(X+1)^2\,,\ \ M_A(X) = (X-1)^2\,(X+1)$
(c) $C_A(X) = (X+1)^5\,,\ \ M_A(X) = (X+1)^3$

QUESTION 24. Parmi les trois polynômes ci-dessous, déterminer celui qui peut être le polynôme minimal d'une matrice réelle trigonalisable dans $\mathcal{M}_n(\mathbb{R})$ mais pas diagonalisable dans $\mathcal{M}_n(\mathbb{R})$:
(a) $X^2 + X$
(b) $X^3 + X^2$
(c) $X^3 + 2X$

QUESTION 25. Soit u un endomorphisme de l'espace vectoriel E de dimension 5. Parmi les trois cas de figure ci-dessous, un est impossible, lequel ?
(a) $\dim\ker u^3 = 5$, $\dim\ker u^2 = 4$, $\dim\ker u = 2$
(b) $\dim\ker u^3 = 5$, $\dim\ker u^2 = 5$, $\dim\ker u = 3$
(c) $\dim\ker u^3 = 5$, $\dim\ker u^2 = 3$, $\dim\ker u = 2$

QUESTION 26. Soit q une forme quadratique dont la matrice dans une base (b_1, b_2) est

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

Dans quelle base a-t-elle pour matrice

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

(a) $(b_1, \sqrt{2}\, b_2)$
(b) $(b_1, 2\, b_2)$
(c) $(\sqrt{2}\, b_1, \sqrt{2}\, b_2)$

QUESTION 27. On considère la forme quadratique q sur $\mathbb{R}^3$ donnée par $q((x, y, z)) = (x - y)^2 + (y - z)^2 + (x - z)^2$.
(a) Cette forme quadratique est de signature $(3, 0)$.
(b) Cette forme quadratique est positive mais pas définie positive.
(c) La matrice de cette forme quadratique sur la base canonique de $\mathbb{R}^3$ n'est pas diagonalisable.

QUESTION 28. Dans un espace euclidien (de dimension finie) E, une projection orthogonale sur un sous-espace vectoriel $F \neq E$ est
(a) orthogonale
(b) symétrique
(c) anti-symétrique

QUESTION 29. Pour $M \in \mathcal{M}_n(\mathbb{R})$ et $A = {}^tM\, M$,
(a) A n'est pas toujours symétrique définie positive
(b) si A est symétrique définie positive alors M est triangulaire supérieure à coefficients diagonaux strictement positifs
(c) si $\lambda \in \mathbb{R}_+$ est valeur propre de A alors $\sqrt{\lambda}$ est valeur propre de M.

QUESTION 30. On se place dans un espace hermitien c'est à dire un espace vectoriel complexe de dimension finie muni d'un produit scalaire. Une seule des assertions suivantes n'est pas toujours vraie. Laquelle ?
(a) tout endomorphisme est trigonalisable dans une base orthonormale.
(b) tout endomorphisme unitaire est de trace nulle.
(b) tout endomorphisme auto-adjoint est diagonalisable dans une base orthonormale avec des valeurs propres réelles.

12.3 Quelques problèmes de révision

Exercice 64 (Examen de février 2001) Dans tout le problème E désigne un espace hermitien. Il s'agit donc d'un espace vectoriel sur $\mathbb{C}$ de dimension finie muni d'un produit scalaire c'est à dire d'une forme sesqui-linéaire à symétrie hermitienne définie positive. Cette dernière sera notée $\langle \cdot, \cdot \rangle$. La norme associée sera notée $||\cdot||$. Dans tout le problème u désigne un endomorphisme de E c'est à dire une application linéaire de E dans E. L'adjoint de u noté u^* est l'unique endomorphisme de E qui vérifie $\langle u(x), y \rangle = \langle x, u^*(y) \rangle$ pour tous $x, y \in E$. On dit que u est normal si et seulement si $u \circ u^* = u^* \circ u$. On rappelle, comme il a été prouvé en cours, que si u est normal alors u est diagonalisable dans une base orthonormale c'est à dire qu'il existe une base orthonormale de E formée de vecteurs propres de u.

Si A est une matrice carrée complexe on note A^* son adjointe c'est à dire la matrice obtenue en transposant et en changeant chaque coefficient en son conjugué. On dit que A est normale si et seulement si $A\,A^* = A^*\,A$.
1) Montrer qu'il y a en fait équivalence entre "u est normal" et "u est diagonalisable dans une base orthonormale".
2) Si u est normal, comparer les valeurs propres de u^* et celles de u. Que dire concernant les sous espaces propres ?
3.a) Soit $\alpha_1, \ldots, \alpha_p$ des nombres complexes distincts et $\mu_1, \ldots, \mu_p$ des nombres complexes. Trouver une polynome à coefficients complexes H tel que $H(\alpha_i) = \mu_i$ pour tout $i \in \{1, \ldots, p\}$. Est il unique ? Y a-t-il existence et unicité si on impose à H d'être de degré $p-1$.
3.b) Prouver que l'endomorphisme u est normal si et seulement si u^* est un polynome en u.
4.a) Montrer que si l'endomorphisme u est normal alors l'image $\mathrm{Im}\, u$ de u et son noyau $\ker u$ sont supplémentaires et orthogonaux dans E.
4.b) En déduire que si u et v sont deux endomorphismes normaux alors $u \circ v = 0$ si et seulement si $v \circ u = 0$.
5) Soit un nombre complexe α et A une matrice carrée complexe $n \times n$ normale telle que, pour tout $i \in \{1, \ldots, n\}$, $\sum_{j=1}^{n} a_{ij} = \alpha$. Prouver qu'alors pour tout $j \in \{1, \ldots, n\}$, $\sum_{i=1}^{n} a_{ij} = \alpha$.
Indication. On pourra par exemple étudier l'effet de A sur le vecteur $\mathbf{1}$ dont toutes les composantes sont égales à 1.
6.a) Si u est normal, montrer que pour tout $x \in E$, $||u(x)|| = ||u^*(x)||$.
6.b) Réciproquement, on suppose que l'endomorphisme u est tel que $||u(x)|| = ||u^*(x)||$ pour tout $x \in E$. Montrer alors que $g = u^* \circ u - u \circ u^*$ vérifie $\langle g(x), x \rangle = 0$ pour tout $x \in E$
6.c) En exprimant la relation précédente pour $x + y$ et $x + i\,y$ où $x, y \in E$

quelconques, prouver que $g = 0$.
6.d) Que conclure alors sur u ?
7) Soit A une matrice carrée complexe $n \times n$ de valeurs propres $\lambda_1, \dots, \lambda_n$ (écrites avec éventuelles répétitions).
7.a) Prouver que si on suppose A normale (dans cette sous-question seulement) alors $\mathrm{Tr}(A\,A^*) = |\lambda_1|^2 + \cdots + |\lambda_n|^2$.
7.b) Montrer qu'il existe U matrice unitaire et T triangulaire supérieure de diagonale $\lambda_1, \dots, \lambda_n$ telles que $A = U^{-1}\,T\,U$.
Indication. On pourra d'abord examiner la possibilité de trouver une base où l'endomorphisme associé à A (sur le base canonique) trigonalise ; puis on pourra modifier cette base pour avoir une matrice de passage (depuis la base canonique) qui est unitaire.
7.c) Montrer que si $\mathrm{Tr}(A\,A^*) = |\lambda_1|^2 + \cdots + |\lambda_n|^2$ alors A est normale.
8) Soit H et K deux matrices complexes hermitiennes positives. On veut prouver que

$$0 \leq \mathrm{Tr}(H\,K) \leq \mathrm{Tr}(H)\,\mathrm{Tr}(K) \qquad (\mathbf{1})$$

8.a) Prouver l'inégalité (1) si on suppose en plus que K est diagonale.
8.b) En revenant au cas général, justifier l'existence d'une matrice unitaire U et d'une matrice diagonale D telle que $H = U\,D\,U^*$.
8.c) Comparer $\mathrm{Tr}(H\,K)$ avec $\mathrm{Tr}(D\,L)$ où $L = U^*\,K\,U$.
8.d) Que dire de la matrice L.
8.e) Prouver finalement (1).

Exercice 65 (Examen de septembre 2001) On désigne par $\mathbb{K}$ le corps $\mathbb{R}$ ou $\mathbb{C}$. Soit $\theta \in \mathbb{K}^* = \mathbb{K} \setminus \{0\}$ et $n \in \mathbb{N}^*$. On considère la matrice carrée A ayant $2n$ lignes et $2n$ colonnes dont le terme général a_{ij} vaut $-\theta$ si $i + j = 2n + 1$ et $j \leq n$, vaut θ si $i + j = 2n + 1$ et $j > n$ et vaut 0 sinon. On note u l'endomorphisme de $\mathbb{K}^{2n}$ de matrice A sur la base canonique $(b_1, \dots, b_{2n})$ de $\mathbb{K}^{2n}$.
1.a) Ecrire la matrice A dans le cas $n = 3$.
1.b) Que vaut la trace de A ? (on revient au cas général pour n)
1.c) Calculer le déterminant de A.
1.d) Que vaut ${}^tA$ la transposée de A ?
1.e) Calculer A^2.
2.a) On note B la matrice 2×2 suivante

$$B = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$$

Donner la nature géométrique de cette transformation dans le cas $\theta = 1$ et $\mathbb{K} = \mathbb{R}$.

2.b) Ecrire la matrice de u dans la base $(b_1, b_{2n}, b_2, b_{2n-1}, \ldots, b_k, b_{2n+1-k}, \ldots, b_n, b_{n+1})$ à l'aide du bloc B.
2.c) Le résultat de la question précédente confirme-t-il les calculs des questions 1.c et 1.e ?
3.a) Quel est le polynome minimal de A ?
3.b) Que vaut le polynome caractéristique de A ?
4) On suppose dans cette question que $\mathbb{K} = \mathbb{C}$.
4.a) Prouver que u est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ et préciser les valeurs propres.
4.b) Dans le cas $n = 1$ –c'est à dire que A est simplement la matrice B– donner une base de vecteurs propres de u.
4.c) Donner une base de vecteurs propres de u dans le cas général $n \in \mathbb{N}^*$.
5) On suppose dans cette question que $\mathbb{K} = \mathbb{R}$.
5.a) u est il diagonalisable dans $\mathcal{M}_n(\mathbb{R})$?
5.b) u est il trigonalisable dans $\mathcal{M}_n(\mathbb{R})$?
6) Calculer A^p pour tout $p \in \mathbb{N}^*$.

Désormais T désigne une matrice antisymétrique (${}^tT = -T$) à coefficients réels ayant n lignes et n colonnes.
7.a) Montrer que la matrice $iT \in \mathcal{M}_n(\mathbb{C})$ est hermitienne.
7.b) Rappeler le résultat du cours concernant la réduction des matrices hermitiennes.
7.c) Que peut on en déduire sur le réduction dans $\mathcal{M}_n(\mathbb{C})$ de T ?
7.d) La matrice T est elle diagonalisable dans $\mathcal{M}_n(\mathbb{R})$?
7.e) Montrer que $\det(I + T) \geq 1$.
8) Soit S une matrice à coefficients réels symétrique définie positive ayant n lignes et n colonnes.
8.a) Rappeler ce que le terme "définie positive" signifie.
8.b) Que sait on sur le réduction d'une telle matrice dans $\mathcal{M}_n(\mathbb{R})$?
8.c) Prouver que $\det(S + T) \geq \det S$. On pourra, par exemple, se ramener au résultat prouvé en 7.e.

Exercice 66 (Examen de janvier 2002) Les parties I et II sont indépendantes

PARTIE I

Soit un entier $n \geq 3$. On considère la matrice $A = (a_{ij})_{1 \leq i,j \leq n}$ carrée $n \times n$ dont le terme général a_{ij} vaut 1 si l'un des deux indices i ou j (mais pas les deux) vaut n. Dans tous les autres cas a_{ij} vaut 0. Ainsi A peut s' écrire :

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

I.1) Quel est le rang de A ?
I.2) 0 est il valeur propre ? Si oui, quelle est la dimension de l'espace propre E_0 associé ?
I.3) Pourquoi peut on affirmer sans aucun calcul que A est diagonalisable ?
I.4) Pour $\lambda \neq 0$, donner une forme équivalente simple du système

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

I.5) En déduire que A admet deux valeurs propres non nulles, l'une λ_+ positive et l'autre λ_- négative et déterminer les sous espaces propres E_+ et E_- respectivement associés.
I.6) Calculer la trace de A. Le résultat est il conforme aux résultats précédents concernant les valeurs propres ?
I.7) Quel est le polynome caractéristique de A ?
I.8) Quel est le polynome minimal de A ?
I.9) Jusqu'à la fin de cette partie on suppose *désormais* $n = 4$. On désigne par $(e_1, \ldots, e_4)$ la base canonique. A partir de la base $(e_1 - e_2, e_2 - e_3)$ de E_0 déterminer selon le procédé de Gram-Schmidt une base de E_0 orthonormale pour le produit scalaire euclidien canonique.
I.10) En déduire une matrice P orthogonale telle que ${}^tP\,A\,P$ est diagonale de diagonale $0, 0, \lambda_+, \lambda_-$.
I.11) On note q la forme quadratique sur $\mathbb{R}^4$ dont la matrice sur la base canonique $(e_1, \ldots, e_4)$ de $\mathbb{R}^4$ est A. Que vaut $q(x)$ pour $x = (x_1, \ldots, x_4)$?
I.12) Quelle est la signature de q ?

PARTIE II

On considère une matrice M réelle carrée $n \times n$ inversible.
II.1) Montrer que la matrice $S = {}^tM\,M$ est symétrique définie positive.
II.2) Prouver l'existence d'une matrice orthogonale Q et d'une matrice diagonale D de termes diagonaux $d_1, \ldots, d_n$ vérifiant $d_1 \geq d_2 \geq \cdots \geq d_n > 0$

telles que $S = {}^tQ\, D^2\, Q$.
II.3) Y a-t-il unicité de la matrice D ? Y a-t-il unicité de la matrice Q ?
II.4) On pose alors $R = M\, {}^tQ\, D^{-1}$. Prouver que R est orthogonale.
II.5) Avec l'identification habituelle entre vecteurs de $\mathbb{R}^n$ et matrices colonnes, on note $||\cdot||$ la norme euclidienne canonique : $||X||^2 = {}^tX\, X$. Montrer que

$$\left(\sup_{X\neq 0} \frac{||MX||}{||X||}\right)^2 = \sup_{Y\neq 0} \frac{{}^tY\, D^2\, Y}{{}^tY\, Y} = d_1^2$$

Les questions qui précèdent prouvent donc qu'étant donnée une matrice réelle inversible M, il existe une matrice diagonale D de termes diagonaux strictement positifs et deux matrices orthogonales Q et R telles que $M = R\, D\, Q$.
II.6) On suppose maintenant que $\tilde{M}$ est une matrice carrée *complexe* inversible. Prouver une décomposition $\tilde{M} = \tilde{R}\, \tilde{D}\, \tilde{Q}$ similaire à ce qui précède où $\tilde{Q}$, $\tilde{D}$, $\tilde{R}$ sont de types particuliers à spécifier.

Exercice 67 (Examen de septembre 2002, second des deux exercices) On considère la matrice 4×4 suivante

$$M = \begin{pmatrix} 0 & 0 & 0 & a \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & c \\ a & b & c & d \end{pmatrix}$$

en supposant tout d'abord que a, b, c, d sont 4 réels avec a, b, c non tous nuls.
1) Calculer le polynôme caractéristique de M.
2) Combien M a-t-elle de valeurs propres ?
3) Pour chaque valeurs propres de M, quelle est la dimension de l'espace propre associé ?
4) M est elle diagonalisable ? Pourquoi était il possible de l'affirmer sans aucun calcul ?
5) Quel est le polynôme minimal de M ?
6) Quelle est la signature de la forme quadratique associée à M ?
7) Si a, b, c, d sont maintenant des nombres complexes, M est elle diagonalisable pout toutes valeurs de a, b, c, d?

Exercice 68 (Examen de janvier 2003) On considère la matrice

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

où a est un paramètre réel.
1) Montrer que A est inversible.

On note $\mathcal{C} = (c_1, c_2, c_3)$ la base canonique de $\mathbb{R}^3$. On définit une base $\mathcal{F} = (f_1, f_2, f_3)$ de $\mathbb{R}^3$ comme étant la base telle que la matrice de passage de $\mathcal{C}$ à $\mathcal{F}$ notée $P_{\mathcal{C}\to\mathcal{F}}$ est A. On note $\langle\cdot,\cdot\rangle$ le produit scalaire euclidien canonique sur $\mathbb{R}^3$:

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3)\rangle = \sum_{i=1}^{3} x_i\, y_i$$

et $q(x) = \langle x, x\rangle = ||x||^2$ la forme quadratique associée.
2) Quelle est la matrice de q dans la base $\mathcal{C}$ puis dans la base $\mathcal{F}$?
3) A quoi se réduit le procédé d'orthonormalisation de Gram-Schmidt quand on l'applique à $\mathcal{F}$ pour obtenir une base orthonormale $\mathcal{B} = (b_1, b_2, b_3)$?
4) Quelle relation y a-t-il entre les matrices de passage

$$P_{\mathcal{C}\to\mathcal{F}}\,,\quad P_{\mathcal{B}\to\mathcal{F}}\,,\quad P_{\mathcal{C}\to\mathcal{B}}$$

5) En déduire l'existence d'une matrice orthogonale Ω et d'une matrice diagonale D telle que $A = \Omega\, D$.
6) En déduire A^{-1}.

On suppose **désormais** que $a = 2$ et donc que

$$\Omega = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & 1 & \sqrt{2} \\ 0 & 2 & -\sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \end{pmatrix}$$

7) Rappeler le théorème du cours sur la réduction des matrices orthogonales.
8) Calculer le polynôme caractéristique de Ω. On pourra faire apparaitre la quantité

$$\beta = \frac{2\sqrt{3} + 3\sqrt{2} + 2\sqrt{6}}{6}$$

9) En déduire la forme de la matrice réduite semblable à Ω. Cette matrice est diagonale par blocs et peut en particulier faire apparaître un bloc 2×2 de rotation d'angle $\theta \in]-\pi, \pi]$.
10) Déterminer $|\theta|$ par une formule qui pourra utiliser la fonction Arccos.
11) Déterminer $\ker(\Omega - I)$.